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

PROCEEDINGS OF SESSION I: RISK ASSESSMENT AND VALUATION OF HEALTH EFFECTS FROM AIR POLLUTION (INCLUDING INTRODUCTORY REMARKS)

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

Table of Contents

Introductory Remarks
William H. Farland, Chief Scientist, Office of the Science Advisor, and Acting Deputy Assistant Administrator for Science, Office of Research and Development, U.S. EPA

Session I: Risk Assessment and Valuation of Health Effects From Air Pollution
Session Moderator: Ron Shadbegian, U.S. EPA, National Center for Environmental Economics

Willingness To Pay for Improved Health: A Comparison of Stated and Revealed Preferences Models
Michael Hanemann, University of California–Berkeley, and Sylvia Brandt, University of Massachusetts–Amherst

Individual Preferences and Household Choices: The Potential Role of Dependency Relationships
Mary F. Evans, University of Tennessee–Knoxville; Christine Poulos, Research Triangle Institute; and V. Kerry Smith, North Carolina State University

Preliminary Results From a Daily, Time-Series Study of Air Pollution and Asthma in the San Francisco Bay Area
Charles Griffiths and Nathalie Simon, U.S. EPA, National Center for Environmental Economics

Discussant: Bryan Hubbell, U.S. EPA, Office of Air Quality Planning and Standards

Discussant: Glenn Blomquist, University of Kentucky

Questions and Discussion

U.S. EPA NCER/NCEE Workshop
Morbidity and Mortality: How Do We Value the Risk of Illness and Death?

Washington, DC
April 10-12, 2006

Introductory Remarks
William H. Farland, Chief Scientist, Office of the Science Advisor, and Acting
Deputy Assistant Administrator for Science, Office of Research and Development

This is the 13th of the Economy and Environment Series workshops that have been
sponsored by the Office of Research and Development's National Center for
Environmental Research and the Office of Policy, Economics, and Innovation's National
Center for Environmental Economics. The opportunity here is really to bring together a
group of colleagues to think about some of the issues around approaches to valuation of
human health effects—both mortality and morbidity. These kinds of issues are
particularly important to the Agency.

The group that we expect to be interacting with over the next two-and-a-half days makes
up a broad section of the scientific community and the economic community. Clearly,
our colleagues from the National Center for Environmental Economics will be
represented here as well as our Science To Achieve Results grantees, who have the
opportunity to use Agency resources to explore many of these kinds of issues. We also
expect EPA economists and other scientists from a number of our program offices as well
as ORD and OPEI. In addition, we had a broad solicitation and have a number of
researchers from academic institutions and other federal agencies, and we're particularly
pleased to welcome our international guests for this program.

Again, our purpose here is really to learn more about research that improves our
understanding with regard to the valuation of health outcomes, and to assure that the
research that is done really feeds into opportunities for improving how we make our
decisions. Now, I want to spend just a few minutes on this idea of the importance of the
workshop, because as we look at research results, as we begin to think about the kinds of
data that come out of the work that you all do, whether they're out of the academic or
federal community, these results need to be credible and relevant and timely so they can
inform the kind of policy decisions that EPA and many other federal agencies are making
every day.

Clearly, the value of cost/benefit analysis is very high. If we look at things like the
Executive Orders that are coming out and new legislation like the Safe Drinking Water
amendments that really focus on having this type of information in order to judge the
quality and relevance of particular options for environmental decisions, we see that this
becomes extremely important. It really affects the way we do business now at the
Agency for many of the things that are important to us. Clearly, accurate cost/benefit
analysis leads to better decisions. It allows us to really understand, where appropriate,

the role of economics in terms of environmental decision making, and we're really proud of the good relationship that has developed between the Office of Research and Development and OPEI with regard to these particular issues and advancing the state of the science.

So, as I said, the idea of human health valuation is a very important topic for the Agency. Clearly, it drives much of what we do. Particularly, as you begin to look at the issue of morbidity and mortality, we realize that our community has recognized that getting a handle on approaches to assessing morbidity is really a top priority for the field, and it strongly highlighted in the most recent environmental economics research strategy that the Agency has put forward. The challenge of trying to deal with valuation of morbidity is one that we'll talk a little bit more about, but it clearly is something that we need to put our collective best minds toward as we begin to think about how we approach it.

As you know, mortality valuation is always high profile and plays a very, very large role in many of the decisions that we make. It is the one area where we have made some significant progress in valuation. For some of the most recent Clean Air Act decisions, for instance the particulate matter decisions that have led to rules recently, more than 90 percent of the monetized benefits come out of these mortality valuations. So, again, we recognize the importance of mortality in trying to get the approaches to mortality valuation right, but we also realize the issue of how we approach morbidity and the very strong role that non-monetized benefits currently play and the importance of trying to value those.

So, over the next two-and-a-half days you'll have an opportunity to deal with a number of very important issues. Some of these are old favorites—if we think about asthma, pollutants in drinking water, lead paint and IQ loss, PCBs, and so on. Those are issues that we've all been dealing with for quite a while, but the field is really opening up for us. There are opportunities here for us to really emphasize things like children's health valuation. As you begin to think about that, think about it from the standpoint of children being differentially susceptible, in many cases, to some of the pollutants that we're dealing with. So, clearly their life stage and their lifestyle—their behavior—is very different from that of adults. At the same time, children are not full economic actors, if you will, so trying to look at the valuation of children's health impacts and even dealing with the issue of early life exposures leading to later life impacts becomes a real challenge for us. Clearly there are opportunities here for methodologic advances in valuing health risk reductions and modeling household decision making—again, a challenge for us as we think about the fact that our households are changing and the types of situations that we're looking at now are particularly important.

We're looking for an opportunity to deal with the issue of updates on cost-effectiveness analysis, mortality valuation, and efforts to include economic questions in our large-scale health surveys. Clearly, this is an integration of the field of monitoring and modeling health effects with the economic valuation of those effects. We are expecting to have a real opportunity to move forward.

One of the things that we're intending to do later on in the session is a panel discussion on the pros and cons of web surveys. Many of you are building web surveys into your protocols, and in the most recent OMB guidance there has been some concern raised about the use of web surveys. This is an approach that is heavily in use, and it's something that we need to look at very carefully so we understand the pros and cons.

Finally, at the end of the session, during the last half day, is something that we've not done before. It is an opportunity to focus, in depth, on a particular set of research results from a single grant that EPA has funded. This is a grant that has gone to UCLA and Oregon State investigators, and we're going to be very interested in your feedback on this particular approach for the workshop. We will focus in on a particular set of results and really have a half-day, in-depth discussion on that.

So, clearly, the research results that we are going to be talking about are going to be very important, and they are currently being used by the Agency. We're looking for opportunities to do things even better than we have in the past. Clearly, the record shows that previous results have been used in important analyses—our Section 812 report in the Clean Air Act, which actually lays out costs and benefits of Clean Air Act decision making, has illustrated the way that the Agency has been successful in laying out the economic benefits of the Clean Air Act. It is actually among the leading regulatory actions with regard to monetized environmental benefits of any of the actions that we take.

The results of our research have been cited by OMB in their guidance on mortality valuation. Some of you may remember the discussions that came forward on exactly how we were going to value mortality for the elderly and the advice that we got from OMB with regard to the so-called "senior death discount" and not discounting the issues with mortality later in life.

Certainly we expect the future to hold expanded use of these techniques, and we're looking forward not only to improving the approaches that we use for this, but also to being able to demonstrate results—to deal with things like our program evaluation ratings and other opportunities that we have to demonstrate how research is used to inform decision making and how those decisions can lead to improved environmental results.

So, looking toward the future, what research will we be looking to fund and support? Clearly, some of that will depend on the kinds of discussions that we have over the next two-and-a-half days. A lot of these issues have been foreshadowed in the Environmental Economics Research Strategy, and this provides a good opportunity for us to work within a framework of important research needs. Some of the things that are clearly going to be part of that research that we fund have to do with the fact that they will be results-oriented types of work; they will be things that we can apply routinely; they will focus on issues such as the question of benefit transfer, which is something that is particularly important as you get into understanding particular situations and applying that to a broader population or a broader situation. Another important issue will be the question of marginal risk changes, so that we can really get at the question of how we monetize over

time with changes that occur with regulatory activities. It's very clear that the field is moving toward an inter-disciplinary approach, much like many of the fields that we interact with at EPA. We're looking forward to meetings like this one to really have an opportunity to hear from various disciplines that have a role to play in the important research that's going on.

With that I'll close and wish you well in terms of the research discussions that will occur over the next two-and-a-half days. I'd like to thank Will and the other organizers from both of the offices who co-sponsored this. I look forward to a very successful workshop. Thanks.

Willingness to Pay for Improved Health: A Comparison of Stated and Revealed Preferences Models

W. Michael Hanemann
Department of Agricultural & Resource Economics
University of California, Berkeley
hanemann@are.berkeley.edu

Sylvia Brandt
Department of Resource Economics
University of Massachusetts, Amherst
brandt@resecon.umass.edu

Abstract:

In this paper we discuss two approaches to estimating the willingness to pay (WTP) for reduced asthma morbidity, contingent valuation and health production function. The study population includes 250 children ages 5-11 with clinically diagnosed asthma, residing in a section of Fresno County, California. Asthma symptoms, including coughing, wheezing and/or shortness of breath, ranged from mild and intermittent to severe and persistent in this group. Detailed health measures (including atopy and pulmonary function), utilization of health services, levels of antigens in the households and exposures to criteria air pollutants were collected as part of a five-year epidemiological study. We administered two economic surveys to measure 1) households' perceptions of risks to an asthmatic child, 2) averting and/or mitigating actions taken, and 3) households' stated willingness-to-pay for a reduction in their children's asthma morbidity.

In the health production model the health outcome is a function of exposure to asthma triggers, mitigating and averting behavior and household's perceived risks. We find that variation in WTP is explained by attitudes towards asthma specific health investments including concerns of associated risks and perceived effectiveness. The survey data indicate that households select from a small number of discrete health investments and that most risk reducing behavior are daily behavioral modifications with no relevant market prices.

We argue that the discrete nature of health investments and socio-cultural patterns of health care utilization make the revealed preference approach inadequate for the case of asthma. As an alternative we present a contingent valuation scenario that was specifically developed to minimize systematic variation in preferences for characteristics related to the scenario rather than the reduction in asthma morbidity. For this purpose, guided by extensive testing in focus groups, we selected a scenario based on a hypothetical asthma monitor that provides to the wearer an indicator of current asthma status.

INTRODUCTION

The economic concept of value implies a tradeoff. The monetary value of any item is defined in economics as the amount of money that a decision-maker – an individual, a household, or a firm, depending on the context – would be willing to exchange for the item. That monetary amount measures the worth of the item in monetary units in the sense that the exchange of this monetary amount has the same impact on the decision maker's wellbeing (utility) as the item itself.[1] The challenge for economic measurement is to identify a trade-off through which value can be measured. Revealed preference approaches work by observing actual choices by decision-makers and inferring the trade-off underlying these choices. Depending upon the nature of the choice (whether it is a discrete, continuous, or mixed discrete/continuous choice) the choice behavior may reveal the trade-off either directly (a simple discrete choice) or indirectly (the cases involving continuous choices) by permitting the identification of an underlying set of preferences which had motivated the observed choice behavior. In the latter case, the trade-off is inferred from the recovered preferences underlying the observed choice rather than directly from the observed choice itself. Stated preference approaches work by placing subjects in a survey or experimental setting and confronting them with choices that, directly or indirectly, reveal their preferences.

In the context of valuing health outcomes, the standard revealed preference approach assumes that health-related choice behavior reflects preferences for health outcomes that are generated by a perceived health "technology". This separation between preferences and production requires the researcher to differentiate between behavior that is an end in itself and behavior that is a means to an end. Consider, for example, assessing the value of good water quality at a beach from this perspective. In the case of amenity value, an individual's choice of which beach to visit (trading off cleaner but more distant beaches versus dirtier but closer beaches) bears directly on the trade-off of interest since going to a nice beach is presumably an end in itself. In the case of health outcomes, an individual's choice of which precautions to take (spending money to purchase goggles, taking an antibiotic before going surfing, etc) is a means to an end – namely, good health – rather than an end in itself from which the individual derives enjoyment per se. In the latter case, the valuation analyst has to disentangle the production component from the pure preference component that underlies the sought-after trade-off. We suggest that this complication may sometimes tilt the balance in favor of stated preference rather than revealed preference as the preferred valuation approach[2].

---

[1] Generically, there are two ways to formulate the exchange: the maximum amount that the individual would be willing to pay (WTP) to obtain the item, if it is favorable, or to avoid it, if unfavorable; and the minimum amount of money that the individual would accept (WTA) to forego the item, if it is favorable, or to endure it, if unfavorable. The relationship between WTP and WTA is a separate issue that will not be pursued here. For simplicity, the discussion below focuses on the WTP measure of welfare.

[2] An important consideration in modeling health outcomes for children is the question of the identity of the decision maker. The decision maker is surely *not* the child but rather one or both of the parents; therefore, the framework is the household rather than individual decision making. Making the household the unit of analysis raises several important but difficult analytical issues that are addressed in other literature. In this paper we focus on the relationship between health preference function and health production function, and we make the simplifying assumption that household decisions regarding children's health reflect a unitary model of household preference and production.

We examine the application of the revealed preference and stated preference approaches to the valuation of reduced asthma morbidity. Our economic study was done in collaboration with an epidemiological study that was the most detailed socio-demographic, indoor air quality and pollution monitoring data collection effort to date (California Air Resources Board). Findings from multiple focus groups and two economic surveys suggest that the discrete nature of health investments and socio-cultural patterns of health care utilization make the revealed preference approach inadequate for the case of asthma. As an alternative we present a contingent valuation scenario that was specifically developed to minimize systematic variation in preferences for characteristics related to the scenario rather than the reduction in asthma morbidity.

This paper is organized as follows. In Section One we describe the epidemiological study and economic surveys used to collect household level data. Second, we summarize the average households expenditures related to asthma morbidity and conceptual limitations to using these costs as a measure of value. In the third section we present the standard household health production model. Conceptual limitations to the standard model are presented in the fourth section. Fifth, we present empirical evidence of these complexities and their implications for the household production model. Next we discuss how we used the findings from the first economic survey to create a contingent valuation scenario. Concluding remarks are included last.

1. Empirical Study

A. Study Setting

This project is a collaboration with an extensive epidemiological study of the effects of air pollution on asthmatic children [Fresno Asthmatic Children's Environment Study, FACES]. The study is located in Fresno, California, which has highest rate of asthma hospitalizations in California at 28.8 per 10,000 (California Facts, 2003). Located in the Central Valley of California, Fresno County has a population of 815,734 and this population has increased by 19.8% since 1990. Forty-four percent of the population is of Hispanic or Latin origin, followed by forty percent of white origin, eight percent Asian and five percent African-American. The Fresno population has lower median income, less education, poorer living conditions and a greater percent of residents below the poverty line as compared to the rest of CA. For example, median household income for 2001 was $34,725 as compared to $47,493   for California. The proportion of residents with a high school degree was 67.5% as compared to 76.8% for the rest of the state, and the proportion of residents below the poverty line was 22.9 % while that in CA was 14.2% (US Census data, 2000).

The FACES cohort included children with clinically diagnosed asthma, residing in a section of Fresno County, California[3]. Children were 6-10 years of age at intake and were followed for approximately 4 years. The study population included children who had a physician's diagnosis of asthma and at least one of the following: 1) reported utilization of or valid prescription for asthma medication in the previous 12 months; or 2) symptoms consistent with asthma in the past 12 months; or 3) an emergent asthma visit or hospitalization in the past 12 months. The requirements for asthma medication use, symptoms, or health care utilization are to minimize the chance of enrolling subjects whose asthma is quiescent (remission). Children who meet these

---

[3] FACES has been recruiting households for the survey since 2000.

criteria may be enrolled regardless of the severity of asthma. Children with major comorbidities that would confound the measurement of pulmonary function were excluded.

The FACES study screened 473 households, completed baseline interviews for 241 households, and retained 205 participating households. The major reasons households who inquired about the study were ineligible to participate include: other chronic disease, lived in house for less than three months, child sleeps at home less than five nights/week, and family planned to move within two years (Mann, 2003).

Demographics and characteristics of the FACES cohort are in tables 1. The percentage of blacks enrolled in the FACES program (13.7%) is greater than the percentage of blacks for the Fresno population (5.3%), while Asian Americans are underrepresented. The average age of children in the FACES cohort is between eight and nine years. The majority of the interviewed households were covered by health insurance (90.3%). Almost 70% households had at least one parent who was affected by asthma. One observable characteristic of the FACES cohort that differs from the Fresno population is the frequency of smoking in the home.

Table 1: Demographics of FACES Cohort

Demographics	Relative Frequency
Race	
White	40.7
Hispanic	41.8
African American	13.7
Other	3.8
Male	58
At least one parent employed	90
At least one parent completed high school	88
Participant's health status	
Ever prescribed oral steroids	61
One or more hospitalization(s)	25
Positive skin test to at least one antigen	65
FEV1	1.06 (range, 0.51-1.59)
Any smokers in home	23%
Asthma severity	
step 1	27
step 2	46
step 3	24
step 4	3

Note: Based on baseline interviews completed as of June 30, 2002 (n=182). Severity scores based on the NHLBI guidelines.

B. Economic Surveys

Two economic surveys were conducted in the FACES cohort. The first survey, a written mail survey was conducted February to August 2004. This survey included detailed questions on asthma related expenditures, asthma related symptoms and activity limitations, and health

beliefs. A total of 202 households completed the first survey (representing 209 children with asthma). The second survey contained a contingent valuation scenario and was conducted October 2005-February 2006. The purpose of completing two surveys was to explore the strengths and limitations of the two approaches to valuing children's health: stated preferences and household health model. In this paper we present the findings from the first survey to motivate our design for the contingent valuation scenario.

The median household size those completing for survey one was 4, (range: 2-9), and 41% of the households had two children under the age of 18. The survey respondent was typically the household member who interacted with the healthcare provider: 95% responded that they are the ones to take children to medical appointments. Employment status of the respondent varied: 37% were employed full-time, 27% were employed part-time, 11% were not employed but were looking, and 24% were not employed outside the home and were not looking for employment. The distribution of household income for participants that completed the first economic survey is reported in Table 2.

Table 2: Household Income

Household income	Relative Frequency
Less than $10,000	11.5
$10,000 to less than $20,000	11.5
$20,000 to less than $30,000	8.2
$30,000 to less than $40,000	15.9
$40,000 to less than $50,000	11.5
$50,000 to less than $75,000	19.2
$75,000 to less than $100,000	12.6
$100,000 or more	9.3

Note: Total responses =182

Tables 3 and 4 describe the asthma status participants who completed the first economic survey. Following the GINA recommendations, asthma severity was based on frequency of daytime and nighttime symptoms. Less than 20% of the children had severe day or night symptoms. More children had moderate symptoms in the day than at night (43% versus 35%), and consequently more children had mild symptoms at night than during the day (46% versus 38%). There was a high degree of correlation between day and nighttime severity.

Table 3: Count of Asthma Severity by Day and Night Symptoms

Daytime symptoms	Nighttime Symptoms			Total
	Mild	Moderate	Severe	
Mild	64	11	1	76
Moderate	26	53	8	87
Severe	2	6	31	39
Total	92	70	40	202

Note: Mild is defined as symptoms less than two times a week. Moderate is defined as symptoms 3-5 times a week. Severe is defined as symptoms every day.

The majority of respondents had prescriptions for a rescue and controller medication.

Table 4: Frequency of Medication Usage

	Percentage who have prescriptions for:	
Number of Medications	**Control Medication**	**Rescue Medication**
None	13	14
1	15	44
2	25	38
3 or more	47	4

Of the survey respondents 79% reported that no one smokes in the home: of respondents that reported smoking in the home 58% reported that the father smoked, 28% reported maternal smoking and 20% reported that another adult smokes in the home. These smoking rates are below that for the Fresno population, but are inline with those for the FACES cohort.

2. Costs related to Asthma

A. Health expenditures

Direct expenditures on asthma were broken into four categories: fixed costs[4], household supplies, pharmaceuticals (prescription and over-the-counter) and alternative therapies. Variable costs are the sum of supplies, pharmaceuticals and alternative therapies.

Table 5: Asthma Related Expenditures

	Median	Mean	Standard deviation	N
Fixed costs	240	357	716	202
Variable costs	110	139	114	199
Household supplies	49	81	86	202
Pharmaceuticals	37	53	61	202
Alternative therapies	0	4	22	199

Note: Fixed costs included service costs or purchases of: air filters, allergy mattress covers and/or pillow covers, humidifiers, dehumidifiers, air conditioners, HEPA vacuum cleaner, landscaping of yard, carpet removal, pest extermination, mold/mildew removal, removing pets, humidity gauge, nebulizers, peak flow meters, spacers for inhalers, replacement for window coverings, and fans. Household supplies include: replacement filters for air filters, filters for air conditioners, HEPA vacuum filters, heater filters, cleansers for mold and mildew, hypoallergenic or non-aromatic cleaners, allergen control sprays, sprays for pest removal. Pharmaceutical costs included: prescription asthma or allergy medication, over the counter allergy or asthma medications, herbal remedies, and home remedies. Costs related to alternative therapies include visits to: chiropractor, acupuncturist, doctor of osteopathy, homeopathic/herbalist, nutritionist, spiritual healer, or other alternative health care provider.

Indirect costs related to asthma morbidity include employment impacts and time used for planned or unplanned medical visits (including in-clinic and emergency room). Of the 202

---

[4] For fixed investments we delineated 1) purchased specifically to help asthma 2) purchased prior to asthma diagnosis 3) never purchased.

households, 43% reported that they usually needed to take time off from work to take their child for medical appointments, and the median time taken off from work for each medical appointment was 75 minutes (mean 83 and standard deviation 37 minutes). Twenty- four percent of the families reported that they had gone to the emergency room for the child's asthma in the previous 12 months, requiring a median of 215 minutes per visit (mean 211 and standard deviation of 128)[5]. Although lost work time due to asthma is one indirect cost of asthma morbidity, 18% of the households reported changes to their employment due to the frequency or severity of their child's asthma. Of those that had changed their employment status: 28% had been subjected to employers reducing their work hours due to previously missing work for child's illness; 22% had been fired or laid-off due to missing work for child's asthma; 38% had chosen to not seek employment outside the home due to child's asthma; 32% chose to work only part-time due to asthma; 27% worked fewer hours during asthma seasons. These statistics suggest that limiting indirect costs to workdays lost underestimates the impact of asthma morbidity on household income.

B. Limitations to Costs as a Measure of Value

There are three reasons why using expenditures and indirect costs as measures of value of reducing asthma morbidity has shortcomings. The first is that there is an important difference between the concepts of cost and value, the latter being the concept of interest to economists. The second issue is the critical distinction between marginal and non-marginal value. The last limitation is that these cost miss larger impacts of asthma morbidity on households. We discuss each of these in this section.

Economists have been aware of the fundamental distinction between *what things cost* and *what they are worth* ever since Adam Smith exposed the diamonds-water paradox (water is essential for life but inexpensive, while diamonds are entirely inessential but extremely expensive). Both may be important to know, but they are different things. What something costs is a question about *supply*; what it is worth is a question about *demand*. What something costs is *objective* and a matter of production and engineering; what it is worth is *subjective* and a matter of preference and taste.

The distinction between marginal and *total* value also underlies the diamond-water paradox is. *At the margin,* an additional kilo of water may have a lower value to people than an additional kilo of diamonds because water is abundant while diamonds are rare, but the *total* value of all water to mankind is likely to be greater than the total value of all diamonds: if we lost all access to water, people would surely judge this a greater harm than losing all access to diamonds. Similarly, Dupuit emphasized the distinction between marginal and *infra-marginal* value. Dupuit observed that, for any consumer and any commodity, while the last unit of an item to be consumed would be just equal in value to its price (which is why it is the last unit to be consumed), the infra-marginal units would be worth *more* than the price because of the phenomenon of diminishing marginal utility. For the infra-marginal units, the consumer would be willing to pay more than the price. Hence, the consumer receives a sort of "profit" or "surplus" on these units which he would lose if the item were not available at that price. This observation provides the foundation for Dupuit's concept of consumer's surplus – the excess of

---

[5] Total time spent on medical visits and emergency room visits will be reported in future work.

what a consumer would be willing to pay for an item over and above what he actually does pay. If the consumer had valued all units of the item exactly at their price his consumer's surplus would be identically zero, but this does not generally happen. Dupuit's larger point was that value is measured by reference to the consumer's demand curve – by what we now call the Marshallian consumer's surplus[6].
.

The third limitation of the expenditure approach is that beyond the indirect and direct costs there are psychosocial impacts on the households from asthma morbidity. These impacts include changing family activities, interactions with peers and the burden of uncertainty surrounding the status of a child's asthma. For example, in focus groups parents frequently discussed difficulties in communicating their child's needs to school officials and to physical education teachers in particular. In our survey 21% (out of 170 responding) disagreed with the statement "My child's classroom teachers are helpful with my child's asthma needs." and 24% (out of 140 reporting) disagreed with the statement "The physical education teacher works with us to include my child." Other impacts were restrictions on normal childhood play: 38% of respondents (n=200) reported that they restricted the *amount of child's activity* more often specifically due to asthma; 46% of respondents (n=201) reported that they restricted the *amount of time outside* more often specifically due to asthma; 44% of respondents (n=201) reported that they restricted the *where the child could play or visit* more often specifically due to asthma. A more dramatic, though less frequently reported change (11 households out of 199 reporting), was moving to a new home to avoid asthma triggers and to improve the child's asthma. The frequency of these impacts and extent to which they affect quality of life suggests that using expenditures to measure value of reduced morbidity misses the complexities of how asthma affects household behavior.

In short, economic valuation is generically about what things are worth, not what they cost. While costs may provide some information about how households value health, costs alone are not adequate measures. In the next section we present a standard household health production model and discuss some conceptual challenges in applying it to valuing children's health.

3. Revealed Preference Approach to Valuing Reduced Asthma Morbidity

A. Health Production Function
We begin with a standard model for household health and to then proceed to show how it is a special case of the Lancaster-Maler utility model. The critical characteristic of the indirect utility function of the household health production model is that it has the same structure of the indirect utility function produced by the Lancaster-Maler utility model, and hence the implications of the Lancaster-Maler for welfare measurement are applicable to the case of the health model. We begin by developing the indirect utility function for the health model.

In the standard model, marketed commodities are divided into two groups, those which have some relation to health (z) – either in preventing ill health or in curing illness once it occurs – and those which have no relation to health (x). The corresponding price vectors are denoted p_x

---

[6] Hicks (1941, 1943) formalized Dupuit's and Marshall's concepts of the difference between the value of infra-marginal units price and their concept of consumer' surplus. Hicks formalized what Dupuit and Marshall asserted in terms of what he called the compensating and equivalent variation.

and p_z with individual elements denoted p_i and p_j, respectively. One could further subdivide the health related market consumption activities into those which promote good health and prevent illness (e.g., taking asthma control medication regularly), z_A, sometimes called averting behaviors, and those which reduce the adverse effects of falling ill (e.g. taking an asthma rescue medication), z_M, sometimes called mitigating behavior, so that $z = (z_A, z_M)$. For our present purpose, we can just work with the vector z. Health status could be a scalar or vector of health states or outcomes but, for simplicity, we will treat H as a scalar here.[7] Finally, q is some measure of environmental pollution that affects health. Thus, for the household there is a health production function given by:

$$H = H(z, q)$$

where

z is a vector composed of averting behaviors (z_a) and mitigating behaviors (z_m).

There are several alternative formulations of the household's preferences, depending on what enters the household's utility function. Obviously, household health (H) and the consumption of non-health-related market commodities (x) enter the utility function. The question is whether any of the elements of z and/or q enter the utility function as well. The point is that, while z and q affect household utility *indirectly* through their influence on health/illness, H, they could also affect household utility directly if it cares about q or z for motives unconnected with their effect on H. The empirical evidence from our surveys suggests that both z and x are important elements of the household's utility. Thus we will use the most general case is where all of the variables affect household utility directly, and the household maximizes utility subject to the health production function and a budget constraint (Y= income):

$$Max_{x,z} \quad U = U(x, z, q, H)$$

$$\text{subject to } H = H(z, q) \quad \text{and} \quad p_i x_i + p_j z_j = Y$$

The result is a set of ordinary demand functions for all market goods, both non-health-related and health-related, $x_i = x_i(p_x, p_z, q, Y)$ and $z_j = z_j(p_x, p_z, q, Y)$ and a corresponding indirect utility function $v(p_x, p_z, q, y)$.

If we compare the indirect utility function produced in the health model $v(p_x, p_z, q, y)$ to that in the generalized Lancaster-Maler utility model, $v(p, q, y)$, we can see that the former is a special case of the latter in which prices have been partition into non-health related good and health-related goods. The difference between the Lancaster-Maler model and the standard household health production model is simply that the household health production model makes the health production function explicit and implies that the production function can be estimated separately from the pure preferences represented by $U(x, z, q, H)$.

---

[7] In this highly simplified version of a unitary model we are not bothering to distinguish between the health or illness of the different members of the household.

B. Welfare Measurement with the Household Production Model

Recall that within the Lancaster-Maler framework, a consumer's utility depends not only on his consumption of market commodities, denoted by the vector x, but also on some other items, q; the utility function is thus u(x,q). While the consumer controls the level of x, subject to his budget constraint, q represents some things that affect the person's welfare but which he does not control. The Generalized Lancaster-Maler model provides both a theory of how q affects the consumer's choice of market commodities (x) and a framework for welfare evaluation of changes in q. The specific implication for purposes of valuing morbidity in children is that the Hicksian compensating and equivalent variation are expressed in terms of the indirect utility function. In the most general case, all elements (p',q',y') can change to a new level (p'',q'',y'') and indirect utility can change from v(p',q',y') to v(p'',q'',y''). Then the compensating variation for this change is the quantity C such that

$$v(p'',q'',y''- C) \ = \ y(p',q',y'),$$

while the equivalent variation is the quantity E such that

$$v(p'',q'',y'') \ = \ v(p',q',y'+ E).$$

If the change is an improvement in the sense that u''> u', the quantity C measure the consumer's willingness to pay (WTP) for the securing the change, while E measures her willingness to accept to forego it, and vice versa if the change entails a reduction in utility. We can use the concepts of compensating and equivalent variation as measures of the economic value of a change in environmental health risks for children.

Consider two important polar cases with regard to the impacts of the change in environmental health risks: A) The change in environmental health risks could simply and automatically trigger a reduction in the family's disposable income, but with no other concurrent effect, so that the change is from (p,q,y') to (p,q,y''). (B) The change in environmental health risks could simply trigger a change in q, with no other concurrent effect on p or y, so that the change is from (p,q',y) to (p,q'',y). In the first case, the direct effect of the environmental change is that the household has less disposable income but everything else remains the same: the impact is equivalent to a lump-sum reduction in income. The only impact is a purely monetary loss and the economic value of this is the monetary loss itself. In the second case, by contrast, the direct effect of the change is a loss of utility – wellbeing – for the household, C and E are different, and they represent alternative ways of expressing this loss monetarily in terms of a loss of income that is equivalent in the magnitude of its impact on the household's wellbeing.

The practical implication of the distinction between (A) and (B) is that, in the first case, one can get along with information on the magnitude of the monetary loss without necessarily knowing anything about the structure of household preferences, u(x,q), while, in the second case, one cannot avoid the need to know about household preferences. In that case, the comparison between revealed- and stated-preference approaches to welfare measurement will hinge on the relative ease and reliability of the two approaches in providing an insight into the structure of household preferences. We argue below that characterizing household preferences is essential to defendable welfare measurement and present the limitations of using the household health

production model, as typically applied, to measure welfare changes. In the next section we present two areas in which these limitations arise:

- Use of the health cost function to estimate the value of a change in health due to a change in pollution.
- Validity of a production function for health.

4. Conceptual Limitations of the Household Production Model for Welfare Measurement

A. Health Cost Function
First we describe a common use of the health cost function to estimate the value of a change in health due to a change in pollution. For purposes of illustration, we assume that there are no changes in the price of any market goods (p_x, p_z) or income (Y), and environmental quality changes from q_0 to q_1. This scenario is equivalent to Case B described above, and here we describe the limitations to the standard approach to estimating a welfare measure in this case. Suppose the change is for the worse, so that $U_0 \equiv v(p_x, p_z, q_o, Y) \geq U_1 \equiv v(p_x, p_z, q_1, Y)$. In this case the equivalent variation measure (denoted E above) is the household's willingness to pay to *avoid* the change. The marginal WTP to avoid is given by:

$$\frac{dE}{dq}\bigg|_{E=0} = \frac{v_q(p_x, p_z, q, Y)}{v_y(p_x, p_z, q, Y)}.$$

Moreover, by suitable manipulation of the first first-order conditions for the solution to the household's maximization problem, one obtains

$$\frac{v_q(p_x, p_z, q, Y)}{v_y(p_x, p_z, q, Y)} = \frac{u_q(x, z, q, H)}{v_y(p_x, p_z, q, Y)} + \underbrace{c_q(p_z, q, H)}_{\substack{\text{derivative of health cost} \\ \text{function w.r.t } q}}$$

Recall that we presented the most general utility specification in which environmental quality, q, entered directly into the household preference function. If instead we restricted environmental quality to entering the health production function only, in which case the preference function takes the form U(x,z,H), then the first term above would drop out. Then under this special case the expression becomes:

$$\frac{v_q(p_x, p_z, q, Y)}{v_y(p_x, p_z, q, Y)} = c_q(p_z, q, H)$$

This simplification has given rise to following pragmatic approach to measuring the marginal value of pollution in a household production context: (1) Estimate the household health production function, H(z,q). (2) From the health production function derive the corresponding

health cost function, $c(p_z, q, H)$. (3) Given the health cost function, calculate the marginal cost of pollution, $c_q(p_z, q, H)$ and assess the value of the given change in pollution, dq, as the product

$$\text{Value of health damage} = c_q(p_z, q, H)\, dq \, .$$

An attractive feature of the expression above for researchers who use it is that it is only requires information derived from the health production function and it avoids the need to use information about the household's utility function. We believe there are two major flaws in this approach. First, as discussed in the second section, the appropriate measure for valuing welfare change reflects the difference between the margin and infra-marginal unit. The expression above considers only the market-clearing price not the Marshallian consumer's surplus. Our second concern with this approach is the assumption that neither q nor z enters the utility function directly: The construction of the utility function as $U(x, H)$, which omits both the health inputs and environmental quality seems contrary to observations about household preferences[8].

In Section Five we present empirical evidence of three ways in which health averting/mitigating behaviors are central to the concept of the preference function.

B. Validity of a Production Function for Health

While the notion of a health production function is illustrative in the discussion of household choice, the extent to which it captures the complexity of trade-offs in the household is questionable. The conceptual concerns regard the deviation between objective and subjective risk assessments and the degree to which health is determined by individual choice.

In the literature on revealed preference valuation of market commodities based on their attributes, researchers have often found that here is a divergence between the objective measures of attributes and people's perceptions of them. Whether people see a beach as clean, an automobile as safe or comfortable, a computer as high-tech looking, say, is a matter of

---

[8] The household production literature often makes reference to an approach to welfare measurement derived from work by Bockstael and McConnell (1983) based on the demand function for z's or I. Bockstael and McConnell *do* permit q to enter the utility function directly, but not the z's. they show that that the Hicksian measure of WTP for a change in q can be measured exactly from information about the demand function for health, H, or the demand function for one or more of the z's that are input to the production of health. There are two qualifications that are critical to the application of their result to the health production context. The first qualification is that their result is about the area under the *compensated* demand function for H or for the z's, not the ordinary demand function. If there are income effects in the demand for H or for the z's, the two demand functions are different and it is not valid to use the area under the ordinary demand function as an approximation to the area under the compensated demand function. These areas involve a price change from the current "price" (marginal cost) of H to the cut-off price at which the demand for H would become zero, which is by no means a marginal change. Hanemann (1980) showed in an analogous situation that the difference in areas can be quite substantial. The second qualification is that preferences satisfy Maler's (1971, 1974) property of weak complementarity with respect to either H or the z's. In this context, weak complementarity implies , that, if a person is in poor health ($I = 0$), she is indifferent to a change in air quality (q). That seems unlikely to be true. Indeed, one can imagine circumstances under which, as long as the person is still alive, a worsening in air quality becomes more serious to her when she is ill ($I = 0$) than when she is healthy ($I > 0$). Bockstael and McConnell were thinking of a household production function for recreation, not health, when they wrote their paper about weak complementarity and the fact is that their analysis seems ill-suited to health applications.

perception. How people see these attributes can be quite different from how an expert would assess them. But, people's choices are likely to be based on their own perception and understanding of the attributes, not on those of the experts. Therefore, researchers often find that, to model choice behavior successfully, they need to elicit the decision makers' subjective perceptions of attributes involved in the choices. The same can be true of household decision making on health production. What may matter is what the household – the parents – see as efficacious courses of action, not what the medical experts or the econometricians determine to be efficacious. Another way of making the same point is to suggest that, while households' decisions are based on ex ante expectations of the effectiveness of health producing actions, what the econometrician measures when fitting a health production function is the ex post outcome. If there were perfect knowledge or rational expectations, the ex ante expectation and the ex post outcome would coincide. To the extent that these conditions are not met, the ex post household production estimated by the econometrician might be misleading as a guide to understanding household choice behavior. If this is so, it has the potential to bias not only the estimation of the production function but also the estimation of household preferences.

The concept of a household production function implies that the household exercises a degree of control over its member's health that is exaggerated and unrealistic from at least two perspectives. First, postulating a household production function $H = H(z,q)$ implies that, for given q, the household can in principle attain *any* desired level of health, H, providing it has sufficient financial resources to cover the cost of the required z's. If it is rich enough to purchase sufficient z's it can make itself as healthy as it wants, regardless of what might befall it in terms of q. From introspection, this notion is implausible. Second, the notion of an interior solution to the household's health production decision is unrealistic. It is conventionally assumed that, in the context of the household's production function, the z's are finely divisible, so that the household arrives at an exact, interior solution to its optimization decision. Households often face a limited and constrained set of options. These constraints may be imposed by the structure of the healthcare sector and the nature of averting/mitigating behaviors.

Thus, while the notion of the household's production of its own health certainly has some basis in reality, it can be pushed too far. People can look after their own health, but this does not mean that they can achieve *any* desired health outcome; therefore some levels of H are not attainable, regardless of the input of z's. The production function $H(z,q)$ is likely to be bounded and it may have some flat segments. Similarly, people do not have an unlimited array of options and therefore they are more likely to be at corner solutions than interior solutions in their household production decisions. As Bartik (1988) has noted: "Defensive options often may be limited; for example, a household seeking to reduce the effects of toxic waste on its water supply might be able to defend itself only by a water filter, bottled water, or moving away."

If these doubts are justified, this can have important implications for health valuation. The usual first-order conditions do not hold and the simple approximations are apt to be unreliable. The household's marginal WTP for improved health might be considerably larger than the marginal cost, c_q, but it may have no viable option for further action. Also, in this case non-marginal valuation can become more complicated because of the need to specify a realistic, non-monotonically increasing health production function or a limited choice set with a few discrete alternatives.

We suggest that these concerns lead to five practical limitations of the approach: 1) defining the health outcome, especially for chronic, episodic conditions such as asthma, is nontrivial 2) households' have varying degree of control for relevant health inputs 3) there is a probable divergence between an objective physiological health production model and the household subjective perception 5) there is likely to be endogeneity between choice of z and H. We present evidence for each of these in Section Five.

5. Empirical Evidence of the Limitations of the Household Production Model for Welfare Measurement

A. Health Cost Function

In Section 4.A, we suggest that simplifying the household model such that environmental quality, q, and health inputs, z, appear only in the health production function and not the utility function, does not capture the complexities of household tradeoffs. In this section we present three examples of ways in which the averting and mitigating behaviors play important roles in the household preference function.

1. Preferences for Health Inputs

Although the majority of the households reported that their children took medications to treat asthma, 30% reported that they had concerns about those medications. A surprisingly common concern was that taking asthma medication as prescribed could lead a child to become addicted or dependent to asthma medication (49% agree or strongly agree, n=187). Other households reported that they believed that having to take medications regularly was embarrassing to children (23% agree or strongly agree, n=189). These concerns over medications affect how a household perceives their benefit and omitting their consideration distorts the model of household choice. In addition to these general concerns, households reported both that their child experienced negative side-effects from asthma medications as well as a belief that these drugs presented a risk to children's health. Table 6 reports the frequency of that household reported their child experiencing side-effects for specific drugs as well as the frequency that household reported that they believe children in general experience side-effects. Note that 29% of households reported their child having side-effects from oral steroids, 27% reported side-effects from rescue medications (albuterol) and 16% reported side-effects from control medications (inhaled steroids). Medication is the largest category of direct costs for a typical asthmatic; however, the market prices of these medications do not reflect the perceived costs to households. Experiencing negative side-effects is likely to have substantial influence on household behavior and this should be incorporated into the preference function to adequately measure and welfare changes.

Table 6: Frequency of Reported Side-Effects and Perceived Side-Effects

Medication	Personal	Others
Oral steroids	29	37
Albuterol	27	20
OTC allergy medications	18	13
OTC cold/flu medications	18	15
Inhaled steroids	16	16
Antiobiotics	16	9
OTC asthma medications	13	10
Intal	10	5
Tylenol	8	3
Vitamins	5	5

Note: Personal is the percentage of households *reporting their child having* side-effects as a result of that medication. Others is the percentage of households reporting they *believe children generally* have side-effects as a result of that medication

2. Adherence to Prescription Medication

Even if prescription medication did not enter the household's preference function directly, there would still be difficulties in using the observed household choice. This difficulty stems from the imperfect information and uncertainty surrounding actual prescription medication usage. Much attention has been given in the public health literature on the discrepancies between national guidelines on prescribing asthma medication and actual prescribing patterns. Furthermore, even if the prescriptions do meet the asthma management guidelines, there is ample evidence of non-compliance on the household level, and moving households towards appropriate usage of medication is a major goal of many asthma interventions. The first hurdle to adherence to prescription medication is ensuring that the prescription is filled: in our sample 25% of the households (n=210) report that they had at some time been given a prescription that they were not able to buy because it was too expensive. It is unclear whether not filling the prescription reflects a fully informed trade-off by the household or the result of a subjective assessment that underestimates the benefits of medication usage. For example, use of asthma medication to control the chronic inflammatory component of asthma is a case of investment under uncertainty. In order to decrease inflammation, control medications need to be taken consistently for 4-6 weeks which leads to a delay between taking the medication and experiencing the benefits. After this fixed investment, the benefit is the reduction of the probability of an asthma exacerbation. Families may be unwilling to make the investment in a prescription medication it the benefits are uncertain and occur in some future period. A more complex issue is the difficulties of communicating to households that the benefits will not be realized until after the initial investment, and there is substantial evidence in the public health literature that households confuse the delayed benefits of control medications with the more immediate benefits of rescue medications.

Within our sample, 17% of those prescribed a control medication were not taking the medication in the manner in which it is intended and of those prescribed a rescue medication, 43% were not taking the medication as intended. These patterns suggest an over-reliance on rescue medications, which has been reported in other populations (Boschert, Sadof, Brandt,

2006). The frequency with which households incorrectly use the rescue medication not only reflect what can be thought of as an inefficiency in health production in the current period, but it also perpetuates the inefficiency into future periods. Approximately 15% of the households reported that the prescription medication used by their child had either worsened their child's asthma or left it unchanged. These assessments are likely to drive the choices over medications in the next period, and if they are a result of non-adherence, then "non-optimal" choices could be perpetuated.

While the causes for non-compliance are complex and not well understood, it does suggest that using observed expenditures on medications may be confounded by factors other than preferences over health states.

3. Non-market Behavioral Choices

In addition to using medication to treat both components of asthma, chronic inflammation and acute bronchial constriction, standard asthma management guidelines include recommended behavioral changes. Most of these behavioral changes are focused on reducing exposure to possible asthma triggers (Boschert, Sadof, Brandt, 2006). Table 7 lists the changes undertaken on a regular basis as well as large, one-time changes that were made to prevent asthma exacerbations.

Table 7: Non-market Household Mitigating and Averting Behaviors

Routine Behaviors	Frequency	N
Check for smog alert 2 or more times a week in summer	83%	190
Change activity on high smog days	78%	189
Dusting frequently	59%	200
Vacuuming frequently	59%	200
Mold removal	51%	197
Close windows	46%	199
Restrict amount of child's time outside	46%	201
Restrict where child can play	44%	201
Restrict amount of child's activity	38%	200
Limit where pet can spend time	31%	154
Restrict child's diet	15%	196
One-Time Changes		
Move household to avoid triggers	6%	199
Stopped smoking	15%	189

In focus groups parents reported that these averting behaviors required both substantial time investments and a level of persistence, which in itself was a burden on household relations. A second observation from the focus groups was that when households were asked what they do to prevent asthma exacerbations, the first reported changes were of the type listed above, not purchases nor taking medication. Our interpretation is that these behaviors are pertinent to how household's perceive the impact of asthma, because these are behaviors that they must maintain over time. The implication for data analysis is that because there are no markets for these regular choices the market data that are used in estimating a cost function are incomplete. Second,

households described the fatigue from constantly monitoring their child's health and modifying the home to reduce triggers and this psycho-social burden is not reflected in any market data.

B. *Validity of a production function for health.*

In section 4.b we discussed conceptual limitations of the health production function. Here we discuss findings from the survey that suggest how these conceptual limitations apply to the case of valuing asthma morbidity.

1. *Defining health status*

Characterizing asthma severity is the subject of substantial epidemiological research, because of the difficulty of capturing the natural variation in frequency and degree of symptoms. In our focus groups we asked households to describe what they consider "typical", "good" and "bad" asthma days. This process generated a set of impacts commonly used to describe asthma morbidity and included many impacts in addition to the standard asthma symptoms or healthcare utilization. The impacts considered important to households included symptoms (wheezing or coughing, shortness of breath, black under eyes, increased mucous/phlegm or sputum, ribs showing, easy of breathing), activity limitations (interrupted playtime, ability to walk stair, ride bike or jump rope, ability to talk and sing) and social impacts (avoiding places with triggers, restricting time outdoors). Social impacts were used 79% of the time to describe asthma morbidity, activity limitations were used 96% of the time, and physical symptoms were used 98% of the time. This finding is consistent with the literature that suggests that households tend to describe the severity of their medical condition in terms of activity limitations or impact on quality of life, where as medical professional tend to categorize severity based on frequency of physical symptoms. For the purposes of estimating welfare effects of morbidity, it is the household perspective that matters and drives behavior. Furthermore these impacts affect households much more regularly than do the extreme events of emergency room visits or hospitalizations. Table 8: Asthma Related Morbidity, presents the percentage of household who report that their child experiences limited activity levels, a social impact or a physical symptom on each type of asthma day.

Table 8: Asthma Related Morbidity

|	Percent experiencing		
	Limited activity level	**Social impact**	**Physical Symptom**
Good day	13	29	28
Normal day	47	43	63
Bad day	93	73	97

These data suggest three modifications of the standard household health models. First, welfare estimates that are based on avoiding unplanned medical visits, emergency room visits or hospitalizations miss the larger more routine impact of asthma morbidity on quality of life. The vector that describes health status should include quality of life impacts that are disease and age specific and should not be limited to physical symptoms. Last, because the outcome of interest is a vector of related impacts, data analysis should utilize a multivariate approach with both symptoms and psycho-social outcomes.

2. Constrained Choice Sets

Reducing exposure to asthma triggers is an important part of asthma management and is often described as the most important behavioral change a household can take. In our sample, households were able to identify triggers for their child. The ten most commonly reported triggers for asthma exacerbation were: an existing cold/flu (86%), air pollution (74%), pollen (66%), exercising outdoors (63%), tobacco smoke (53%), dust (53%), outdoor smoke (52%), strong winds (50%), mold (40%), animal dander (33%). Given the parents' perception of asthma triggers, the household model would then posit that households would make the relevant choices to reduce (avert) these risks; however, reducing exposure is not feasible for an important class of asthma triggers. The total height of the bars in Figure 1 shows the frequency households reported the exposure to be an asthma trigger for their child. The shaded are of the bars indicate the degree to which households that reported the item as a trigger felt they had control over their child's exposure. While households reported the ability to limit or reduce exposure to many of the commonly cited asthma triggers (exercising outdoors, tobacco smoke, dust, mold, animal dander, cleaning solutions, food allergies, and roaches), this was not uniformly the case. Of the top ten potential triggers, there were five triggers that more than 20% of the household reported that they had no control over their child's exposure (pollution 41%, strong winds 31%, cold/flu 28%, pollen 25%, and outdoor smoke 20%).

Figure 1: Perception of control of asthma triggers
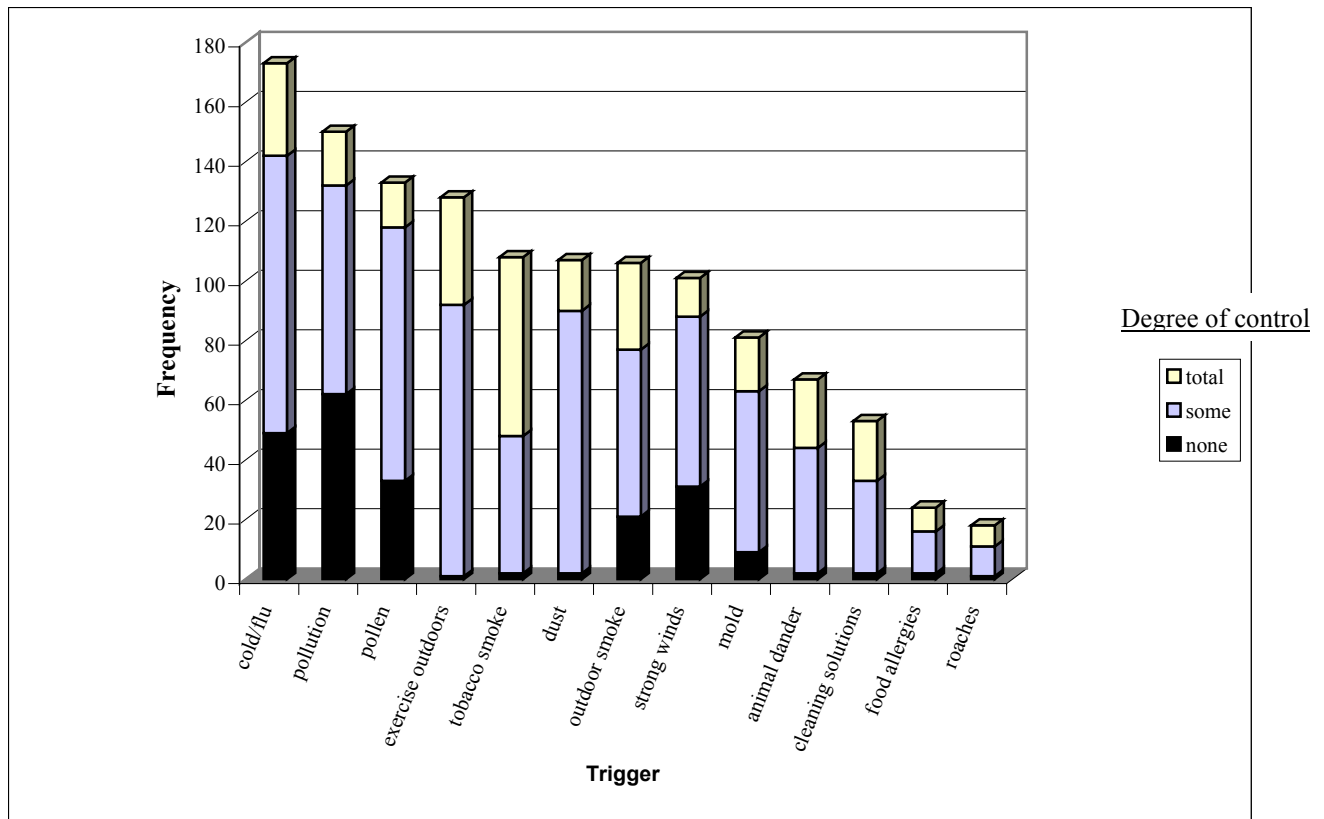
One assumption of the health production function is that the household is able to purchase any number of units of inputs if they are willing to pay the price. The reality was different for 11% of our sample (n=202) who reported difficulties in making appointments with their medical doctor

when needed, and the 7% (n=202) who report that their medical doctor is not helpful when their child's asthma worsens.

These findings suggest that one of the cornerstones of asthma management, averting risk through reducing exposure to triggers, does not readily translate into a production function framework, because the ability to control these exposures is limited. In addition to limited control of asthma triggers, households also reported limitations in the quantity and quality of available medical inputs. The results of our survey corroborate our concern that there may be no interior solutions to the household's maximization problem as commonly formulated.

3. Divergence Between Objective and Subjective Health Production Function

A fundamental assumption of the household production model is that households perceive a health production function and make health choices accordingly. We found substantial differences in how households conceptualize the process that determines asthma status. Of our sample, 16% (n=191) disagreed that asthma can be managed so that a child does not have symptoms; 13% (n=189) agreed with the statement that asthma episodes can cause problems but are not really harmful or dangerous; 58% (n=189) agreed with the statement that asthma episode usually occur without warning; 16% (n=202) report being uncertain about what to do when a child begins to have asthma symptoms. These statistics suggest that households often have incomplete and imperfect information with which to make choices, in other words there is substantial divergence between an objective physiological model of health production and the household's perceptions. They also suggest that rather than households conceiving of a production function, they consider their child's asthma status as exogenous and try to optimize welfare given the asthma status.

We present two avenues in which these divergences may arise. First, prior to making a health related expenditure, households do not have an assessment of how helpful an input will be. To explore this, we asked households both the amount spent on fixed health inputs and their ex post evaluation of the effectiveness of each investment. Table 9 lists the households' assessments of purchases, and three patterns should be noted. First, although each of these investments are commonly suggested averting/mitigating behaviors none were unanimously helpful (ranging from 44% to 98% reporting that the investment was helpful). Second, the investment that was most often reported to be helpful was the nebulizer, which provides relief during a current asthma exacerbation followed by a spacer which helps in delivery of medication, while those investments that were less likely to be reported as helpful were those that reduce triggers thus reducing the probability of an exacerbation in the future (e.g. air filters, removing pests, and HEPA vacuums). Third, households perceived a value to the peak flow meter, which provides information useful in asthma management but which in itself does not reduce triggers or alleviate symptoms. These patterns suggest that households' ex ante expectation and ex post outcome do not coincide for all investments, and the time frame for delivery of benefits may play an important role in household's subjective assessment.

Table 9: Household Investments and Assessments

Investment	% purchased	% reported helpful	N
Air filter	30	69	200
Mattress cover	28	73	200
Pillow cover	27	82	201
Humidifier	31	74	202
HEPA Vacuum	38	70	202
Removed carpet	25	78	201
Removed pests	25	44	202
Remove mold	39	72	201
Nebulizer	31	98	199
Peak flow meter	35	73	199
Spacer	54	93	200

These empirical observations substantiate our concern that a production function for health may impose a relationship between health and choice that is artificially strict and complete.

4. Endogeneity Between Health Inputs and Health Status

One source of endogeneity between the choice of health inputs (z) and asthma morbidity (H) arises from families "benchmarking" their concept of what is normal or attainable respiratory health. For example, one observation in asthma case management programs is that families either do not perceive their respiratory difficulties as asthma symptoms or they come to accept the asthma symptoms as unavoidable. Even within the FACES population, we found that the concept of asthma control deviated from the standard medical concept of asthma control (normal or near normal respiratory function). In our survey, 26% of the households who described their child's asthma as well to completely controlled, actually would be classified as moderate to severe based on the frequency of their daytime symptoms during the winter 2003-2004. As was shown in Table 8: Asthma Related Morbidity, even on a "normal day" from the perspective of the household, children commonly experienced limitations in their level of physical activity (47%), constraints on social interactions (43%) and symptoms (63%). As households that have children who routinely experience asthma morbidity come to expect these impacts as normal or "best possible" level of asthma control, their health investments will reflect this perceived limitation. This benchmarking could be thought of as a creating categories of households with differing perceptions of the frontier for H(z,q), and the perceived frontier would be correlated with the unobservable characteristics that affect the baseline asthma severity.

In the case of household health, the typical instrumental variables approach to the problem of endogeneity is confounded by the complexity of fully specifying the production model. As shown by Griliches and Mairesse (1999), instrumental variables estimates of the production relationship will not produce valid estimates of the coefficients on the health inputs if there are omitted variables from the health production function that are correlated with the elements in the vector of health inputs, z. As shown in the previous sections, households vary in their perceived risks and benefits of health inputs, and unless a model can adequately capture the factors that determine this variation instrumental variables will be an incomplete solution.

7. Stated Preference Approach

A major advantage of the stated preference approach relative to the revealed preference approach is that it allows the researcher to create a trade-off with which to confront survey respondents, thus it makes it possible for the researcher to control the specification of the household production function. Instead of having to estimate an unknown production function, commingled with unknown household preferences, and complicated by the household's unknown subjective perceptions of what it can do to protect or improve its health, the researcher may be able to create his own specification of the trade-off, thereby limiting the unknowns to be estimated from the data to the respondent's preferences.

A second survey was conducted in-person at the FACES office over October 2005-February 2006. This second survey included questions on frequency of asthma symptoms that correspond to the updated Gina asthma severity classification (Luppi, 2004), severity of asthma symptoms, asthma triggers, asthma specific health beliefs, rating and ranking of the impact of asthma on quality of life, causes of household stress and a contingent valuation scenario. A total of 130 FACES households completed the second survey[9].

After the interviewer completed the health status questions, (s)he presented the contingent valuation scenario. The participant was given a brochure that described a hypothetical asthma monitor that could be worn like a watch. The description of how the watch worked included a diagram of a normal airway and an asthmatic airway with both constriction and inflammation.

Figure 2: Brochure Diagrams

Normal Airway

Asthmatic Airway
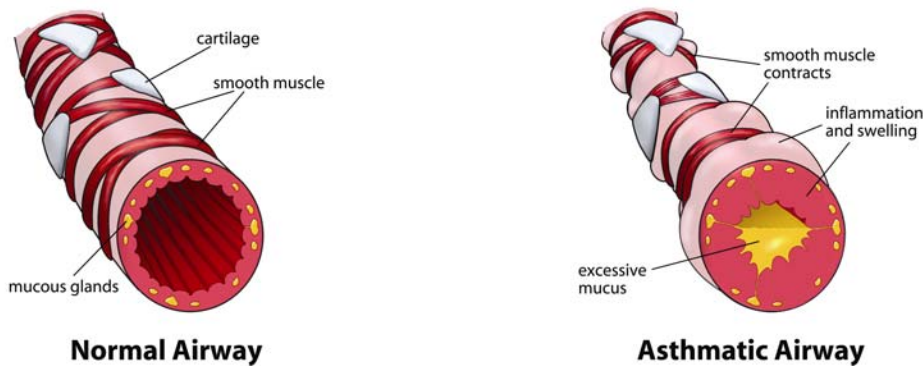
Asthma Monitor

---

[9] We began conducting the contingent valuation survey in Oakland, California in March 2006. Our target is an additional 200 surveys by the end of 2006.

The brochure explained that the watch monitors the level of oxygen in the child's blood and provided an indication when it varied. A green face on the watch indicated that oxygen was optimal whereas a yellow indicated caution and a red face indicated an emergency. By monitoring the child's asthma, it was suggested, action could be taken to stop the asthma from progressing to the point that physical symptoms developed. The hypothetical monitor, the *BreatheRight* watch, was said to have been shown to cut the number of days with asthma symptoms by one-half. We used a one and a half bounded dichotomous choice format to elicit bids for the hypothetical scenario. Initial bids were based on the distribution of responses from a pilot of twenty-two non-FACES households in the Fresno area conducted in August of 2005[10]. Stating bids and subsequent bids were updated following Cooper, Hanemann and Signorello (2002).

We crafted the hypothetical scenario to have six characteristics relevant to the findings of the first survey that were discussed in Section 5:
> 1. The scenario reduced morbidity without relying on medication, and thus would not be confounded by preferences for medication.
> 2. The device did not require behavioral changes to be effective, thereby reducing the issue of non-adherence.
> 3. The tool reduced both the physical symptoms and the stress of monitoring the child's asthma, which addresses the larger issue of how asthma morbidity affects quality of life.
> 4. The device helped families communicate quantitative information about their child's asthma, improving access to health care when needed.
> 5. The instrument provided objective information on the child's health status and assisted families in assessing health risks and effectiveness of averting and mitigating behavior.

Results from the contingent valuation survey will be reported in future research.

Conclusions
In this paper we discuss two approaches to estimating the willingness to pay (WTP) for reduced asthma morbidity, contingent valuation and health production function. In the health production model the health outcome is a function of exposure to asthma triggers, mitigating and averting behavior and household's perceived risks. We find that variation in expenditures is explained by attitudes towards asthma specific health investments including concerns of associated risks and perceived effectiveness. The survey data indicate that households select from a small number of discrete health investments and that most risk reducing behavior are daily behavioral modifications with no relevant market prices.

We argue that the discrete nature of health investments and socio-cultural patterns of health care utilization make the revealed preference approach inadequate for the case of asthma. As an alternative we present a contingent valuation scenario that was specifically developed to minimize systematic variation in preferences for characteristics related to the scenario rather than the reduction in asthma morbidity. For this purpose, guided by extensive testing in focus groups, we selected a scenario based on a hypothetical asthma monitor that provides to the wearer an indicator of current asthma status.

---

[10] Participants for the pilot were recruited through a newspaper ad in the local paper (*Fresno Bee*) and a recruitment table at the American Lung Association's annual walkathon in Fresno.

References
To be added

Individual Preferences and Household Choices: The Potential Role of Dependency Relationships

Mary F. Evans, Christine Poulos, and V. Kerry Smith

Background

- STAR grant: applying weak substitution to value air quality improvements that improve health

- 3 phases of research for two study populations:

	Children	Older adults
1. Develop theoretical model		✓
2. Verify environmental health impacts using secondary data	ongoing	✓
3. Survey parent-child and older adult-caregiver pairs to value air quality improvements		Focus groups in 2005

Motivation

- Phase 3 requires understanding of dependency relationships (young child/adult parent, older adult/caregiver).

- Central to characterizing and interpreting behavior, and to designing stated preference surveys:
  - Intra-household allocation process

Purpose

- Describe a conceptual framework and testable theoretical model of dependency relationships that are expected to influence the value of environmental quality improvements that affect children and older adults

- Describe internet survey activities that:
    - Informs conceptual structure
    - Complement/substitute for focus group activities and/or cognitive interviews

Theoretical background

- Collective model (Chiappori and co-authors) permits the recovery of individual preferences from household behavior. Offers strategy for analyzing household structure and dependency relationships.

- What information is used for identification?
  - Chiappori: Focus on the *intensive* margin.
  - Our proposal: Focus on the *extensive* margin.

- Current tests of the collective model rely on observing the full system of demands.

- Is there an alternative test that does not rely on this information?

Review of Chiappori collective household framework

- ☐ Background
  - ▪ Household is best viewed as a collection of individuals with different preferences; analysts usually observe *household* not individual demands for goods and services.
  - ▪ Household "behavior" is unlikely to be described adequately by the unitary model – *i.e.* treating choice as if it was motivated by a single agent's decisions.
- ☐ Assumptions
  - ▪ Each household member has own preferences that are known by other household members
  - ▪ Collective decisions of household are Pareto efficient

Focus of Chiappori and co-authors' research

Browning and Chiappori [1998], Chiappori and Ekeland [2006; forthcoming]

- **Key Issue**: What does the efficiency assumption imply for household demands and specifically for the matrix of Hicksian price effects?

  - Question directs attention to the *intensive* margin of choice

Basic structure of argument

Context and definitions

- Assume two members (I and II) of household
  - p = prices of market goods (T x 1 vector)
  - X = quantity of market goods consumed by household, $x^I + x^{II} = X$
  - Z = private good that is public consumption to members of household
  - $U^i\left(x^i, x^j, Z\right) = i^{\text{th}}$ individual's utility function ($i \neq j$)
  - y = household income, $y = p^T \cdot \left(X + Z\right)$

Key efficiency assumption

- There exists a differentiable, homogeneous of degree zero, function μ(p, y) such that for any (p, y) the vectors (x^I, x^{II}, Z) are solutions to the following optimization problem:

$$\underset{x^I,x^{II},Z}{Max}\ \mu(p,y)U^I\left(x^I,x^{II},Z\right)+\left(1-\mu(p,y)\right)U^{II}\left(x^I,x^{II},Z\right)$$

$$subject\ to\ p^T\cdot\left(X+Z\right)=y$$

- Yields household (Marshallian analog) demand functions: $f_s(p,y,\mu)$

- Expenditure minimization problem yields (Hicksian analog) demand functions: $h_s(p,U,\mu)$

Key generalization

- Duality implies, holding μ constant

$$\frac{\partial h_s}{\partial p_t} = \frac{\partial f_s}{\partial p_t} + \frac{\partial f_s}{\partial y} \cdot f_t$$

- Allowing μ to vary with prices and income,

$$\frac{\partial h_s}{\partial p_t} = \frac{\partial f_s}{\partial p_t} + \frac{\partial f_s}{\partial y} \cdot f_t + \frac{\partial f_s}{\partial \mu} \left(\frac{\partial \mu}{\partial p_t} + \frac{\partial \mu}{\partial y} \cdot f_t \right)$$

Element of matrix of Hicksian price effects (pseudo-Slutsky matrix)

Element of conventional Slutsky matrix (symmetric)

Leads to matrix that is at most rank one (in two-person household) = **basis of test**

Illustrative example

- Provides intuition for alternative test of the collective model

- Informs development of choice questions

- Assumptions
  - Up to a two-person household
  - Linear indirect utility function
    - Form varies according to structure of household
  - Consider change in indirect utility from improvement in air quality, q, that reduces the amount of care giving time required for self or another individual (small child, teenager, older adult)

Individual only (no altruism, no income sharing)

- Indirect utility function with initial air quality

$$V_0 = \alpha_0 + \alpha_1 w + \alpha_2 y + \alpha_3 q_0 - \lambda \overline{L}$$

  - with w the wage rate, y non-wage income, $\overline{L}$ (fixed) care giving time to self, q_0 initial level of air quality.

- Indirect utility function with improved air quality

$$V_1 = \alpha_0 + \alpha_1 w + \alpha_2 (y - T) + \alpha_3 q_1$$

  - with T the cost of the program to improve air quality (and reduce care giving time).

- Change in indirect utility

$$\Delta V = \lambda \overline{L} + \alpha_3 (q_1 - q_0) - \alpha_2 T$$

Altruism (no income sharing)

- Indirect utility function with initial air quality

$$V_0 = \alpha_0 + \alpha_1 w + \alpha_2 y + \alpha_4 h(q_0) - \lambda \overline{L}$$

  - with h describing the health of the dependent as a function of air quality, $\overline{L}$ care giving time to dependent.

- Indirect utility function with improved air quality

$$V_1 = \alpha_0 + \alpha_1 w + \alpha_2 (y - T) + \alpha_4 h(q_1)$$

- Change in indirect utility

$$\Delta V = \lambda \overline{L} + \alpha_4 [h(q_1) - h(q_0)] - \alpha_2 T$$

Income sharing (no altruism)

- Indirect utility function with initial air quality

$$V_0 = \alpha_0 + \alpha_1 w + \alpha_2 \left[b_0 + b_1 y + b_2 w + b_3 q_0 + b_4 \overline{L} \right] - \lambda \overline{L}$$

  - where the term in brackets represents the individual's share of household income.

- Indirect utility function with improved air quality

$$V_1 = \alpha_0 + \alpha_1 w + \alpha_2 \left[b_0 + b_1 (y - T) + b_2 w + b_3 q_1 \right]$$

- Change in indirect utility

$$\Delta V = (\lambda - \alpha_2 b_4) \overline{L} + \alpha_2 b_3 (q_1 - q_0) - \alpha_2 b_1 T$$

Altruism and income sharing

□ Indirect utility function with initial air quality

$$V_0 = \alpha_0 + \alpha_1 w + \alpha_2 \left[b_0 + b_1 y + b_2 w + b_3 q_0 + b_4 \overline{L} \right] + \alpha_4 h(q_0) - \lambda \overline{L}$$

□ Indirect utility function with improved air quality

$$V_1 = \alpha_0 + \alpha_1 w + \alpha_2 \left[b_0 + b_1 (y - T) + b_2 w + b_3 q_1 \right] + \alpha_4 h(q_1)$$

□ Change in indirect utility

$$\Delta V = (\lambda - \alpha_2 b_4) \overline{L} + \alpha_2 b_3 (q_1 - q_0) + \alpha_4 \left[h(q_1) - h(q_0) \right] - \alpha_2 b_1 T$$

Matrix of household types

Label	Coefficient on Δq	Coefficient on $\overline{L}$	Coefficient on T	Number in household
Individual only	α_3	λ	$-\alpha_2$	1
Altruism only	?	λ	$-\alpha_2$	1
Income sharing only	$\alpha_2 b_3$	$\lambda - \alpha_2 b_4$	$-\alpha_2 b_1$	2
Altruism and income sharing	?	$\lambda - \alpha_2 b_4$	$-\alpha_2 b_1$	2

Choice questions

- ☐ Objective: understand how people make choices (not to obtain WTP estimates for policy use)
- ☐ Strategy:
  - ■ Two double-bounded questions per respondent
    - ☐ One focusing on respondent
    - ☐ Another focusing on another individual: young child (2-5), teenager (13-17), or older adult (>62)
  - ■ Design attributes constant within questions for respondent
  - ■ Interval censored model to estimate valuation function

Choice questions

- Respondent asked to suppose that subject (self or other individual) has asthma

- General choice question: "Would you pay T for a program that will improve air quality and reduce the amount of time you allocate to care giving for [yourself / young child / teenager / older adult] from L to zero?

  - Double-bounded question with respect to T

  - L (2 levels), T (4 levels) randomly assigned

  - Sequence of responses to each double-bounded question sort individuals into four bins (yes/yes, yes/no, no/yes, no/no)

Survey details

- Add survey questions to weekly internet panel
  - 1000-2000 respondents (18+ years); $600-$1000/question
  - Socioeconomic data provided: education, income, household size and composition
  - Data available in days
- Questions
  - Choice questions
  - Additional characteristics (presence and age of children in HH, presence and age of older adults in HH, asthma in HH members)

Conclusion

- Dependency relationships and household structure are potentially important in estimating the benefits to improved environmental quality.

- Proposed tests have focused on the *intensive* margin. We propose an alternative that focuses on the *extensive* margin that does not require observation of full system of demands.

Air Pollution and Asthma:
Preliminary Results from a Daily Time-Series Study of San Francisco

By

Charles W. Griffiths and Nathalie B. Simon
U.S. Environmental Protection Agency

Air Pollution and Asthma:
Preliminary Results from a Daily Time-Series Study of San Francisco

By Charles W. Griffiths and Nathalie Simon[1]

Asthma is a chronic lung disease characterized by intermittent, recurring episodes of wheezing, breathlessness, tightness of the chest, and coughing. These episodes are caused by inflammation of the airways that carry air into and out of the lungs. Asthma is considered to be a growing problem in the United States, especially among children. The prevalence of asthma increased 46 percent between 1982 and 1993 in the United States. While increases in prevalence have been documented in all age, race, and gender groups, the increase has been most significant among children -- individuals under the age of 18 -- with a staggering 80 percent since 1982.

While the exact causes of the illness remain unknown, asthma attacks can be triggered by a number of factors including exposure to allergens (e.g., dust mites, pollen, mold, pet dander, and cockroach waste), strong fumes, respiratory infections, exercise, dry or cold air, as well as air pollution (including ozone and particulate matter). Despite recent efforts to reduce ambient levels of air pollution, approximately 46 million people lived in counties that did not meet the air quality standards for at least one of the six criteria pollutants in 1996. The combination of poor air quality with other triggers is often most extreme in urban centers where a disproportionate number of minority and low income households reside.

A relatively large number of studies exist that focus on the temporal relationship between

air pollution and asthma attacks resulting in Emergency Room visits or Hospital Admissions. However, these studies by design focus on severe outcomes and miss milder ones – asthma attacks that are alleviated through medication use and do not require immediate medical attention. Those studies that do examine milder forms of asthma symptoms (e.g., respiratory symptoms or increased use of asthma medication) generally take the form of cohort or panel studies and have tended to focus on children. This paper presents the preliminary results of a time-series analysis of the effects of acute exposure to ambient air pollution on the incidence of asthma attacks, as measured by prescription counts for short term "quick relief" medications. Using prescription data from San Francisco, California, we estimate the relationship between exposure to ozone and PM10 and asthma symptoms.

Background

The relationship between short-term increases in ambient levels of air pollution and asthma outcomes has been documented in a number of venues using two types of studies: daily time series studies and cohort or panel studies. Daily time-series studies have been used to model the relationship between air pollution and a number of health outcomes including daily mortality and other relatively severe respiratory outcomes such as hospital admissions, emergency room visits and doctor visits. A relatively large segment of these studies have focused on asthma. In a study by Walters et al. (1993), for instance, daily levels of SO2 and black smoke were found to have a positive association with hospital admissions for asthma in Birmingham, UK. A similar result was found in Birmingham, Alabama in a study focused on hospital admissions due to pneumonia and Chronic Obstructive Pulmonary Disease (of which asthma is a component) among elderly inhabitants (Schwartz 1994). Also found was a positive

association between air pollution levels and doctor visits for asthma in London (Hajat et al. 1999). In Barcelona, Spain, a positive association between emergency room visits for Chronic Obstructive Pulmonary Disease and air pollution levels was found (Sunyer et al.1993). While these studies are indicative of the detrimental effects of short-term increases in air pollution on rather severe asthma outcomes, they give no indication of the effects of air pollution exposure on asthma outcomes that are milder in nature.

Panel or diary studies can provide some indication of the effects of air pollution on less severe health outcomes. They model symptoms experienced by panel members as a function of air pollution levels. While most cohort studies are focused on children, some studies have found positive and significant effects of air pollution exposure on exacerbation of asthma symptoms at other ages. Neukirch et al. (1998) found measurable short-term effects of low-level air pollution in Paris France on nonsmoking asthmatic adults diagnosed with mild or moderate asthma. Similarly, Newhouse et al. (2004) found that ozone concentrations on the previous day were associated with a number of symptoms including wheezing, headache, and fatigue in their panel of 24 individuals aged 9-64 with physician diagnosed asthma. Ostro et al. (1991) also found a strong association between daily air pollution levels (specifically airborne acid aerosols, particulates, and sulfates) and increased asthma symptoms among a panel of asthmatics in Denver, Colorado. Similar results have been reported in the Utah Valley (Pope et al. 1991), Glendora California (Krupnick et al. 1990), and the Netherlands (Hiltermann et al. 1998) among other places.

While diary studies are useful in isolating the effects of short-term increases in pollution on milder outcomes, these studies face several difficulties. Among these difficulties, as noted by Schwartz et al. (1991), is the fact that daily symptom rates are often highly correlated from one

day to the next and the heterogeneity among subjects causes dependencies in the data. Some study results are also limited by the availability of particulate pollution measures while others are limited by panel size or length of study period.

In contrast to the studies described above, our study examines the effect of short term or acute exposures to air pollution on asthma symptoms as measured by the purchase of quick relief asthma medications in San Francisco, California. Zeghnoun et al. (1999) explore a similar relationship in Le Havre, France and find statistically significant effects of black smoke, NO2 and SO2 on respiratory drug sales for mucolytic and anti-cough medications for children and adults. Our study in contrast is focused on quick relief asthma medications. We hypothesize that acute exposure to air pollution may make an individual more susceptible to asthma attacks, causing an increase in the use of quick relief medications.

Methodology

This study looks at the effects of differences in short term air pollution exposures on the occurrence of asthma attacks, where asthma attacks are proxied by the number of prescriptions for quick relief asthma medication filled. The total count of prescriptions for quick relief asthma medication is explained using measures of asthma triggers and other cofactors. The study utilizes a dataset of asthma drug prescriptions for a large percentage of the pharmacies in the state of California and GIS layers of spatial factors.

In this study, our "health" outcome (filling asthma prescriptions) is not a "direct" effect of air pollution exposure, but rather a secondary effect. That is, the true sequence of events goes as follows: short-term exposure to air pollution makes an individual more susceptible to asthma triggers leading to an exacerbation of asthma symptoms which in turn causes an increase in

5

asthma medication use. The increase in asthma medication use eventually (perhaps with a lag) leads to the filling of a prescription.   The urgency with which a prescription needs to be filled will vary across individuals and their initial stock of asthma medication, making short term effects more difficult to observe.

An individual suffering from asthma will use his inhaler with some probability based upon the amount of pollution present, current weather conditions, and seasonal factors.

$$\text{Pr(Inhaler use)}_t = f(\text{Pollution}_t, \text{Weather}_t, \text{Seasonal Factors}_t) \qquad (1)$$

We assume that each day the individual makes an independent decision, where the choice is whether or not to use the inhaler based on the contemporaneous pollution, weather, and seasonal factors. If we were able to witness the individual's use of his inhaler, then we could model this behavior using daily cofactors. Minor modification would be possible if asthma attacks were serially correlated, or if the use of an inhaler one day was related to the conditions on the previous day as well as the current day.

In our case, we do not witness the individual's use of the inhaler, only the purchase of a new inhaler when the old one is empty (or close to empty). Therefore, the observable event, the number of prescriptions, is a function not only of the contemporaneous factors, but the total amount of pollution and weather conditions over the recent past, as well as the seasonal factors.

$$(\text{Number of Prescriptions})_t = f(\text{Pollution}_t, \text{Pollution}_{t-1}, \ldots, \text{Pollution}_{t-m}, \text{Weather}_t,$$
$$\text{Weather}_{t-1}, \ldots, \text{Weather}_{t-m}, \text{Seasonal Factors}_t) \qquad (2)$$

The number of days, m, which defines the recent past needs to be long enough to capture the signal of inhaler use, but should not be so long as to add additional noise to the model. Seasonal variation does not need to be modeled as an aggregation over time since it can be captured using other methods – in this paper, through the use of continuous trigonometric cycles of varying

length. If the number of prescriptions follows a Poisson distribution, then we can model the

mean incidence rate, that is, the number of prescriptions on any given day, as

$$
\begin{aligned}
r_t = \exp(\alpha &+ \beta[Pollution_t + Pollution_{t-1} + \ldots + Pollution_{t-m}] \\
&+ \gamma[Weather_t + \gamma Weather_{t-1} + \ldots + \gamma_m Weather_{t-m}] \\
&+ \delta \, Seasonal \, Factors_t
\end{aligned}
\tag{3}
$$

Since each day should be given the same weight, β and γ can be estimated for the sum of

pollution and weather cofactors over the recent past.

Since we are modeling a single CSMA over time, we assume that the exposure rate does

not change from one day to the next and, therefore, do not include it explicitly in the estimation.

Modeling the number of prescriptions in this fashion means that β is then a semi-

elasticity of the impact of a one-unit change in pollution. In other words,

$$
\frac{\partial r_t}{\partial P} \cdot \frac{1}{r_t} = \beta
\tag{4}
$$

where a one unit change in pollution produces a β percent increase in inhaler prescriptions.

Data

The number of prescriptions for quick acting asthma medication was obtained from

NDChealth (hereafter, NDC), a Phoenix-based company that maintains prescription-related data

for marketing research. NDC maintains two datasets of use for this study, a "retail pharmacy"

database and a "patient" database. The pharmacy database contains dispensing records from

approximately 36,000 pharmacies nationwide, and captures approximately 70% of the volume of

traditional pharmacy-dispensed prescriptions. Hospital, military and mail order pharmacies and

prescriptions dispensed to institutionalized patients are not included in this database, which may

pose a problem in the future as mail order prescriptions grow, but is probably not important here. The patient database is a subset of approximately 14,000 of the pharmacies in the pharmacy database. The patient database is a more complete database, in many cases including the patients age and gender, along with a unique patient identifier so that the history of a patient may be followed. Not included in the database, and unknown to NDC, is any information that could personally identify a patient (such as a name, address or phone number) and NDC has been very careful not to release any individual patient data, even with the anonymous identifier.

Prescription data were provided for San Francisco by NDC, segregated by the level of asthma severity of the patient. Asthma severity is classified as mild intermittent, mild persistent, moderate persistent, and severe, based upon the number and combination of prescriptions that the patient fills for both quick-relief and maintenance asthma medicine over the 12 month calendar year (NIH, 1997). Generally, asthma medications fall into one of two categories: (1) short-term treatments intended to provide quick relief in the event of an asthma attack and (2) long-term maintenance therapies intended to prevent asthma attacks. Mild asthmatics are those patients prescribed a quick-relief medication only. Patients with mild persistent asthma not only are prescribed a quick-relief medication but are also prescribed a single controller or maintenance therapy. Moderate asthmatics are prescribed two controllers operating by different modes of action in addition to the quick relief medications, while severe asthmatics are prescribed three controllers with different modes of action. Should an individual's asthma severity level shift over the 12 month period, the individual is assigned to the most severe of the categories for which he/she qualifies. A list of the quick acting and controlling asthma medication is listed in Table 1.

```
Table 1: Asthma Medication

Symptomatic Therapy (Quick Relief)
Albuterol
Bitolterol
Isoetharine
Metaproteronol
Pirbuterol
Terbutaline

Controller Therapy (Long-term preventative)
Inhaled Corticosteroids
  Beclomethasone
  Budesonide
  Flunisolide
  Fluticasone
  Triamcinolone
Leukotriene Antagonists
  Motelukast
  Zafirlukast
  Zileutin
Long Acting Beta Agonists
  Salmeterol
Xanthine Derivatives
  Aminophylline
  Dyphylline
  Oxtriphylline
  Theophylline
Mast Cell Stabilizers
  Cromolym
  Nedocromil
```

medication in a five digit zip code for each quarter from 1998 to 2001 were used.   Data are

given by dispense quarter and the zip code of the dispensing pharmacy. These data are further

disaggregated by asthma severity.

The prescription data used in this analysis are limited in the following way. They only

include counts of prescriptions for quick relief asthma medication from those pharmacies that

"consistently" report this information. "Consistent" reporting is defined by NDC as pharmacies

for which fewer than 11 days of data are missing in any 30 day period.

The air pollution data are publicly available from the California Air Resource Board. Daily observations on the levels of PM10, SO2, NOx, and ozone are available for 71 monitors in San Francisco.

The weather data come from the National Climatic Data Center. Daily observations for the average, minimum, and maximum temperature, as well as relative humidity, and the minimum and maximum relative humidity were obtained for 97 active weather stations in San Francisco.

The summary statistics for the data used in this analysis are listed in Table 2.

Table 2: Summary Statistics

Variable (units)	Number of Observations	Mean	Min	Max
Total Daily Prescriptions (number)	1428	852.0	28	1714
Mild Intermittent	1428	389.5	15	950
Mild Persistent	1428	276.4	9	579
Moderate	1428	128.8	1	260
Severe	1428	57.3	3	122
Daily Minimum Temperature (^{o}F)	1672	39.69	15	67
Daily Average Relative Humidity (%)	1672	68.09	37	98
Daily Average PM10 ($\mu g/m^3$)	1499	38.01	2.60	227.72
Max of Ozone 1Hr (ppm)	1672	0.08	0.04	0.18

Results

We estimate equation 3 above using a standard negative binomial regression – the more generalized form of the Poisson model. Total counts of prescriptions filled over the course of a

week are regressed against minimum temperature, relative humidity, pollution measures, time trends and trigonometric terms designed to capture cyclical trends ranging from 1 year to 2.4 months in length. Weekly counts are used to minimize day of week effects and effects of pharmacy closures due to holidays. Temperature and pollution measures are summed over 180 days preceding the weekly counts. As described above, summing these factors in this way allows us to more accurately capture the effects of pollution and weather on the dispensing of quick relief asthma medications. A 180-day window was selected upon inspection of the plot of the residuals of prescription counts (with seasonal variation and time trends removed) against pollution measures.[2]

The model was constructed in a step-wise manner in which we first added our season and time controls, using the Akaike Information Criterion to inform the choice of model before incorporating other factors. Once the seasonal factors were selected, we incorporated meteorological metrics including temperature and relative humidity. Minimum, maximum and daily average measures were tested against one another. Minimum daily temperature and average daily relative humidity provided the best fit according to AIC.

With weather factors controlled for, we added ozone and PM10 measures to the model. We experimented with 8-hour and 1-hour ozone measures, and found that daily maximum observations of 1-hour ozone readings provided the best fit. Daily average PM10 was similarly selected for inclusion.

Once the model construction was completed for counts of total prescriptions, we applied the same model construct to counts of prescriptions by severity level. Results for all five regressions are reported in Table 3.

---

[2] Troughs and peaks in the residuals were matched with those observed in the pollution data.

11

Table 3: Regression Results

Variable Name	Total Prescriptions		Mild		Mild Persistent		Moderate		Severe	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
trend	-0.0001	0.0001	-0.0001	0.0001	-0.0001*	0.0001	0.0000	0.0001	-0.0002*	0.0001
Year 1999	0.1238***	0.0315	0.1206**	0.0386	0.0978***	0.0302	0.1299***	0.0269	0.2836***	0.0290
Year 2000	0.2051***	0.0589	0.2285***	0.0714	0.1473**	0.0568	0.1559**	0.0506	0.4056***	0.0545
Year 2001	0.0062	0.0884	0.0097	0.1075	-0.0867	0.0851	0.0354	0.0758	0.3489***	0.0816
sin1yr	-0.4187***	0.0302	-0.5011***	0.0372	-0.3914***	0.0287	-0.3140***	0.0255	-0.2129***	0.0274
cos1yr	0.2556***	0.0098	0.3547***	0.0120	0.1974***	0.0093	0.1440***	0.0082	0.1044***	0.0087
sin6mo	-0.0707***	0.0058	-0.0924***	0.0070	-0.0774***	0.0055	-0.0221***	0.0049	-0.0101*	0.0053
cos6mo	0.0102**	0.0034	0.0179***	0.0042	-0.0034	0.0032	0.0075**	0.0029	0.0111***	0.0031
sin4mo	0.0317***	0.0046	0.0477***	0.0056	0.0207***	0.0044	0.0167***	0.0039	0.0177***	0.0042
cos4mo	0.0286***	0.0034	0.0447***	0.0042	0.0200***	0.0033	0.0042	0.0029	0.0091**	0.0031
sin3mo	-0.0162***	0.0040	-0.0194***	0.0049	-0.0177***	0.0039	-0.0080*	0.0034	-0.0112**	0.0037
cos3mo	-0.0102**	0.0034	-0.0075*	0.0042	-0.0158***	0.0033	-0.0095**	0.0029	-0.0061**	0.0031
sin2.4mo	-0.0191***	0.0037	-0.0220***	0.0046	-0.0213***	0.0035	-0.0082***	0.0032	-0.0105**	0.0034
cos2.4mo	0.0200***	0.0035	0.0331***	0.0042	0.0105***	0.0033	0.0064*	0.0029	0.0002	0.0031
Minimum Temperature	-0.0003***	2e-05	-0.0004***	2e-05	-0.0003***	2e-05	-0.0002***	2e-05	-0.0001***	2e-05
Average Relative Humidity	0.0002***	2e-05	0.0002***	2e-05	0.0002***	2e-05	0.0001***	2e-05	0.0001***	2e-05
Average PM10	2e-05***	5.13e-06	2e-05***	6.29e-06	2e-05***	4.90e-06	1e-05**	4.36e-06	5.89e-06	4.65e-06
Max of Ozone 1Hr	-0.0536***	0.0085	-0.0705***	0.0105	-0.0397***	0.0081	-0.0462***	0.0072	-0.0279***	0.0077
constant	9.5639***	0.2153	9.0780***	0.2633	8.1843***	0.2059	7.4835***	0.1833	6.0107***	0.1956

Note: ***= statistically significant at 99% confidence level; **=statistically significant at 95% confidence level; *=statistically significant at 90% confidence level

Looking across the five models presented in Table 3, we find mixed results. Minimum temperature enters all five equations with the expected sign. As temperature decreases, we expect to see an increase in asthma symptoms as exhibited in the results. The decrease in the magnitude of the effect across severity levels is not altogether surprising given the increased use of maintenance therapies as asthma severity increases. This is in keeping with a recent study by Delfino et al. (2002) that found stronger associations between air pollution and exacerbation of asthma symptoms among asthmatic children who were not taking anti-inflammatory medications.

The effects of relative humidity are somewhat puzzling, however, as we expected dryer air to exacerbate asthma symptoms. In our results, it seems we have the opposite effect -- as relative humidity increases, so does the number of asthma prescriptions filled.

The effect of pollution exposure on asthma symptoms is also mixed. Daily levels of PM10 have a positive and significant effect on asthma prescriptions with a 1 microgram per cubic meter increase in PM10 resulting in a 0.00002 percent increase in asthma prescriptions each week (or approximately 0.12 additional prescriptions). Assuming each inhaler contains approximately 200 metered doses of the quick relief medication with the recommended usage to relieve symptoms being 2 doses, this translates to approximately 12 "attacks" per week. The effect of ozone exposure however is contrary to what we expected, with our model showing negative but statistically significant effects across all severity levels.[3]

Next Steps

Our regression results are unsatisfying in many respects. We fully expected to see a

---

[3] We intended to incorporate the effects of NOx and SO2 as well, but the data coverage was too spotty for these pollution measures.

negative (and statistically significant) effect of ozone exposure on asthma prescriptions. Still, our analysis is preliminary and we hope that a number of improvements to our model will likewise improve our results.

First, we recognize that our negative binomial model is rather crude. The model does not take into account the autoregressive nature of the data and as such could be producing biased results. Now that we have selected the appropriate terms to incorporate into the model, we intend to explore the use of generalized additive models.

We also recognize that we applied the model we constructed for total prescriptions to other endpoints (counts by severity level) with little regard for whether it produced the best fit in these other contexts. Ideally, we would apply the same methodology for constructing the model to these other endpoints – controlling for the factors we believe to be important (e.g., weather, pollution, and seasonality) but using the AIC to assess which measures best control for these factors and eventually applying generalized additive models here as well.

The selection of the window over which to sum our observations was admittedly rather ad hoc. We intend to explore other more "rigorous" means of selecting the proper window. One option is to conduct a rather crude "expert elicitation" and identify (fewer than 10) physicians whom we could survey on this point. Essentially, we would want to learn from them approximately how often their patients refill their prescriptions.

In addition, we have access to the counts of prescriptions by age group as well as by severity. While time did not permit us to present the results of regressions by age here, we intend to explore the effects of pollution exposure on asthma prescriptions by age group in future analyses.

References:

Delfino, Ralph J., Robert S. Zeiger, James M. Seltzer, Donald H. Street, and Christine E. McLaren (2002). "Association of Asthma Symptoms with Peak Particulate Air Pollution and Effect Modification by Anti-inflammatory Medication Use*." Environmental Health Perspectives.*110(10): A607-617.

Hajat, S., A. Haines, S. A. Goubet, R.W. Atkinson and H.R. Anderson (1998). "Association of Air Pollution with Daily GP Consultations for Asthma and Other Lower Respiratory Conditions in London." *Thorax*. 54: 597-605.

Hiltermann, T.J.N., J. Stolk, S.C. van der Zee, B. Brunekreef, C.R. de Bruijne, P.H. Fischer, C.B. Ameling, P.J. Sterk, P.S. Hiemstra, and L. van Bree (1998). "Asthma Severity and Susceptibility to Air Pollution." *European Respiratory Journal*. 11: 686-693.

Krupnick, Alan J., Winston Harrington, Bart Ostro (1990). "Ambient Ozone and Acute Health Effects: Evidence from Daily Data." *Journal of Environmental Economics and Management*. 18:1-18.

Neukirch, Francoise, Claire Segala, Yvon Le Moullec, Myriam Korobaeff, and Michel Aubier (1998). "Short-term effects of low-level winter pollution on respiratory health of asthmatic adults." *Archives of Environmental Health*. 53(5): 320-9.

Newhouse, C.P., B.S. Levitin, and E. Levitin (2004). "Correlation of Environmental Factors with Asthma and Rhinitis Symptoms in Tulsa, OK. *Ann. Allergy Asthma Immunol*. 92: 356-366.

Ostro, Bart D., Michael Lipsett, Matthew B. Wiener, and John Selner (1991). "Asthmatic Responses to Airborne Acid Aerosols." *American Journal of Public Health*. 81: 694-702.

Peters, Annette, Inge F. Goldstein, Ulrich Beyer, Kathe Franke, Joachim Heinrich, Douglas Dockery, John D. Spengler, and H. Erich Wichmann (1996). "Acute Health Effects of Exposure to High Levels of Air Pollution in Eastern Europe." *American Journal of Epidemiology*. 144(6): 570-81.

Pope, C. Arden, Douglas W. Dockery, John D. Spengler, and Mark E. Raizenne (1991). "Respiratory Health and PM10 Pollution." *American Review of Respiratory Disease*. 144: 668-674.

Schwartz, Joel (1994). "Air Pollution and Hospital Admissions for the Elderly in Birmingham, Alabama." *American Journal of Epidemiology*. 139 (6): 589-98.

Schwartz, Joel, David Wypij, Douglas Dockery, James Ware, Scott Zeger, John Spengler, and Benjamin Ferris, Jr. (1991). "Daily Diaries of Respiratory Symptoms and Air Pollution: Methodological Issues and Results." *Environmental Health Perspectives*, 90: 181-187.

Sunyer, Jordi, Marc Saez, Carles Murillo, Jordi Castellsague, Francesc Martinez, and Josep M. Anto (1993). "Air Pollution and Emergency Room Admissions for Chronic Obstructive Pulmonary Disease: A 5-year Study." *American Journal of Epidemiology*. 137(7): 701-5.

Walters, Sarah ,R.K. Griffiths, J.G. Ayers (1994). "Temporal Association Between Hospital Admissions for Asthma in Birmingham and Ambient Levels of Sulphur Dioxide and Smoke." *Thorax*, 49: 133-140.

Zenghnoun, Abdelkrim, Pascal Beaudeau, Fabrice Carrat, Veronique Delmas, Onealy Boudhabhay, Francois Gayon, Dominque Guincetre, and Pierre Czernichow (1999). "Air Pollution and Respiratory Drug Sales in the City of Le Havre, France, 1993-1996." *Environmental Research*, Section A, 81: 224-230.

No documents are available regarding Bryan Hubbell's discussion comments.

Comments by Glenn Blomquist on
"Risk Assessment and Valuation of Health Effects from Air Pollution"
EPA Workshop on Morbidity & Mortality: How Do We Value the Risk of Illness & Death?
April 10, 2006, Washington, DC

Comments on "Willingness to Pay for Improved Health: A Comparison of Stated and Revealed Preference Models" by W. Michael Hanemann and Sylvia Brandt

This paper discusses two approaches to valuing changes in asthma morbidity among children: (1) contingent valuation and (2) household production of health. The study is designed to facilitate comparison of estimates based on the two approaches by estimating willingness to pay (WTP) for the same households and children. While the comparison is incomplete at this stage of the project, the potential for comparison between the stated WTP in contingent valuation and WTP inferred from household averting and mitigating behavior is great. The focus of this paper is on the results of Hanemann and Brandt's survey of households with asthmatic children, parents' perceptions of risks of asthma attacks, and their behavior with regard to those risks.

The study group was recruited by Fresno Asthmatic Children's Environment Study (FACES) in Fresno, California. All children had clinically diagnosed asthma and nearly 70% of the households had at least one parent who was affected by asthma. These households are familiar and experienced with asthma. The written, mail survey of 202 households was conducted in 2004. Information about asthma severity, medication use, asthma-related expenditures, items and services purchased, household income, and time spent dealing with children's asthma was collected. The comprehensiveness and detail are incredibly good. Any doubt that asthma health risks are partly endogenous surely disappears when confronted with these data.

Table 5 is particularly informative. Seventeen types of fixed costs (expenditures) are listed including purchases of air conditioning, air filters, pest extermination, carpet removal, and pet removal. Eight types of variable costs for household supplies are listed including heater filters, cleansers for mold, and hypoallergenic cleaners. Four types of pharmaceutical costs are listed including prescription asthma medication, over-the-counter drugs, and herbal remedies. Seven types of alternative therapies are listed including nutritionists. The time periods for fixed and variable costs are not specified in Table 5, but if they are (recklessly) lumped together average (mean) expenditures appear to be roughly 1% of median household income for group.

As Grossman (1972) showed more than 30 years ago, time inputs in the production of health matter. My educated guess is that ways in which households change their time allocations due to their children's asthma will be at least as important as changes in money expenditures. About this behavioral response, Hanemann and Brandt have information reminiscent of diary data collected for EPA for a small sample of adult asthmatics in the 1980s. These data for adults are extraordinary and, as evidenced by the recent study by Yen, Shaw, and Eiswerth (2004), are still being gleaned. The new data for households with asthmatic children should be at least as useful. They deserve their own table comparable to Table 5. Hanemann and Brandt report that time costs are sizable. They should be able to estimate them well. They report that 43% of the

households took time off from work to take a child to a medical appointment and that the average (mean) time taken off was 83 minutes. They report 24% of the households took time off from work in the previous year to take a child to an emergency room with an average (mean) of 211 minutes spent. If wage data are available for each household, dollar values of these time costs can be estimated and added to the dollar expenditures on marketed goods and services. If wage data are not available, then given the available information on household income and characteristics of the household, Hanemann and Brandt should be able to use a data set such as the Current Population Survey and estimate wages and/or shadow wages for each of the members of the household. While these time costs may be less than 1% of annual household income, there are indications of substantially higher time costs. These higher costs are due to changes in employment status due to having hours reduced (28%), being fired or laid off (22%), or choosing to work fewer hours during asthma season (27%). More than two-thirds (70%) of the households reported that a parent had chosen to work part-time or be a stay-at-home parent due to a child's asthma. While the value of time not spent at work for those who have decided not to work in the market is not zero, the value of these time costs could be much greater than the money expenditures for some households. We should look forward to estimates that exploit these data using the best techniques that labor economics has to offer. The fact that some of the household time is unpriced is a barrier that can be overcome.

Hanemann and Brandt note several conceptual and empirical limitations of the household production approach that they believe are threats to the "validity" of the production function approach to estimating WTP for reducing children's asthma morbidity. My view is that their concern is legitimate, but that none of the limitations is a fatal flaw that should prevent them from making a meaningful comparison with the WTP estimates from contingent valuation. My assessment is that a number of high quality studies have been done valuing changes in morbidity using a household production approach despite limitations. A recent example is Dickie's (2005) article on valuing children's health, work for which he received the Georgescu-Roegen Prize. Hanemann and Brandt have excellent information on what goes on in, what to some is, the black box of household production of health. Perhaps it is the richness of the information that makes them hint that a meaningful comparison cannot be made. For example, they report that more than 20% of the households said they had no control of the top ten potential asthma triggers. Viewed differently, the fact that nearly 80% said they had at least some control would be reassuring to many who would apply the household production approach.

In Table 7, nonmarket, household averting and mitigating behaviors is listed based on Hanemann and Brandt's first survey. They include activities such as checking for smog alerts, closing windows, and restricting where the child can play as well as parents giving up smoking cigarettes. These activities can be difficult for researchers to value in dollars for a household production estimate, as they note. This information guided them in designing their contingent market for a hypothetical BreatheRight watch for monitoring. In addition to trying to put dollars values on the activities directly, they might consider adding some contingent time tradeoffs and/or contingent behavior questions to value the nonmarket averting and mitigating behaviors.

All in all, I think that Hanemann and Brandt believe that they have laid the groundwork for a first-class contingent valuation study to estimate the values that parents place on improved control of their children's asthma. I share that belief and think further that they have laid the

2

groundwork for a first-class household production study, and comparison between the two. I look forward to seeing the successful completion of all three.

Comments on "Individual Preferences and Household Choices: The Potential Role of Dependency Relationships" by Mary F. Evans, Christine Poulos, and V. Kerry Smith

Since this project is a work in progress and no paper was available, my comments, unlike good Kentucky bourbon, have aged only one hour. Here are a few quick reactions. One is that the motivation for valuing changes in the environment is not entirely clear. Modeling households as groups in which individuals' roles are treated as economic decisions is offered as a way to gain insights into household relationships. Evans, Poulos, and Smith believe that insights gained could be used as an alternative to focus groups in the development of surveys and survey instruments. Presumably the values that are elicited for changes in environmental quality could depend on which member of the household is asked. My comment is that several good reasons exist for using focus groups and that even if we learn something about household relationships, there are still potentially great benefits to focus groups. If focus groups are held, the marginal cost of exploring household relationships is probably low.

A second comment is that it would be interesting to work through the system of equations for an exogenous change in a quantity, as we often do in environmental economics, instead of a change in price. I am not sure that makes sense based on the brief presentation, but it might be worth considering.

My third and last comment is that the household structures presented did not include the one that I consider the most important, namely a household with at least three individuals. Because I have been father in a household of four in which two parents were involved in raising a child with chronic asthma, I think that an important household configuration has been omitted in the modeling so far. In our household, I think the values you would have elicited in surveys would have been fairly close regardless of which of us you asked. It would be good to have a model that allowed for that situation. What comes from this research could be fascinating.

Comments on "Air Pollution and Asthma: Preliminary Results from a Daily Time Series Study of San Francisco" by Charles W. Griffiths and Nathalie B. Simon

This paper is similar to Hanemann and Brandt's in that deals with morbidity related to asthma and air pollution and studies residents of a city in California. The idea is that since air pollution can trigger asthma attacks, the pattern of filling prescriptions for medications that relieve asthma symptoms, should be influenced by the pattern of air pollution. Air pollution episodes lead to prescription episodes with some lag. My main comment is that the probability of inhaler use, as shown in equation 1, should be broadened to incorporate human behavior. In the context of household production of health, the use of a market input such as an inhaler will depend on exogenous factor such as pollution, weather, and seasonal factors, as Griffiths and

Simon indicate. In addition, however, inhaler use will depend on averting and mitigating behavior such as limiting outdoor exercise, and the myriad of things documented by Hanemann and Brandt. Inhaler use will depend on how well the individual manages controller therapy. Use of these long-term maintenance drugs is an investment that yields a return weeks later when asthma attacks are prevented or reduced in severity.

Weekly count data of prescriptions for quick relief asthma medication for San Francisco for the years 1998-2001 were analyzed using a standard negative binomial regression. Increases in average PM10 increased counts and increases in average minimum temperature decreased counts as expected. Increases in average relative humidity are found to increase prescription counts, a result which Griffiths and Simon did not expect. As one who lives in the Bluegrass Region of Kentucky and associates humidity with lush growth, abundant pollen, thriving mold, and other asthma triggers, I am not surprised by the positive sign on humidity. The result that ozone is associated with a decrease in prescription counts is unexpected. My suggestion, based on my main comment, is to think about inhaler use as determined by individual factors to see if that suggests other variables. If a pollution episode is correctly anticipated, asthmatics are carefully engaged in averting behavior, and long-term maintenance drugs are effective, the increase in use of rescue drugs and increase in prescription counts will be small. This explanation does not distinguish between PM10 and ozone, but perhaps it will lead to a better specification that produces results that are more in line with expectations.

My last comments are that I agree with Griffiths and Simon that they should do more time series diagnostics and more sensitivity analysis of the 180 day window that they use for pollution measures. The windows for PM10 and ozone may be different.

References

Dickie, Mark. "Parental Behavior and the Value of Children's Health: A Health Production Approach" *Southern Economic Journal* 71 (April 2005): 855-872.

Grossman, Michael. "On the Concept of Health Capital and the Demand for Health" *Journal of Political Economy* 80 (March 1972): 223-255.

Yen, Steven T., W. Douglas Shaw, and Mark E. Eiswerth. "Asthma Patients' Activities and Air Pollution: A Semiparametric Censored Regression Analysis" *Review of Economics of the Household* 2 (March 2004): 73-88.

Summary of the Q&A Discussion Following Session I

Reed Johnson, (RTI)
Directing his comment to Mary Evans and Christine Poulos, Dr. Johnson stated, "I thought I followed your model pretty well, Mary, and then you offered the format of the question you were going to ask, which was as I recall: For a given improvement in air quality that would reduce the amount of care giving time, how much would you be willing to pay? I don't quite understand, in the context of the other presentations we've heard today, what you're assuming about the household production function. That is, are you assuming some fixed proportion of time for a decrease in . . . ?"

Mary Evans, (University of Tennessee)
"Yes, at least in the pilot survey the care giving time is going to be exogenous, so that will be determined in the stated preference survey. So we exogenously specify both the initial care giving time and the reduction in the final care giving time."

Reed Johnson
Dr. Johnson continued, "And what are they going to get in terms of improved health? Is it possible to value the time independently of the change in the health outcome experienced in the household?"

Mary Evans
Dr. Evans clarified, "In some household structures, they'll value both the health impact as well as the reduction in time allocation."

Reed Johnson, (RTI)
Dr. Johnson responded, "So if it's not altruistic, they wouldn't value the health outcome."

Mary Evans
"Right. If there's only that income-sharing model, the only way that air quality would impact that particular model is through the mu function essentially—through the individual's share of the household income."

Reed Johnson
Dr. Johnson added, "I guess it wasn't clear to me how the model handles substitution among various household production inputs and how that plays out in the willingness to pay."
_____

Lauraine Chestnut, (Stratus Consulting, Inc.)
Ms. Chestnut addressed Charles Griffiths: "Regarding the negative ozone coefficient— that colder temperatures give higher medicine use—I was just wondering how much you looked into the correlation between that, since the ozone tends to be worse in warmer weather."

Charles Griffiths, (U.S. EPA, NCEE)
Dr. Griffiths responded, "That's actually an excellent point that was raised to me just recently, which is why I kept saying that the way we've modeled ozone is counter-intuitive." Acknowledging that the factor "currently enters in straight," he offered that he "should account for the fact that the ozone effect may be seen only during a certain season." He said, "It may be washed out by the fact that I'm not accounting for the seasonality of the ozone effect."

_____

END OF SESSION I Q&A

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

**PROCEEDINGS OF SESSION II: ISSUES WITH MORBIDITY VALUATION
AND
KEYNOTE ADDRESS BY BRIAN MANNIX**

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

Table of Contents

Keynote Address
Brian Mannix, Associate Administrator, U.S. EPA, Office of Policy, Economics, and Innovation
Introduction by: Al McGartland, Director, U.S. EPA, National Center for Environmental Economics

Session II: Issues with Morbidity Valuation
Session Moderator: William Wheeler, U.S. EPA, National Center for Environmental Research

IOM and Cost Effectiveness
Nathalie Simon, U.S. EPA, National Center for Environmental Economics

Altruism and Environmental Risks to Health of Parents and Their Children
Mark Dickie and Shelby Gerking, University of Central Florida

Is an Ounce of Prevention Worth a Pound of Cure?
Ryan Bosworth and Trudy Cameron, University of Oregon; and J.R. DeShazo, University of California–Los Angeles. Presented by: J.R. DeShazo.

Discussant: Kelly Maguire, U.S. EPA, National Center for Environmental Economics

Discussant: Kevin Boyle, Virginia Polytechnic Institute

Questions and Discussion

U.S. EPA NCER/NCEE Workshop
Morbidity and Mortality: How Do We Value the Risk of Illness and Death?

Washington, DC
April 10-12, 2006

Keynote Address
Brian Mannix, Associate Administrator,
U.S. EPA, Office of Policy, Economics, and Innovation

He [Al McGartland] told you I'm not a lawyer, but he didn't say what my background is—I'm actually a chemist, although you couldn't tell by my career. . . . What I want to do is step back and take a look at the metrics that we use to describe the benefits of mortality reductions that we attribute to environmental regulations. In particular, I want to raise questions about the statistical robustness of the "lives saved" metric that is now commonplace. I should say that years ago I was an advocate for VSL analysis at the beginning of my career, and I encouraged EPA to focus on lives saved. Now that I'm back at EPA (and I did start here at EPA in 1977), I'm surprised at how much progress has been made in incorporating VSL into Agency analyses and decisions. I'm surprised, too, that I'm not very comfortable with where that progress has left us, and I'm *most* surprised to find that the most serious difficulty in my mind turns out not to be with the "V" but with the "SL." That is, economic valuation of mortality benefits is a tractable problem analytically and politically, but figuring out the right metric for mortality benefits is much more problematic. I'll illustrate this with a contrived example:

Suppose on Monday a hospital in a small town publishes a press release announcing that over the busy weekend it had managed to save a dozen lives. The local TV station sends down a camera crew and asks if it can interview a few of the lucky survivors. The ER nurse tells them, "I'm sorry, that won't be possible—he died." "What do you mean— who died?" the reporter asks. "The man who was having the heart attacks," the nurse replies. "We managed to save him 12 times in 13 attempts."

The point of this story is that while we can easily count "lives" or "deaths," we cannot easily count "lives saved." It is not well defined, and it is inherently unbounded. The airbag may save your life in the event that your brakes fail, but how many times has your life been saved when the brakes didn't fail? The number of lives saved during my commute this morning is already beyond my ability to reckon. In some narrow context, we might be able to come up with a workable definition of a "life saved." As a lifeguard, Ronald Reagan would put a notch in a log every time he saved a life, and I don't doubt that it was accurate and meaningful. If he had kept the notched log during his presidency, however, I can't imagine how we would come up with an accurate count—or interpret it if we had one.

I don't believe it is possible to come up with a definition of "lives saved" that is robust, that can be applied to a wide variety of situations, and that can be aggregated in a

statistically meaningful way. The underlying difficulty is that "lives saved" lacks a time dimension. We know that all lives are temporary, and while the valuation problem is quite complex, we are generally in agreement that a longer life is better than a shorter one. If we don't capture the time dimension, we are unlikely to come up with a metric for mortality that is versatile and that behaves well in statistical usage.

There is a standard statistic for measuring longevity that everyone is familiar with—the expected value of the length of life, or life expectancy. It has several advantages in communicating with the public. Everyone has a pretty good idea of what it measures. People also have a good sense of what the units mean. They may have a great deal of difficulty picturing what a "ten to the minus six" risk of death is, but they know how long a minute is and how long 10 years is, and that covers more than six orders of magnitude. This also solves the problem of divisibility—some find it difficult to think about a fraction of a life saved or about the same life being saved multiple times, but they have no trouble dividing time into units of arbitrary size. The public will also have less difficulty attaching a monetary value to changes in life expectancy, even those who cannot imagine attaching a finite value to a life saved.

I should mention here that just yesterday I saw a new Ford commercial in which Bill Ford says, "Every life saved is worth it." I couldn't agree more. The real advantage of using life expectancy, though, is that it is a well-defined and well-behaved summary statistic that reflects mortality risks across an entire population, including risks of all kinds and at all ages without discriminating against any particular subgroup. Let's suppose we're evaluating a range of policy options, all of which have a small marginal effect on mortality risks. If we take as our mandate to maximize life expectancy using limited resources, we can easily solve the problem. We know that the solution will give us a cost-effectiveness criterion, a fixed dollar amount per incremental year of life expectancy. The decision rule would be to adopt those measures that met the cost-effectiveness criterion and to avoid committing resources to those that didn't.

Note that if we use another decision criterion in place of this one, we will get a shorter life expectancy for the same expenditure of resources. If we use a VSL rule, for example, we might save more lives, whatever that might mean, but on average, people will live shorter lives. In most cases I think the two criteria would likely lead to similar outcomes. When they don't, however, we have to ask whether policies that cause a shorter life expectancy can really be said to be improving public health. Similarly, if we adopt maximum life expectancy as the goal, but make adjustments to the metric for age, quality, or willingness to pay, the result will be that people live shorter lives—better, maybe, in some sense, but shorter. I believe this creates a strong presumption for using life expectancy as a standard metric in evaluating regulatory decisions, using a flat VSLY (value of a saved life year) as the cost-effectiveness criterion. As the first-order approximation of mortality benefits, I think this is vastly superior to the VSL approach, and I think anyone advancing some other decision rule needs to explain how we can justify adopting policies that will lead to a shorter life expectancy. I don't rule out that such justifications may exist, but I think we should be cautious in entertaining them.

Al McGartland has pointed out to me that there's a contradiction here. I embrace the use of willingness-to-pay data in figuring out what our cost-effectiveness criterion should be, but I shrink from looking any deeper into the data to find out how it might vary from group to group or person to person. I think this is a contradiction I can live with. An individual, perhaps because he is wealthy, who is willing to pay and does pay much more than average to reduce his own mortality risks, should certainly be able to do so. However, I am not ready to concede that that same individual is entitled to tilt public health measures in his favor simply because he is willing to pay—but does not pay—for them. When writing rules or spending public funds, there is an egalitarian consideration that does not apply when individuals are spending their own money. As analysts we may feel that we can improve the analysis by making adjustments for age or quality or to incorporate the latest willingness-to-pay data, but as a government official I'm reluctant to go very far down that road. In part, that's because I question whether government has any legitimate business making such adjustments, and in part it's because if the government did get into that business, the adjustments would likely be made according to the rules of politics, not necessarily those of economic analysis. So, perhaps a flat VSLY is desirable for the same reason that a flat tax is appealing to some—it minimizes the opportunities for making mischief.

I'll stop there and look for reactions.

(A question and answer session followed.)

IoM Committee Recommendations and Cost Effectiveness Analysis at EPA

Nathalie B. Simon, USEPA

Presentation Prepared for "Morbidity and Mortality:
How do We Value the Risk of Illness and Death
April 10-12, 2006

1

Disclaimer

The views expressed in this presentation are entirely those of the presenter and do not necessarily represent those of the USEPA. No endorsement by the Agency should be inferred.

2

A Very Brief History

- In 2003, OMB released Circular A-4
  - Requiring Agencies to perform health-based cost effectiveness analyses of economically significant rules where health is the primary benefit
  - Prior to this, EPA seldom performed CEAs using health-related quality of life measures.
- Later that year, the Institute of Medicine (IoM) convened a committee, at the request of John Graham and with funding from several Agencies, to address a number of questions on how best to perform CEAs in a regulatory context.
- Specifically, the Committee was asked to:
  - Describe current agency practices in estimating benefits and costs of regulatory actions
  - Review measures currently used in CEAs to aggregate health improvements
  - Develop criteria for choosing among available measures
  - Assess the various measures for data requirements, feasibility, theoretical validity and ethical implications
  - Recommend measures appropriate for federal agency use
- IoM released its long anticipated report in January 2006

3

Recommendation 1: Regulatory CEAs that integrate morbidity and mortality impacts in a single effectiveness measure should use the quality-adjusted life year to represent net health effects.

Recommendation 2: Regulatory analyses should report four measures of cost- effectiveness:

- *Compliance cost per death averted*
- *Compliance cost per life year gained*
- *A health-benefits-only ratio* using the net change in QALYs as the outcome measure.
- *A comprehensive ratio* using QALYs as the outcome measure and incorporating the value of other benefits as offsets to compliance costs.

4

Recommendation 3: The life year and QALY estimates used in regulatory analyses should reflect actual population health as closely as possible

Recommendation 4: Incremental cost-effectiveness ratios are generally the most useful summary measure for comparing different regulatory interventions.

Recommendation 5: In addition to reporting effects in the aggregate, regulatory analyses should report QALY impacts separately for each health endpoint.

Recommendation 6: The reporting of all CEA results should be accompanied by information on related uncertainties and non-quantified effects.

5

Recommendation 7: Regulatory analyses should not assign monetary values to estimates of health-adjusted life years as a method for valuing health states.

Recommendation 8: The regulatory decision-making process should explicitly address and incorporate the distributional, ethical, and other implications of a proposed intervention along with the quantified results of BCA and CEA.

Recommendation 9: Policy makers and program administrators should work to ensure the substantive involvement of a broad range of individuals and groups at all stages of policy development for regulating risks.

6

Recommendation 10: A high research priority should be improving the data used to assess the health risks (effects on incidence of particular types of illness, injuries, and deaths, and the duration and latency of effects) addressed by regulatory actions.

Recommendation 11: The Department of Health and Human Services (DHHS) and other federal agencies should collect HRQL information through routinely administered population health surveys and other major studies and data collection efforts related to risk assessment and monitoring.

Recommendation 12: DHHS should coordinate, with the involvement of federal regulatory offices and agencies, the development of an integrated research agenda to improve the quality, applicability, and breadth of HRQL measures for use in regulatory CEA. The Committee identifies the following areas as priorities for research:

- Current elicitation methods such as the standard gamble and time trade-off, while theoretically well founded, may be difficult for respondents to understand and prone to generate inconsistent responses. Research to facilitate improved methods is needed. In addition, methods for eliciting societal values for investments in health (in contrast to individual preferences for health states), such as person trade-off techniques, should also be investigated.

- Methods for measuring children's health-related quality of life, including characterization of the impact of illness and injury and the valuation of these impacts, need continued development and refinement.

- Methods to correlate QALY estimates based on different generic HRQL indexes should be developed so that estimates from different underlying valuation studies are consistent and can be used in the same analysis.

Some Key Issues for EPA

- What do we do if BCA and CEA produce different rankings of policies/programs…or if the four CEA ratios produce different rankings?
- We sometimes have policies that reduce morbidity in one population and prevent deaths in another. How do we combine these effects using a single HRQL index?
- Are QALYs biased since they undervalue health gains to the disabled, elderly and chronically ill?
- How should we report uncertainty? Do we somehow develop ranges of QALYs as we've been asked to do with benefits estimates?
- What do we do when we are unable to quantify the impacts but we know what they are?
- How will the CEA results be used? Will they be used to construct league tables? These may lead to systematic bias away from environmental policies to direct health policies simply because we are unable to quantify some of the environmental effects.

9

Next Steps for EPA

- The Agency is in the process of updating its *Guidelines for Preparing Economic Analysis*
- A cross-office workgroup was recently convened to write a chapter on CEA for the Guidelines
  - Working through IoM recommendations
  - A public review draft is anticipated by end of calendar year.

For those of you interested in learning more about the IoM report:

see "recent reports" at www.iom.edu

10

Altruism and Environmental Risks to Health of
Parents and their Children[*]

Mark Dickie
and
Shelby Gerking**

Department of Economic
University of Central Florida
Orlando, FL 32826

June 2, 2006

Running Title: Altruism and Environmental Risks

JEL Codes: Q51, Q58, D13, D64.

** Corresponding author. E-mail: Shelby.Gerking@bus.ucf.edu. Mail: Department of Economics, University of Central Florida, P.O. Box 161400, Orlando, FL 32816-1400. Phone: (407) 823-4729. Fax: (407) 823-3269.

**Altruism and Environmental Risks to Health of
Parents and their Children**

ABSTRACT

This paper tests an equilibrium condition from a model that incorporates: (1) altruism of parents toward their young children and (2) household production of latent health risks. The model demonstrates that an altruistic parent's marginal rate of substitution between an environmental health risk to herself and to her child is equal to the ratio of marginal risk reduction costs. Econometric estimates support this prediction based on data from a stated preference study involving 488 parents of children aged 3-12 years. This outcome implies that parents reallocate family resources to at least partly offset the effectiveness of public programs that aim to reduce their children's environmental risks.

Key words: Altruism, household production, environmental risk, child health.

**Altruism and Environmental Risks to Health of
Parents and their Children**

1.   Introduction

Special protection of young children from environmental hazards has become a

worldwide priority in government policies to improve human health.[1] Effectiveness of these

measures depends on what steps parents voluntarily take to keep children out of harm's way. If

parents are naive about hazards, do not care about their children, or lack the resources to protect

their health, implementation of well-designed public policies to increase protection of children

may have the intended effect. On the other hand, if parents are informed, altruistic, and

sufficiently well off financially, measures aimed at increasing protection of their children from

particular hazards will be offset to some extent as parents redistribute family resources. In any

case, the fundamental tension between altruism and self-interest in family exchange looms as a

crucial behavioral factor determining the effectiveness of government policies to protect

children's health.

What is known about altruism in families? Several prominent empirical studies (e.g.,

Cox and Rank 1992, Altonji, Hayashi, and Kotlikoff 1992, 1997, Laitner and Juster 1996) do not

support the implication of altruism for transfer-income derivatives in examining inter-household

financial transfers between parents and adult children. Other papers (e.g., Liu *et al*. 2000,

Jenkins, Owens, and Wiggins 2001, Nastis and Crocker 2003, Agee and Crocker 2004, Dickie

and Messman 2004) look at how parents protect themselves and their pre-teenage children from

---

[1] For example, Executive Order 13045 (Federal Register, 1997) directs U.S. federal executive branch agencies to assign a high priority to addressing health and safety risks to children, coordinate research priorities on children's health, and ensure that their standards take into account special risks to children. The U.S. Environmental Protection Agency has formulated a seven-step strategy to protect children's health (U.S. EPA 1996). Some of the more visible federal decisions in which protection of children's health figured prominently include tightening of air quality standards for ozone and particulate matter and implementation of the 1996 Safe Drinking Water Act Amendments and the 1996 Food Quality Protection Act. Scapecchi (2006) summarizes similar efforts undertaken in other countries.

environmental and other hazards. In this branch of the literature, altruism is sometimes

mentioned as a possible parental motivation, but equilibrium conditions implied by altruism are

not tested.

This paper tests a model of altruistic family behavior (Becker 1974, 1981 and Barro

1974) that incorporates household production of latent health risks. The model demonstrates that

the parent's marginal rate of substitution between risks faced by herself and her child is equal to

the ratio of marginal risk reduction costs. This prediction is tested using survey data on skin

cancer risks faced by 488 parents in Hattiesburg, MS and their biological children between the

ages of 3 and 12 years. Marginal rates of substitution are obtained from stated preference values

for a hypothetical sun lotion. While stated preference valuation remains a controversial method

of obtaining willingness to pay for reduced environmental risk, its application here supports

consistent estimation of the desired marginal rates of substitution because of the way the survey

(described more fully later on) is designed. Test outcomes support the model and imply that

parents are altruistic toward their young children.

2.   *Conceptual Framework*

2.1   *Model*

This subsection presents an extension of Becker's (1981) model of altruism that

incorporates household production of latent health risks. The model envisions a "family"

composed of one altruistic parent and one child. Because only one child is included in the

model, the analysis focuses on how parents allocate resources between themselves and their

children, rather than on how parents make tradeoffs among different children. By including only

one parent in the model, a unitary perspective is adopted in which possible divergent interests

between parents in a family are not considered. Although the unitary model has been rejected in

several empirical tests (e.g., Lundberg et al. 1997), tests presented in Section 4 reveal no significant differences in valuation of latent health risks between fathers and mothers.

To facilitate treatment of latent health risks, assume that the parent has two periods of life remaining while the child has three. During the present period ($t = 0$), the parent receives all family income, purchases market goods for her family, and behaves as a paternalistic altruist in that she derives utility from her own consumption as well as from the combination of goods that she provides to her child.[2] Thus, the parent allocates goods to the child according to her own views as to what is best and disregards the child's preferences (if any) except in situations in which they are congruent with her own. In period $t = 1$, the child will be an adult with his own income, which the parent may supplement with transfers, and will make his own consumption decisions. In this period, the parent will derive utility from her own consumption and may also derive satisfaction from the level of utility achieved by the child. The model therefore envisions that the parent's altruism may switch from paternalistic altruism to the more all-encompassing concern for the child's well-being considered by Altonji, Hayashi, and Kotlikoff (1997) after the child is mature enough to exhibit well-defined preferences and the parent can no longer dictate the combination of goods that the child will consume.[3] In the third and final period ($t = 2$), the child continues to receive income and purchase market goods while the parent is deceased.

The survey, described more fully in Section 3, elicits willingness to pay to reduce two latent environmental health risks facing both the parent and the child. In the model, these two risks are denoted a and b. To consider a latency period that is longer for the child than for the parent, assume that the events at risk may occur in the last period of either individual's life.

---

[2] Paternalistic altruism is more fully discussed by Jones-Lee (1991, 1992)
[3] Both types of altruism are incorporated into the model to assist in clarifying the interpretation of statistical tests presented in Section 4. All-encompassing concern for another's well-being has also been termed "benevolence" (Bergstrom 2006) or "pure" altruism (Jones-Lee 1991, 1992).

Constraining the lifetime risk to lie in a single period simplifies the task of communicating changes in risk to survey respondents (see Section 3). Perceptions of the jth latent risk to the ith person are denoted R_i^j, where superscript j distinguishes between the two risks (a and b) while subscript i distinguishes the parent (p) from the child (k). Perceived lifetime risks are influenced by the use of market goods that otherwise have no utility:

$$R_p^j = R_p^j(G_{p0}^j, G_{p1}^j),$$
$$R_k^j = R_k^j(G_{k0}^j, G_{k1}^j, G_{k2}^j), \qquad j = a, b.$$

(1)

where G_{it}^j denotes individual i's use in period t of a market good affecting the jth risk. Simplifying assumptions here are that: (1) the risk production functions do not shift over time, (2) the child when grown is assumed to share his parent's assessment of both risks, and (3) marginal products of the G_{it}^j are strictly negative in both production functions.

When the child begins to make his own consumption decisions as an adult in period $t=1$, he will maximize his lifetime utility given by $U_k(C_{k0}, C_{k1}, C_{k2}, R_k^a, R_k^b)$ subject to his perceived risk production functions given in equation (1), the choice of $(C_{k0}, G_{k0}^a, G_{k0}^b)$ that already will have been made by the parent, and his lifetime budget constraint,

$T + y_{k1} + (1+r)^{-1} y_{k2} = C_{k1} + P^a G_{k1}^a + P^b G_{k1}^b + (1+r)^{-1}[C_{k2} + P^a G_{k2}^a + P^b G_{k2}^b]$. Here and in

equations (2) and (3), variables y_{it} and C_{it} respectively denote individual i's income and consumption of an aggregate market good in period t, T denotes the income transfer from parent to child in period $t=1$ ($T \geq 0$), r denotes the market interest rate and P^j denotes the market price of the protective good affecting the jth risk.

In period $t = 0$ the parent maximizes the utility function

$U_p(C_{p0}, C_{p1}, C_{k0}, R_p^a, R_p^b, R_k^a, R_k^b) + \eta U_k^*(C_{k0}, G_{k0}^a, G_{k0}^b, T, y_{k1}, y_{k2}, r, P^a, P^b)$

(2)

subject to the four perceived risk production functions in equation (1), the restriction $T \geq 0$ and her lifetime budget constraint

$$
\begin{aligned}
y_{p0} + (1+r)^{-1} y_{p1} &= C_{p0} + C_{k0} + P^a (G_{p0}^a + G_{k0}^a) + P^b (G_{p0}^b + G_{k0}^b) \\
&\quad + (1+r)^{-1} [C_{p1} + T + P^a G_{p1}^a + P^b G_{p1}^b],
\end{aligned}
\tag{3}
$$

where $\eta \geq 0$ is the weight the parent places on the child's lifetime utility and $U_k^*(\bullet)$ denotes the indirect utility function from the child's maximization problem. When $t = 0$, the parent chooses quantities of all market goods that she and her young child use and when $t = 1$, the parent makes these choices only for herself while deciding how much income to transfer to her child.

The parent's paternalistic altruism in period $t = 0$ is reflected in her concern for her child's present consumption and his risk . If $\eta = 0$, the parent has no further concern for the child in future periods and will not care how his future choices may affect the lifetime risk he ultimately faces. If $\eta > 0$, the parent continues to care about the child in the future, but she exhibits benevolence or all-encompassing altruism in that she respects the child's adult preferences and cares about his overall level of well-being rather than the specific bundle of goods he consumes.

First order conditions[4] for period $t = 0$ quantities imply that for $j = a, b$

$$
\begin{aligned}
\partial U_p / \partial C_{p0} &= \partial U_p / \partial C_{k0} + \eta \partial U_k / \partial C_{k0} \\
(\partial U_p / \partial R_p^j)(\partial R_p^j / \partial G_{p0}^j) &= (\partial U_p / \partial R_k^j)(\partial R_k^j / \partial G_{k0}^j) + \eta (\partial U_k / \partial R_k^j)(\partial R_k^j / \partial G_{k0}^j).
\end{aligned}
\tag{4}
$$

---

[4] These equations make use of the relationships $\partial U_k^* / \partial G_{k0}^j = (\partial U_k / \partial R_k^j)(\partial R_k^j / \partial G_{k0}^j)$. Equations (4) and (5) also make use of the assumption that the parent exhibits paternalistic altruism only in period $t=0$. Thus her paternalistic altruissm encompasses concern for how her present choices affect her child's risk but does not extend to concern for how his future choices may alter his risk. Any concern for the child in future periods is reflected by $\eta > 0$, not by $\partial U_p / \partial R_k^j$. This assumption means that the parent does not have to consider the dependence of her child's future choices on her decisions today. A more formal analysis of this point is available on request.

Thus, in period $t = 0$, the model predicts the familiar result that if both individuals consume C

and G in positive quantities, the parent's marginal rate of substitution between the child's

consumption of C (G) and her own consumption of C (G) is equal to unity.[5] This outcome holds

independently of the magnitude of η, the weight that the child's utility receives in the parent's

utility function, and also holds if the parent exhibits either type of altruism. If instead the parent

exhibits neither type of altruism (i.e., is not an altruist toward the child), then these marginal

rates of substitution equal zero. If the parent exhibits either or both types of altruism toward the

child but does not care about her own consumption or about the level of risks that she faces, then

these marginal rates of substitution are arbitrarily large.

In periods $t = 1$ and $t = 2$, first order conditions imply that

$$
\begin{aligned}
&\partial U_p / \partial C_{p1} = \lambda_p (1+r)^{-1} \\
&\partial U_k / \partial C_{kt} = \lambda_k (1+r)^{1-t} && t = 1,2 \\
&(\partial U_p / \partial R_p^j)(\partial R_p^j / \partial G_{p1}^j) = \lambda_p P^j (1+r)^{-1} && j = a,b && (5)\\
&(\partial U_k / \partial R_k^j)(\partial R_k^j / \partial G_{kt}^j) = \lambda_k P^j (1+r)^{1-t} && j = a,b \quad t = 1,2 \\
&\eta \lambda_k = \lambda_p (1+r)^{-1} && if \ T > 0.
\end{aligned}
$$

Equation (5) shows that if $\eta > 0$ and if $T > 0$, then in period $t = 1$ the parent's marginal rate of

substitution between the child's consumption of C (G) and her own consumption of C (G) also is

equal to unity. In the case in which $\eta > 0$, therefore, transfers from the parent to child ensure

that the parent's marginal rate of substitution between the child's consumption of market goods

and her own consumption of market goods is equal to unity in all periods in which both

individuals are alive. If $\eta > 0$, but $T = 0$ (as may occur in period $t = 1$ if the child is rich and the

parent is poor) then the parent's marginal rates of substitution between her child's consumption

---

[5] Throughout the paper, the convention adopted for calculating marginal rates of substitution is that the parent's marginal utility of the child's consumption is in the numerator and the parent's marginal utility of her own consumption is in the denominator.

and her own consumption are positive, but in general are not equal to unity because $\eta \lambda_k \neq \lambda_p (1+r)$. On the other hand, if the parent is a paternalistic altruist only and has no concern for the child's well-being after period $t = 0$ has ended ($\eta = 0$), then in period $t = 1$ the parent's marginal rates of substitution between her child's consumption and her own consumption are equal to zero. Finally, just as in period $t = 0$, if the parent cares about her child's well-being but not about her own consumption of market goods, then her marginal rates of substitution between the child's consumption and her own consumption become arbitrarily large.[6]

The empirical analysis presented in Section 4 looks at risk reduction, not consumption of G^j. So, in period $t = 0$, the first order equation for G^j in (4) is rewritten as equation (6) to show that when corner solutions are set aside, the parent's marginal rate of substitution between risk to her child and risk to herself is equal to the ratio of marginal products of a risk-reducing market good that both individuals consume.

$$\frac{(\partial U_p / \partial R_k^j) + \eta(\partial U_k / \partial R_k^j)}{(\partial U_p / \partial R_p^j)} = \frac{(\partial R_p^j / \partial G_{p0}^j)}{(\partial R_k^j / \partial G_{k0}^j)} = \frac{MC_{k0}^j}{MC_{p0}^j} \qquad j = a,b. \qquad (6)$$

The ratio of marginal products, in turn, equates to the ratio of present value marginal costs because the price per unit of G^j is the same no matter who uses it.

Equation (5) also implies that each individual equates the present-value marginal costs of risk reduction over time, provided that risk production functions are constant over time. Thus, in

---

[6] Equation (5) also implies that when the parent and child consume positive quantities of all goods in all periods, the inter-temporal marginal rate of substitution between consumption of C (G) in period $t + 1$ and consumption of C (G) in period t equals the discount factor $(1 + r)^{-1}$ for both the parent and the child. The inter-temporal marginal rate of technical substitution between risk-reducing goods in different periods likewise equals the discount factor for both individuals. If $\eta > 0$ and if $T > 0$, then the parent's marginal rate of substitution between her child's consumption of C (G) in period $t + 1$ and her own consumption of C (G) in period t is equal to the discount factor as well.

period $t = 1$, the present-value marginal cost of risk reduction for the parent will be the same as in period $t = 0$, and the present-value marginal cost of risk reduction will be the same for the child in periods $t = 1$ and $t = 2$. In addition, if $\eta > 0$ and $T > 0$, then the marginal costs of risk reduction for the child are the same in all three periods.[7] Evidently, the parent's all-encompassing concern for the child's well-being together with her monetary transfers enables her to choose marginal cost of risk reduction values that the child will use for the rest of his life. In consequence, if $\eta > 0$ and $T > 0$

$$\frac{\eta(\partial U_k / \partial R_k^j)}{(\partial U_p / \partial R_p^j)} = \frac{(\partial R_p^j / \partial G_{p1}^j)}{(\partial R_k^j / \partial G_{kt}^j)} = \frac{MC_k^j}{MC_p^j} \qquad j = a, b \qquad t = 1, 2. \qquad (7)$$

On the other hand, this marginal rate of substitution equates to zero if $\eta = 0$ and will not equate to the marginal cost ratio if either $T = 0$ or if the parent does not care about risk to herself.

Together, equations (6) and (7) imply that if $\eta > 0$ and $T > 0$, and both the child and parent consume positive quantities of all market goods in all periods when they are alive, then the parent's marginal rate of substitution between her child's and her own latent risk equals the ratio of present-value marginal costs of reducing risk in any period. Three further implications of equations (6) and (7) are that even if the parent is a paternalistic altruist in period $t=0$ and if $\eta > 0$ and $T > 0$: (1) the ratio of marginal risk reduction costs for the child and the parent is not expected to equal unity because the technologies used to produce perceived risk reduction may differ and, even if the technologies are the same, levels of perceived risk faced by the two people may not be the same, (2) for either individual, the ratio of marginal costs for reducing the first risk need not equal the ratio of marginal costs for reducing the second risk, and thus (3) for either

---

[7] $MC_p^j = (1+r)^{-t} P^j / (\partial R_p^j / \partial G_{pt}^j), t = 0, 1,$ and $MC_k^j = (1+r)^{-t} P^j / (\partial R_k^j / \partial G_{kt}^j), t = 0, 1, 2.$

individual, the marginal rate of substitution between the two types of risks equals the

corresponding ratio of marginal costs in reducing the two risks.

Empirical estimates described in Section 4 test the null hypothesis that the equilibrium

conditions stated in equations (6) and (7) hold. This test is facilitated by considering percentage

risk changes rather than changes in risk by absolute amounts. For instance, when the parent and

child experience the same percentage reduction in a risk, the ratio of marginal products in

equation (6) equals the ratio of initial risk levels, as illustrated below for period $t = 0$.[8]

$$\frac{(\partial R_p^j / \partial G_{p0}^j)}{(\partial R_k^j / \partial G_{k0}^j)} = \frac{R_p^j}{R_k^j} \qquad j = a, b$$

Thus, in this case, as shown in equation (8), the parent's marginal rate of substitution between

equal percentage risk changes for herself and for the child equates to unity.

$$\frac{[(\partial U_p / \partial R_k^j) + \eta(\partial U_k / \partial R_k^j)]R_k^j}{(\partial U_p / \partial R_p^j)R_p^j} = \frac{(\partial R_p^j / \partial G_{p0}^j)/R_p^j}{(\partial R_k^j / \partial G_{k0}^j)/R_k^j} = 1 \qquad j = a, b \qquad (8)$$

If $\eta > 0$ and $T > 0$, then the corresponding condition will hold for periods $t = 1$ and $t = 2$, as

shown in equation (9).

$$\frac{\eta(\partial U_k / \partial R_k^j)R_k^j}{(\partial U_p / \partial R_p^j)R_p^j} = \frac{(\partial R_p^j / \partial G_{p1}^j)/R_p^j}{(\partial R_k^j / \partial G_{kt}^j)/R_k^j} = 1 \qquad j = a, b \quad t = 1,2 \qquad (9)$$

Evidence that equation (8) holds supports the notion that parents are altruistic toward

their children, but does not indicate whether parents are paternalistic altruists only, whether

parents only exhibit the broader type of altruism associated with $\eta > 0$ and $T > 0$, or whether

---

[8]This outcome also yields a useful corollary for transferring adult morbidity estimates to children when equal proportionate changes in risk to both groups are considered. If the parent and child experience the same percentage reduction in risk, the ratio of marginal products in equation (4) equals the ratio of initial risk levels. This means that the ratio of the parent's willingness to pay to reduce risk to the child to the parent's willingness to pay to protect herself equates to this ratio of risks. The ratio of actual risks faced might be estimated in some cases using existing health science and biomedical information. The ratio of perceived risks might be established by studies of parents' perceived risks to children and to themselves.

parents exhibit both types of altruism. Evidence that equation (9) holds, on the other hand, says nothing about paternalistic altruism, but supports the notion that $\eta > 0$ and $T > 0$. Evidence supporting equations (8) and/or (9) does not indicate whether η or the provisions the parent makes for the child (C_{k0}, G_{k0}^{j}, T) are large or small. As discussed more fully in Section 4, tests applied do not distinguish between paternal and all-encompassing altruism and do not identify the value of η if $\eta > 0$.[9]

2.2   Policy implications

The model developed in the previous subsection suggests that effectiveness of government programs aimed at reducing risk through behavior modification will be compromised to some extent because they motivate parents to reallocate family resources, as illustrated by the following three examples.[10] First, suppose that in a country composed of M identical families[11], the government initiates an administratively costless program in time period $t = 0$ to provide special protection of children from risk, as envisioned by Executive Order #13045 (Federal Register 1997) in the United States and by similar policies pursued by other countries (Scapecchi 2006). Assume that: (1) the government has access only to the "family technology" for risk reduction described by equation (1), (2) the program provides the parent

---

[9] The model presented can be modified or extended in a variety of ways without altering the basic result that the altruistic parent's marginal rate of substitution between her child's and her own risk equals the ratio of marginal costs of risk reduction. For example, a discounted expected utility model in which individuals produce risk but probabilities condition expectations rather than utility itself also implies equality between the parent's marginal rate of substitution and the ratio of risk-reduction costs.

[10] Although the model does not address issues related to government risk information provision or how parents might respond to such information, it is at least plausible that such programs might be more effective than behavior modification programs. Also, along these lines, note that if in addition to paternalistic altruism, $\eta > 0$ and $T > 0$, parental learning about risks will be retained by the child through adulthood in the sense that his marginal costs of avoiding a risk are equated through all periods of his life. In this situation, parental learning may be passed to future generations as well, but a formal investigation of this matter would require reformulating the model to allow the child to have children of his own as, for example, in Becker (1974).

[11] Further examples based on heterogeneity of parent incomes, two-parent families, and families with multiple children easily can be constructed based on those presented below. Similar examples also can be developed for models where government policy operates by determining the level of an environmental hazard that affects child and/or parent risk rather than by providing G, although in that case the rate of substitution between G and the environmental hazard in the risk production functions must be considered.

with an extra unit of G earmarked for the child's use, (3) the program is financed by levying a tax on each parent in the amount of $\$P$, the price per unit of G,[12] and (4) and parents exhibit one or both types of altruism. As long as prices of market goods and the parent's income remain unchanged, parents and children in each family end up consuming the same quantities of all goods as before. In consequence, the program does not alter behavior and has no effect on the level of risk faced by either person.

Second, suppose instead that the government program sets out to protect everyone (i.e., both adults and children) from risk by giving each family one unit of G for either person to use, rather than earmarking it for the child's use. In this situation, each family simply "purchases" one unit of G for $\$P$ from the government rather than from the private market. Again, if incomes and market prices remain unchanged and parents behave altruistically, each family member consumes the same quantities of C and G as before so that the program has no effect on behavior or on risk levels faced by either parents or children.

Third, suppose that the government is more efficient than families in lowering risk, perhaps because of economies of scale in providing risk reduction. In this case, each family might receive more than one unit of G in return for the tax payment of $\$P$, thereby experiencing the equivalent of an increase in income. Pure paternalistic altruists would then divide the income increase between their own consumption of C and G in periods $t = 0$ and $t = 1$ and their child's consumption of these goods in period $t = 0$, with the increment in G allocated between the parent and the child so that the parent's marginal rate of substitution between risk to the child and risk to herself remained equal to the ratio of marginal costs of risk reduction. If in addition to or instead of paternalistic altruism, parents also exhibit all-encompassing altruism

---

[12] Becker (1981, Chapter 8) presents a closely related example with extended discussion in the context of an income transfer between an altruistic person and his/her spouse.

with $\eta > 0$ and $T > 0$, more substitution possibilities arise because a portion of the income increase could be transferred to the child for use later in his life. Thus, while the program could succeed in lowering risk, the efficiency gain is diffused because both family members now consume more of all goods in the present period and possibly in future periods.

3.   Data and Experimental Design

3.1   Background

Field data were collected from parents of pre-teenage children during summer of 2002 using a self-paced, interactive, computerized instrument.[13] An early version of this instrument was used in a pilot study of parents' willingness to pay to reduce perceived skin cancer risks (Dickie and Gerking 2003). Two subsequent versions of the instrument were pre-tested and de-briefing sessions with pre-test participants guided development of the final version. Parents who participated in this study were residents of the Hattiesburg, MS metropolitan statistical area and were initially identified by random digit dialing. When calls reached adults, interviewers asked whether they had at least one biological child between the ages of 3-12 living at home, and whether they were willing come to the University of Southern Mississippi to participate in a federally funded study of health risks to parents and their children. Biological children were singled out for inclusion in the study because skin cancer risk is partly determined by genetic characteristics inherited from parents (e.g., fairness of skin and sensitivity of skin to sunlight). Parents were offered a $25 payment for participating in the study.[14]

---

[13] A more complete description of these data is provided in Dickie and Gerking (2006).

[14] Approximately 30% of calls to presumed working residential numbers yielded no contact with an adult after three attempts at different times of day and days of the week. In 64% of cases in which a call reached an adult, the adult declared that the household did not meet eligibility requirements (had no biological children aged 3-12 living at home). Parents agreeing to participate in the study constituted 3.5% of working residential numbers, 5% of contacts with adults, and 14.3% of contacts with adults who did not declare the household ineligible. Finally, 68% of persons agreeing to participate completed the instrument.

The sample consisted of 610 parents; children did not participate.[15] Of the parents, 75% were white, 20% were African-American, and 5% were members of other races.   Data from the 122 African-American parents are not considered further in this paper (but are analyzed in Dickie and Gerking 2006) because blacks face low levels of risk and therefore have fewer incentives than whites to think about precautions against solar radiation exposure and how their own risk might differ from that of their children. Of the 488 non-black parents, 25% were male, 75% were under the age of 40, mean household income was $60,000 per year, 83% were married, and 60% worked full time. Parents generally were aware of skin cancer: 83% knew someone personally who had been diagnosed with this disease, 18% knew of someone (public figures, friends, or relatives) who had died from skin cancer, and 82% had considered the possibility that one of their children might get skin cancer. At an early stage in the interview, one biological child aged 3-12 of each parent was randomly selected (if there was more than one in this age range) and designated as the sample child. Questions asked mainly focused on the parent and the sample child. Half (50.4%) of the sample children were male and the average age of sample children was 7 years.

3.2.   *Elicitation of Risk Beliefs*

Two types of risk to both parents and children were elicited: (1) the unconditional risk of getting skin cancer during one's lifetime and (2) the conditional risk of dying from this disease given that it occurs.[16] Parents Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein made

---

[15] Responses from 25 parents were disregarded either because they did not answer all questions (21 parents) or because they did not follow instructions given by the experiment administrator (4 parents).

[16] The ability of respondents to understand the risk concepts presented and to clearly distinguish between these two types of risk was a concern from the beginning of the study because of difficulties people have thinking about probabilities (Slovic, Fischhoff and Lichtenstein 1985). This concern was amplified for the present study because few previous surveys have dealt with compound risks. In de-briefing sessions conducted after the pre-tests, the meaning of the morbidity risk and conditional death risk questions were extensively discussed with participants. Participants suggested a number of wording changes in the questions, but through this discussion and through their direct statements, they demonstrated facility with the risk concepts involved.

preliminary assessments of lifetime skin cancer risk using an interactive scale similar to that used

by Krupnick *et al.* (2002) and Corso, Hammitt, and Graham (2001). The scale, which underwent

a number of design changes based on the pre-tests, depicted 400 squares in 20 rows and 20

columns and all 400 squares were initially colored green. Parents changed green squares to red

ones to represent amounts of risk. Before using the scale to estimate skin cancer risk, parents

practiced using the risk scale for an unrelated event (a possible auto accident) and were told

about the meaning of "chances in 400". Also, they were told to consider only the chances of

getting skin cancer (or of getting it again if they had already had it), rather than how serious the

case might be. Parents then used the risk scale to estimate lifetime chances of getting skin

cancer, for themselves and then for their sample child. Frequency distributions of these

responses presented in Table 1 indicate considerable variation in risk estimates with some

parents believing that skin cancer is highly unlikely and a smaller number of parents believing

that skin cancer is inevitable. Risk estimates tended to pile up at the 5, 10, 15, etc. percent

marks.

As shown in Table 2, parents estimated that their own lifetime risk of getting skin cancer

exceeded that of their sample child (26.9% vs. 22.5%). The null hypothesis that mean perceived

skin cancer risks are equal for parents and children is rejected at the 1% level in a matched-

samples test. This outcome may reflect a number of factors possibly including parents' beliefs

that they take greater precautions to protect their children from skin cancer risk than their parents

did in an earlier period when less was known about the hazards of solar radiation exposure.

Parents also appear to have overestimated skin cancer risk. Ries *et al.* (1999) found that whites

have a lifetime chance of 21% of getting either melanoma or non-melanoma skin cancer. The

fact that the survey introduced the possibility of getting skin cancer again if the parent had

already had it does not appear to be an important complicating factor in this regard. Sample parents are relatively young and 4.3% reported having been previously diagnosed with this disease.

Parents were given an opportunity to revise their beliefs about the chances of getting skin cancer after receiving information about this disease. They were told that: (1) according to the National Cancer Institute, the average person in the United States has a lifetime risk of getting skin cancer of 18% and (2) a person's risk may differ from this average because of skin color and sensitivity to sunlight, family history of skin cancer, amount of time spent in direct sunlight, experience with sunburns, and use of sun protection products. Parents were questioned about observable skin characteristics, sun exposure history, and use of sun protection products both for themselves and their sample children. Over 90% of parents and 97% of children use sun protection products such as sun lotion. Children use sun protection products a greater fraction of the time that they are outside and use products with a higher sun protection factor than do their parents (Table 3). About 40% of parents revised their own lifetime risk estimates, but upward and downward revisions balanced to yield zero mean revision. Revised risk estimates for children were on average 2 percentage points lower than initial risk estimates.

To obtain a rough indication of beliefs about latency of skin cancer risks, parents were asked, "Suppose you do get skin cancer sometime in the future. At what age do you think you would get it for the first time (or for the next time if you have already had it)?" Responses to this and a parallel question about the children are summarized in Table 4. About 65% of parents saw skin cancer as a disease that would strike them or their children at age 50 or later. Based on the midpoints of the age intervals listed in Table 4, parents on average expected that skin cancer, if it occurs, would strike them at age 53 or their children at age 55. Comparing expected age at onset

16

to current age, the average implied latency period is 18 years for parents and 48 years for children, a difference that is significant at the 1% level. These rough measures of perceived latency suggest that parents see skin cancer as a disease that occurs later in life and see their children's risk as lying farther in the future than their own.

Parents also provided estimates of mortality risk from skin cancer both for themselves and for their sample children assuming a doctor had diagnosed this disease. Parents were unaware that they would be asked about the likelihood of dying from skin cancer when they answered the previously described questions about getting this disease.[17] Parents provided their perceptions of conditional mortality risk of skin cancer given a diagnosis of this disease using the previously described risk scale. Table 1 presents the frequency distribution of responses. About two-thirds of parents believed that their conditional risk of death given a diagnosis of skin cancer is 10% or less and about three-fourths of parents believed that if similarly diagnosed, their sample child's conditional risk of death is 10% or less. Many parents felt that the conditional risk of death is less than 5% both for themselves and for their children. This outcome suggests that parents were aware that skin cancer is seldom fatal. Parents reported higher mean conditional death risk estimates for themselves (12.1%) than for their sample children (9.4%), a significant difference at the 1% level.

3.3   *Experimental Design and the Choice Experiment*

Parents valued risk reductions by expressing willingness to pay for a hypothetical sun lotion.[18] The product was described using labels (see Figure 1 for an example) designed to look like those on bottles of over-the-counter sun lotions. Except for differences in the type and

---

[17] Respondents were instructed not to look ahead or to go back to previous questions but rather to see the experiment administrator if they needed to correct a mistaken answer. Data from 4 respondents who did not comply with this instruction were among the previously mentioned observations that were deleted.

[18] This approach also was used in a recent cross-country study of skin cancer risks (see Brouwer and Bateman 2005).

amount of skin cancer protection offered, the labels were identical in all respects to control for

other possible motivations for purchasing sun lotion, such as to prevent sunburn or to get a

suntan and to guard against aging or wrinkling of skin (see Dickie and Gerking 1996). Eight

labels were used in the study: Four labels varied reductions in risk of getting skin cancer

(10%/50% for parent/child) and four labels varied reductions in conditional death risk (10%/50%

for parent/child).[19]   As demonstrated in Section 2, use of percentage changes simplifies the

econometric tests. Use of percentage changes in risk also has an advantage over presenting

absolute risk reductions in that the post-treatment risk levels always are non-negative.[20]

Each parent was randomly assigned two of the eight labels and asked for willingness to

pay for each.[21] One of the assigned labels offered reduced risk of getting skin cancer and the

other offered reduced conditional death risk from skin cancer. Labels were presented one at a

time in randomized order. After parents were given time to read a label as if considering buying

the product for the first time, they were shown their previously marked risk scales both for

themselves and their children showing the level of perceived risk the parent originally indicated,

---

[19]The survey presents exogenous changes in risk to avoid issues that arose in a previous study (Dickie and Gerking 1996) in which risk changes were treated as endogenous. In the earlier work, labels were presented without the stated risk changes and respondents indicated the amount by which risk would be reduced if the product were used as directed. Survey participants, however, expressed little confidence in their response to this question and responses obtained were unavoidably correlated with unobserved participant characteristics. In the present context, telling parents what to believe about the magnitude of risk change is at least arguably better than asking a difficult question. Also, random assignment of labels means that risk changes are orthogonal to respondent characteristics. Nonetheless, because changes in risk actually are endogenous, interpretation of the econometric estimates presented in the next section must necessarily be guarded.

[20] Data on actual purchases of currently marketed sunscreen lotions would not support valuation of the two risks separately from other motivations for using sunscreen (Dickie and Gerking 1991, 1996) and would not reflect random assignment of exogenous risk changes. These two features of the field study are critical for estimating the marginal rate of substitution.

[21] Means of the four perceived risks, family income, number of children in the family, and age and gender of parent and children were compared across labels, separately for the four morbidity labels and four conditional mortality labels. Statistical tests fail to reject the null of a constant mean across labels at 10% for all characteristics except gender of parent across the four morbidity labels. With that one exception, the randomly assigned labels are orthogonal to important parent and child characteristics.

and the risk reduction the sun lotion would offer. In this way the magnitude of the risk change

for the parent and the child was described in absolute as well as in percentage terms.

For the first of the two labels, parents were asked, "Now please think about whether you

would buy the new sun protection lotion for yourself or your child. Please do not consider

buying it for anyone else. Suppose that buying enough of the lotion to last you and your child

for one year would cost $X. Of course, if you did buy it, you would have less money for all of

the other things that your family needs. Would you be willing to pay $X for enough of the

sunscreen to last you and your child for one year?" The value of X was randomly selected from

among nine values ranging between $20 and $125. The narrative also reminded parents that

lifetime use of the sun lotion is necessary to obtain the stated skin cancer protection benefits. For

the second label, parents were told, "Suppose that instead of the previous label, we showed you

the following label." Willingness to pay then was elicited as before.

4.   *Empirical Estimates*

4.1.   Methods and Interpretation

Following Cameron (1988), the null hypothesis that parents' stated purchase intentions

for the hypothetical sun lotion are consistent with equations (8) and (9) is tested based on a

specification of the willingness-to-pay function rather than on an explicit specification of a

difference in random utility functions. The approach taken uses the model developed in Section

2 to derive present period ($t = 0$) willingness to pay (WTP^j) for the hypothetical sun lotions to

reduce the unconditional risk of getting skin cancer ($j = a$) and the conditional risk of dying

from this disease if it is contracted ($j = b$).

Each new sun lotion is treated as a newly available private good that if purchased would

provide an increment, S_{it}^j, in the planned amount of protective goods that was optimal in the

absence of the new sun lotion. If individual i uses sunscreen j during period t then $dG_{it}^j = S_{it}^j = 1$; otherwise $dG_{it}^j = S_{it}^j = 0$. The resulting changes in lifetime risk are $dR_i^j = \sum_t (\partial R_i^j / \partial G_{it}^j) S_{it}^j$.[22]

Parents participating in the field study were told the lifetime risk reductions that would result from use of the new sun lotion and that achieving these risk reductions would require lifetime use of the product. Therefore assume that the parent would prefer not to purchase the sun lotion for herself now, unless she envisioned continuing to use it in the future. Likewise, she would prefer not to purchase the sun lotion for her child now, unless she believed that he would find it in his interest to use it in the future. Also, the first period's supply of the sun lotion is offered as a single purchase decision for the parent and child together, rather than as a separate purchase decision for each. In consequence, the parent decides that neither she nor her child will use the sun lotion at all ($S_{it}^j = S = 0$), or that both will use it now and in the future ($S_{it}^j = S = 1$). The possibility that only one of the two individuals would use the sun lotion is addressed below.

Suppose that the required expenditure for the lotion for the parent and child together during $t = 0$ is denoted X^j, and that in subsequent periods, when the child makes his own allocation decisions, each individual may purchase the sun lotion in an amount for one person at half of this expenditure, $X^j / 2$. Then the parent's maximal lifetime utility assuming continuing use of the sun lotion is $U_p^*(y_{p0} - X^j, y_{p1} - X^j / 2, y_{k1} - X^j / 2, y_{k2} - X^j / 2, r, P^a, P^b; S = 1)$, where

---

[22] This specification assumes that users of the new sun lotions would not neutralize the risk reductions by making other substitutions, for example by spending more time outdoors in sunlight. In two previous skin cancer surveys, attempts were made to account for possible substitutions that might influence endogenously perceived risk changes associated with hypothetical sun lotions. In Dickie and Gerking (1996), an indicator for whether respondents used current sunscreen in order to stay outdoors longer was not significantly related to the perceived risk reduction associated with a hypothetical sun lotion. In Dickie and Gerking (2003), respondents were asked whether using a hypothetical sun lotion would lead them or their children to spend more time outdoors in sunlight. Fewer than 10% of parents responded affirmatively, and indicators for this type of substitution were not significantly related to perceived risk changes associated with the hypothetical sun lotion, or with willingness to pay for it. These results suggested that the possibility of offsetting substitutions would not be a major factor considered by parents when they initially evaluated the new sun lotions and consequently no questions concerning this type of behavior were included in the present study.

$U_p^*(\bullet)$ denotes the indirect utility function and where $\partial U_p^* / \partial y_{kt} = 0$ if $\eta = 0$. Derivatives of this

function include

$$(\partial U_p^* / \partial S) = (\partial U_p / \partial R_p^j)dR_p^j + (\partial U_p / \partial R_k^j)dR_k^j + \eta(\partial U_k / \partial R_k^j)dR_k^j$$

$$(\partial U_p^* / \partial X^j) = -\left[\lambda_p + (1/2)(1+r)^{-1}(\lambda_p + \eta\lambda_k) + (1/2)(1+r)^{-2}\eta\lambda_k \right] \qquad (10)$$

$$= -\lambda_p \sum_{t=0}^{2} \left(n_t \Big/ 2 \right)(1+r)^{-t}$$

where the dR_i^j denote the lifetime risk changes resulting from use of the sun lotion in all periods

and n_t denotes the number of users of the sun lotion in period t whom the parent cares about (if

$\eta > 0$, $n_0 = 2 = n_1$, $n_2 = 1$ because the parent cares about the child in all periods, while if

$\eta = 0$, $n_0 = 2$, $n_1 = 1$, $n_2 = 0$ because the parent cares about the child only in $t = 0$). As shown in

equation (10), the child's decision to purchase the sun lotion in periods $t = 1$ and $t = 2$ affects the

parent's welfare if $\eta > 0$.

The parent's willingness to pay for the sun lotion per period, WTP^j, is the value of

X^j that equates $U^*(\bullet) \equiv \bar{U}$, where $\bar{U}$ denotes the parent's maximal lifetime utility if neither she

nor her child uses the sun lotion. Applying the implicit function theorem to this identity and

using equation (10) implies that marginal willingness to pay for the first period of sun lotion use

is

$$d(WTP^j) = (1/\lambda_p)\left[\sum_{t=0}^{2}(n_t/2)(1+r)^{-t} \right]^{-1}\left[\begin{array}{l} (\partial U_p / \partial R_p^j)(dR_p^j) \\ + [(\partial U_p / \partial R_k^j) + \eta(\partial U_k / \partial R_k^j)](dR_k^j) \end{array} \right] \qquad (11)$$

$$= \beta\left(\delta_p^j(-dR_p^j / R_p^j) + \delta_k^j(-dR_k^j / R_k^j) \right). \qquad j = a,b,$$

In this equation $\delta_p^j = -(\partial U_p / \partial R_p^j)R_p^j / \lambda_p$ and $\delta_k^j = -[(\partial U_p / \partial R_k^j) + \eta(\partial U_k / \partial R_k^j)]R_k^j / \lambda_p$ denote

the parent's marginal willingness to pay for proportionate reductions in her own and her child's

lifetime risk, and $\beta = \left[\sum_{t=0}^{2} (n_t / 2)(1+r)^{-t} \right]^{-1}$ denotes the fraction of the present value of total

planned expenditures on the sun lotion that occur in the first period. Because $\beta < 1$, coefficients

of lifetime risk reductions understate the parent's marginal willingness to pay for risk reduction;

i.e., first-period expenditures on sun lotion do not reveal the full willingness to pay for lifetime

risk reduction. Nonetheless, the ratio of coefficients of lifetime risk changes

$\beta \delta_k^j / \beta \delta_p^j = [(\partial U_p / \partial R_k^j) + \eta(\partial U_k / \partial R_k^j) R_k^j] / ((\partial U_p / \partial R_p^j) R_p^j)$ equals the parent's marginal rate of

substitution between equal percentage risk changes for herself and for the child. If the parent is

altruistic, this marginal rate of substitution equals unity.[23]

For econometric estimation, equation (11) is specified for parent h as

$$WTP_h^j = \gamma_0^j + \gamma_p^j \left[\Delta_p^j / R_p^j \right]_h + \gamma_k^j \left[\Delta_k^j / R_k^j \right]_h + controls_h + \varepsilon_h^j. \tag{12}$$

In equation (12), Δ_p^j and Δ_k^j are interpreted as the discrete reduction in the jth risk for the parent

and the child that would occur if the sun lotion was used, the R_i^j denote the last estimate the jth

perceived risk elicited for individual i in the field study, and $\gamma_i^j = \beta \delta_i^j$, $i=p,k$. Thus the variables

in square brackets denote the percentage risk reductions (divided by 100) shown on the sun

lotion labels for the jth type of risk and take the value 0.1 or 0.5. Treating the γ_i^j as constants

implies that willingness-to-pay per unit of risk reduction $\partial WTP^j / \partial \Delta_i^j = \gamma_i^j / R_i^j$ decreases with

---

[23] Nonmonetary costs of using the sun lotion such as time costs of ensuring proper application and disutility from odor or other product attributes are assumed equal for parent and child. The description of the sun lotion attempted to minimize time requirements by indicating that one application would last all day and to control for potential sources of disutility such as odor, allergic reactions and blocking of pores. The description was constant across all labels. To the extent that nonmonetary costs differ between parent and child, however, the costs would be confounded in the δ_i^j coefficients.

the magnitude of perceived risk initially faced.[24] Also: (1) *controls* refers effects on willingness

to pay of measured parental characteristics such as income and family size, and (2) ε_h^j denotes a

random disturbance term with standard properties included to capture unobserved characteristics

of parent h. These characteristics might include willingness to try new products, the ability to

process the information presented on the sun lotion label, evaluation of joint outputs such as

sunburn protection and skin aging, as well as other factors that influence whether the product

would be purchased.

Five aspects of equation (12) warrant further discussion before turning to the results of

estimation. First, altruism implies that $\gamma_k^j / \gamma_p^j = 1$. But a test of this hypothesis does not

distinguish between types of altruism that may motivate parents' stated intentions to purchase the

sun lotion, because $\partial U_p / \partial R_k^j$ and η are not separately identified; both are components of γ_k^j.

Distinguishing between the types of altruistic motivations considered in Section 2 must await

further research that contrasts parental behavior toward both young and adult children. In any

case, the test does not rest on directly estimating WTP for risk reduction, but instead on

estimating the ratio of estimated contributions of risk reduction to willingness to pay. This

means that γ_k^j and γ_p^j must be consistently estimated, but it is not necessary to obtain a consistent

estimate of γ_0^j.

Second, the percentage risk reduction variables are randomly assigned experimental

design points. Thus, they are orthogonal to other experimental design points as well as to parent

---

[24] In other words, the marginal value of risk reduction $\partial U_i / \partial R_i^j$ diminishes as R_i^j rises so that γ_i^j remains constant.
To test the adequacy of this specification, which treats willingness to pay as a linear function of percentage risk
changes, separate regressions were run for low-risk and high-risk groups. The null hypothesis that slope coefficients
in both the morbidity and conditional mortality equations are equal in the high and low risk groups was not rejected
at conventional levels. This result occurred whether morbidity risk or conditional mortality risk of the parent or the
child was used to distinguish between low and high risk groups. The test was based on the first specification
reported in Table 5 below.

characteristics included in *controls* and to parent characteristics captured by ε_h^j. This means that if the functional form of equation (12) is correct: (1) endogeneity problems in estimating the γ_i^j are avoided and (2) estimates of the γ_i^j are unaffected by the choice of variables to include in *controls*.

Third, willingness to pay for the sun lotion is treated in an errors-in-variables framework in which stated willingness to pay (W_h^j) by parent h to reduce the jth risk differs from true willingness to pay (WTP_h^j) by both systematic and random factors according to

$$W_h^j = WTP_h^j + \alpha_h^j = WTP_h^j + \alpha^j + v_h^j, \qquad j = a, b. \qquad (13)$$

In equation (13), α^j is the nonzero mean of α_h^j and v_h^j is a random disturbance. α^j is assumed to represent systematic misstatement of true willingness to pay. For example, parents may misstate willingness to pay because the choice of whether to buy the sun lotion was presented as a hypothetical question and/or may not have been adequately considered in light of preferences, and financial constraints.[25] Also, v_h^j captures unobserved parent-specific heterogeneity as well as purely random factors that may affect a parent's stated willingness to pay for the label presented. The v_h^j are assumed to be normally distributed with mean zero and constant variance and the possibility that $E(v_h^a v_h^b) \neq 0$ motivates joint estimation of willingness-to-pay equations for the two types of risk.

The marginal rate of substitution (γ_k^j / γ_p^j) is estimated by substituting equation (13) into equation (12) to obtain

---

[25] As discussed by Carson, Groves, and Machina (2000) the overstatement of purchase intentions arising from incentive incompatibility of hypothetical, binary discrete-choice questions for private goods is unrelated to the scope of the good and its costs. Also, joint benefits of the sun lotion are held constant across labels but the parent's evaluation of any perceived difference between joint outputs of the lotion and existing products would be reflected in the constant term.

$$W_h^j = (\gamma_0^j + \alpha^j) + \gamma_p^j \left[\Delta_p^j / R_p^j \right]_h + \gamma_k^j \left[\Delta_k^j / R_k^j \right]_h + controls_h + \varepsilon_h^j + v_h^j, \qquad j = a, b. \quad (14)$$

Notice that estimators of the constant term (γ_0^j) will be inconsistent if, as expected, $\alpha^j \neq 0$.

Also, estimators of coefficients of parent characteristics included in *controls* will be inconsistent

if the controls are correlated with the composite error ($\omega_h^j = \varepsilon_h^j + v_h^j$). Nevertheless, consistent

estimators of γ_k^j and γ_p^j still can be obtained as long as equation (14) is correctly specified,

because the two risk reduction variables are experimental design points that were assigned

independently of parent characteristics.

Fourth, the dependent variable W_h^j (stated willingness to pay for a one year's supply of

sun lotion) is latent: Parents only were asked to state whether they would be willing to make a

randomly assigned expenditure. Parents are assumed to answer in the affirmative if $W_h^j > P_h^j$,

where P_h^j denotes the expenditure for a one year supply of sun lotion *j* that was randomly

assigned to parent *h*. Thus a parents states that she will purchase the sun lotion if

$$\omega_h^j / \sigma^j < (\gamma_0^j + \alpha^j) / \sigma^j + (\gamma_p^j / \sigma^j) \left[\Delta_p^j / R_p^j \right] + (\gamma_k^j / \sigma^j) \left[\Delta_k^j / R_k^j \right] - (1 / \sigma^j) P_h^j,$$

where the *controls* are suppressed for notational simplicity, $E(\omega_h^j) = 0$ and $\mathrm{var}(\omega_h^j) = (\sigma^j)^2$, and

ω_h^j is symmetrically distributed. These features together with an assumption of normally

distributed composite errors that have an expected non-zero covariance across equations

$E(\omega_h^a \omega_h^b) = \sigma_{ab} \neq 0$ motivates estimation by bivariate probit, where $\rho = \sigma_{ab} / \sigma^a \sigma^b$.[26] Following

Cameron and James (1987), the coefficient of the randomly assigned sun lotion price is

interpreted as an estimate of $-1 / \sigma^j$ that can be used to recover unnormalized coefficients of risk

reductions (γ_i^j) from the normalized estimates of γ_i^j / σ^j.

---

[26] Of course, the assumption of normally distributed errors will not be exactly satisfied when non-normally distributed parent characteristics (e.g., income) are not included as covariates.

Fifth, a concern is that use of stated preference data to estimate the willingness to pay

function will result in a comparatively large variance of the composite error ($\omega_h^j = \varepsilon_h^j + v_h^j$).

Stated preference data are often "noisy" and this feature could lead to wide confidence intervals

around the estimated values of marginal rates of substitution, thus making it more likely that the

null hypothesis being tested will not be rejected.

4.2   Results

Full information maximum likelihood bivariate probit estimates are shown in Table 5.[27]

Sample means of covariates are presented along with the regression estimates. Two pairs of

estimates are reported. The first uses only design points as covariates and the second shows the

outcome when two controls for parent characteristics (family income and number of children in

the family) are added. Two design points measure skin cancer risk changes for the parent and

the child (see equation (14)) and a third measures the randomly assigned sun lotion price. A

fourth design point variable is added to control for the order in which the morbidity and

conditional mortality labels were shown.

Consider first the pair of estimated regressions that use only design points as covariates.

The estimated value of ρ (=0.778) is positive, as expected, and significantly different from zero,

indicating an efficiency gain from joint estimation of the two equations. The coefficients of the

required annual expenditure are negative and differ significantly from zero at 1%, suggesting that

parents were more reluctant to purchase the sun lotion at higher costs than at lower costs.

Additionally, coefficients of variables measuring percentage reductions in the two types of risk

to both parent and child are positive and significantly different from zero at the 1% level in each

---

[27] Ordinary least squares estimates were used as initial values in computing the binomial probit estimates used as
starting values for the bivariate probit routine. Coefficient estimates and estimates of the marginal rate of
substitution between child and parent risks from the binomial probit estimates are broadly consistent with those
reported in Tables 4 and 5, but are less precisely estimated.

of the two equations. This outcome suggests that parents are willing to pay more for larger than

for smaller reductions in the two types of risk and is consistent with the conceptual model

presented in Section 2. Comparing these coefficients to the estimated intercept, however,

appears to suggest that increases in risk reduction do not bring about proportionate increases in

willingness to pay. Many previous studies have found that stated willingness to pay does not

increase proportionately with increases in risk reductions (see Hammitt and Graham 1999 for

further discussion of this issue). Nevertheless, this conclusion may not apply because the

(unnormalized) intercepts actually are estimates of $(\gamma_0^j + \alpha^j)$ rather than γ_0^j, and $\alpha^j > 0$ if

parents tend to overstate purchase intentions. Also, as mentioned previously, coefficients

understate willingness to pay for reduced risk because $\beta < 1$. Estimates show that the order in

which the morbidity and conditional mortality labels were presented is unimportant.

When controls for income and family size are introduced, estimates again indicate

positive correlation between the errors in the two equations (0.788). Coefficients of family

income are positive while coefficients of the total number of children in the family are negative

as expected. These coefficients, however, are not consistently estimated if income and family

size are correlated with unobserved family characteristics influencing the sun lotion purchase

decision. Income coefficients are significantly different from zero only at the 10% level under a

two-tail test, suggesting a weak tendency for parents' willingness to pay to increase with income.

The small effect of income may simply reflect the relatively low costs of the sun lotion, with the

highest cost reaching only about $10/month. Coefficients of the number of children are

significant at the 1% level, providing evidence that parents reduce protective expenditures per

family member when more children are present. Because the risk change variables are

orthogonal to these parent characteristics, coefficients and standard errors of risk changes are

little altered from their corresponding values discussed previously. Supplementary regressions

(Appendix) specified like those in the last pair of columns but also including covariates for

marital status, education, age and gender of parent, age and gender of child, and whether a close

relative had been diagnosed with skin cancer also demonstrated this same result. Only two of the

additional 14 coefficients differed significantly from zero at 10%.[28] Also, in this expanded

regression, coefficients of the risk change variables were almost unchanged as compared with

those presented in Table 5.

Table 6 reports tests of whether the equilibrium condition implied by altruism holds

($\gamma_k^j / \gamma_p^j - 1 = 0$, $j = a, b$). Column (2), Table 6, labeled "full sample," reports results based on

Table 5 estimates that control only for design points. Standard errors are computed using the

delta method. As shown, the null hypothesis that this equilibrium condition holds is not rejected

at conventional significance levels in either the unconditional morbidity or conditional mortality

equations. This null hypothesis also is not rejected using a Wald test of the restriction

$\gamma_k^j / \gamma_p^j - 1 = 0$ in both equations jointly.

Remaining columns of Table 6 summarize outcomes of parallel tests in six subsamples

defined according to the gender of parent, gender of child, and age of child. Results for

subsamples were obtained by re-estimating the willingness-to-pay equations separately by

subsample using only the four experimental design points as covariates. Parent gender is

considered because the unitary model assumes that families act as if maximizing a single utility

function, so that decisions made by mothers should be consistent with those made by fathers.

Gender and age of child are considered because parental marginal rates of substitution should not

---

[28] The two variables with significant coefficients were parent gender in the morbidity equation and child age in the
conditional mortality equation. Also, in regressions including only experimental design points and the constructed
measures of perceived latency for parents and children, three of the four latency coefficients were negative as
expected, but none was significant.

differ between children as long as marginal costs of risk reduction are the same, as in this field study. As shown in Table 6, results are consistent with the hypothesis $\gamma_k^j / \gamma_p^j = 1$ in all six subsamples. Furthermore, likelihood ratio tests detect no significant differences in willingness to pay functions by gender of parent, or by age or gender of child.[29]

Although not reported in Table 6, a comparable analysis was undertaken based on subsamples defined by family income, by age and education of parent, and by presence of one versus more than one child in the family. This analysis is motivated by the assumed constancy of coefficients of the willingness to pay functions, relative to the possibility that the marginal utility of income, the β term, or other parameters may vary with characteristics of the parent.[30] Also, the model in Section 2 includes only one child in the family and the survey asked parents to consider using the sun lotion for only one of their children, even though most parents in the sample reported having more than one child. However, the null hypothesis that parameters of willingness to pay functions are equal between families with high or low income, or between parents with and without college educations, or between older and younger parents, or between single or multi-child families, is not rejected. Also, the hypothesis $\gamma_k^j / \gamma_p^j = 1$ is not rejected in any of these additional subsamples.

---

[29] The null hypotheses that slope coefficients of the equations do not differ by gender of parent, or by gender or age of child, after allowing for different intercepts, were each separately tested using likelihood ratio tests. Results indicated that the null hypothesis would not be rejected at conventional significance levels in any comparison. Further analysis of the role of parent gender was conducted by re-estimating the model in the last two columns of Table 5 while including a dummy variable for parent gender and interactions of this variable and all covariates. The only statistically significant difference between male and female parents was found in the coefficient of the number of children in the morbidity equation, where female willingness to pay for the sun lotion declined less than male willingness to pay with increases in the number of children. Coefficients of risk changes, annual cost and income appear to be the same for mothers and fathers. Also, outcomes of all of these tests by parent gender are the same if the comparison is restricted to married parents.

[30] A related issue involves whether parents differed in their perceptions of available substitutes for the hypothetical sun lotion. The survey would have been improved had parents been asked how skin cancer risks could have been reduced by the amounts shown on the labels if the product were not available or if they chose not to buy it. In the absence of this information, we assume that either substitution opportunities are negligible or are the same for all parents.

The analysis presented assumes that the parent would use the sun lotion for herself and her sample child but not for anyone else. The apparent decline in willingness to pay for the sun lotion with increases in the number of children in the family (Table 5) along with the lack of significant differences in slope coefficients of willingness to pay functions between single- and multiple-child families suggests that parents did not envision using the sun lotion to protect additional children when stating their purchase intentions. Also, parents who indicated that they would buy the sun lotion were asked about the intended users. The majority of parents indicated that the lotion would be used for the parent and the sample child (85% for the morbidity labels and 90% for the conditional mortality labels), with almost all of the remaining purchasers intending to use the lotion for the child only.[31] Excluding parents who envisioned purchasing the sunscreen but using it for only one individual does not change the outcome of any of these statistical tests. Additionally, because parents were told that achieving the stated risk reductions required use of the lotion as directed, the above tests were performed again after adjusting the risk change measures of Table 5 so that the risk change would be zero for the parent or child if the parent did not envision that person using the sun lotion. The null hypothesis is not rejected using these adjusted measures of risk changes.

Finally, empirical results obtained can be used to test another aspect of the model presented in Section 2. Wald tests are carried out to determine whether the marginal rate of substitution between the two types of risk for both parent and child are equal to the corresponding ratio of marginal costs in reducing these risks. This amounts to testing whether the cross-equation coefficient restrictions $\gamma_i^a / \gamma_i^b = 1,\ i = p, k,$ are valid. To control for different values of σ^j in the two equations, the tests were conducted using the unnormalized coefficient

---

[31] Four parents who indicated that they would purchase one of the sun lotions envisioned using it for themselves only (three for the morbidity labels and one for the conditional mortality labels).

estimates. Standard errors of ratios of these coefficients were computed using the delta method.

In separate tests involving the coefficients of risk reduction for parents and children, the null

hypothesis is not rejected at conventional significance levels. Additionally, a joint test of the

null hypothesis for parents and children together yields the same result. These results are

consistent with altruism and suggest that parents responded to the assigned changes in the two

types of risk consistently with the theoretical model of Section 2.[32]

5.   *Summary and Conclusions*

Special protection of young children from environmental hazards has become a

worldwide priority of government policies to improve human health. The fundamental tension

between altruism and self-interest in families looms as the crucial behavioral factor determining

the effectiveness of these policies. This paper estimates parents' marginal rates of substitution

between skin cancer risks faced by 488 parents and their children between the ages of 3 and 12

years. A model of altruistic family behavior that incorporates household production of latent

health risk guides the estimates. The model demonstrates that the marginal rate of substitution

between risks faced by the parent and child is equal to the ratio of marginal risk reduction costs.

Resulting empirical estimates then focus on whether this equality holds.

Tests rest on an examination of stated preference values for a hypothetical sun lotion.

Although stated preference valuation is a controversial method of obtaining willingness to pay to

reduce environmental risks, it supports consistent estimation of parents' marginal rates of

substitution between health risks to themselves and corresponding health risks to their children in

the field study described here. Consistent estimation of marginal rates of substitution is made

possible by: (1) allowing for both systematic and random errors in parents' stated willingness to

---

[32] The outcome of this test reinforces the conclusion that respondents sensibly considered the compound
probabilities involved in the study.

31

pay for the sun lotion and (2) randomly assigning skin cancer risk reductions offered by sun lotion to the sample of parents. Together, these innovations imply that the skin cancer risk reductions assigned are orthogonal both to parent characteristics and to errors parents may make in assessing their willingness to pay for the sun lotion.

In the theoretical model, an altruistic parent's marginal rate of substitution between risk to her child and risk to herself equates with the corresponding ratios of marginal skin cancer risk reduction costs. This prediction is the basis of the null hypothesis for econometric tests using data from the field study. The null hypothesis is not rejected, so test results support the notion that parents are altruistic toward their young children. This outcome stands in contrast to findings in related studies that present evidence against altruism of parents toward their children. This study, however, looks at behavior of parents toward pre-teenage children living at home, rather than behavior of parents toward their adult children who have formed households of their own. An important implication is of findings from this study is that effectiveness of public intervention programs to reduce environmental risks faced by children may be compromised to some extent because parents will respond by redistributing family resources.

References

Agee, Mark D. and Thomas D. Crocker. 2004. "Transferring Measures of Adult Health Benefits to Children: A Review of Issues and Results," *Contemporary Economic Policy* 22: 468-482.

Altonji, Joseph, Fumio Hayashi, and Laurence Kotlikoff. 1992. "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," *American Economic Review* 82, 1177-98.

Altonji, Joseph, Fumio Hayashi, and Laurence Kotlikoff. 1997. "Paternalistic Altruism and Inter- vivos Transfers: Theory and Evidence," *Journal of Political Economy* 105: 1121-1166.

Barro, Robert J. 1974. "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82: 1095-1117.

Becker, Gary S. 1974. "A Theory of Social Interactions," *Journal of Political Economy*, 82: 1063-1094.

Becker, Gary S. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Bergstrom, Theodore C. 2003. "Benefit-Cost Analysis and the Entanglements of Love," University of California, Santa Barbara manuscript.

Bergstrom, Theodore C. 2006. "Benefit-Cost in a Benevolent Society," *American Economic Review* 96: 339-351.

Brouwer, Roy and Ian J. Bateman. 2005. "Benefits Transfer of Willingness to Pay Estimates and Functions for Health Risk Reductions: A Cross-Country Study," *Journal of Health Economics* 24: 591-611.

Cameron, Trudy A. and Michelle D. James. 1987. "Efficient Estimation Methods for 'Closed-Ended' Contingent Valuation Surveys," *Review of Economics and Statistics* 69: 269-276.

Cameron, Trudy A. 1988. "A New Paradigm for Valuing Non-market Goods Using Referendum

Data: Maximum Likelihood Estimation by Censored Logistic Regression," *Journal of Environmental Economics and Management* 15: 355-379.

Carson, Richard T, Theodore Groves and Mark J. Machina. 2000. "Incentive and Informational Properties of Preference Questions." University of California, San Diego, Department of Economics working paper.

Corso, Phaedra S., James K. Hammit and John D. Graham. 2001. "Valuing Mortality-Risk Reduction: Using Visual Aids To Improve The Validity Of Contingent Valuation," *Journal of Risk and Uncertainty* 23: 165-84.

Cox, Donald and M. Rank. 1992. "Inter-vivos Transfers and Intergenerational Exchange," *Review of Economics and Statistics* 74: 305-314.

Dickie, Mark and Shelby Gerking. 1991. "Valuing Reduced Morbidity: A Household Production Approach," *Southern Economic Journal* 57: 690-702.

Dickie, Mark and Shelby Gerking. 1996. "Formation of Risk Beliefs, Joint Production and Willingness to Pay to Avoid Skin Cancer. *The Review of Economics and Statistics* 78: 451-464.

Dickie, Mark and Shelby Gerking. (2003). "Parents' Valuation of Latent Health Risks to their Children," in *Risk and Uncertainty in Natural Resource and Environmental Economics*, (eds.) Justus Wesseler, Hans-Peter Weikard, and Robert D. Weaver, Edward Elgar: Cheltenham, UK.

Dickie, Mark and Shelby Gerking. (2006). "Valuing Children's Health: Paternalistic Perspectives," in *Economic Valuation of Environmental Health Risks to Children*, OECD: Paris.

Dickie, Mark and Victoria L. Messman. 2004. "Paternalistic Altruism and the Value of Avoiding Acute Illness: Are Kids Worth More than Parents," *Journal of Environmental Economics and Management* 48: 1146-1174.

Federal Register. 1997. Executive Order 13045 of April 21, 1997—Protection of Children from Environmental Health Risks and Safety Risks, 62, FR Doc. 97-10695 (April 23).

Hamilton, William D. 1964. "The Genetical Evolution of Social Behavior, i and ii," *Journal of Theoretical Biology* 7: 1-52

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

Jenkins, Robin R., Nicole Owens, and Lanelle B. Wiggins. 2001. "Valuing Reduced Risks To Children: The Case Of Bicycle Safety Helmets," *Contemporary Economic Policy* 19: 397-408.

Jones-Lee, Michael. 1991. "Altruism and the Value of Other People's Safety," *Journal of Risk and Uncertainty* 4: 213-219.

Jones-Lee, Michael. 1992. "Paternalisticistic Altruism and the Value of Statistical Life," *The Economic Journal* 102: 80-90.

Krupnick, Alan, Anna Alberini, Maureen Cropper, Nathalie Simon, Bernie O'Brien, Ron Goeree, and Martin Heintzelman. 2002. Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Study of Ontario Residents. *Journal of Risk and Uncertainty* 24: 161–86.

Laitner, John and F. Thomas Juster. 1996. "New Evidence on Altruism: A Study of TIAA-CREF Retirees," *American Economic Review* 86: 893-908.

Liu, Jin-Tan, James K. Hammitt, Jung-Der Wang and Jin-Long Liu. 2000. "Mother's Willingness to Pay for Own and Her Child's Health: A Contingent Valuation Study in Taiwan," *Health Economics* 9: 319-26.

Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales. 1997. "Do Households and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit", *Journal of Human Resources,* 32, 463-480.

Nastis, Stefanos A. and Thomas D. Crocker. 2003. "Pregnant Mother's Valuation of Own and of Child Health," presented at USEPA/University of Central Florida conference, "Valuing Environmental Health Risk Reductions to Children," Oct 20-21, http://yosemite.epa.gov/EE/epa/eerm.nsf/vwRepNumLookup/EE-0475?OpenDocument.

Ries, L.A.G., M.P. Eisner, C.L. Kosary, B.F. Hankey, B.A. Miller, L. Clegg, A. Mariotto, M.P. Fay, E.J. Feuer, B.K. Edwards (eds). 2003. *SEER Cancer Statistics Review, 1975-2000*, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2000.

Scapecchi, Pascale. 2006. "Valuation Differences between Adults and Children," in *Economic Valuation of Environmental Health Risks to Children*, OECD: Paris.

Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein, "Regulation of risk: A psychological perspective," in Roger Noll (ed.), *Regulatory Policy and the Social Sciences* (Berkeley: University of California Press, 1985).

U.S. Environmental Protection Agency. 1996. *A National Agenda to Protect Children's Health from Environmental Threats*, Washington, D.C.: U.S. Government Printing Office.

Table 1. Frequency Distribution of Parents' Perceived Risks.

N=488.

Risk Range (%)	Risk of Getting Skin Cancer[a]		Conditional Risk of Dying from Skin Cancer	
	Parents	Children	Parents	Children
0 - 4.75	53	46	78	111
5 - 9.75	24	48	140	169
10 - 14.75	53	78	112	97
15 - 19.75	55	62	59	40
20 - 24.75	55	59	33	28
25 - 29.75	61	63	22	17
30 - 34.75	39	32	9	5
35 - 39.75	22	16	7	5
40 - 44.75	33	23	4	5
45 - 49.75	6	4	5	1
50 - 54.75	49	29	16	9
55 - 59.75	4	2	1	1
60 - 64.75	5	5	0	0
65 - 69.75	0	1	0	0
70 - 74.75	4	2	2	0
75 - 79.75	6	5	0	0
80 - 84.75	2	3	0	0
85 - 89.75	2	2	0	0
90 - 94.75	9	5	0	0
95 – 100	6	3	0	0

[a]Initial risk assessment.

Table 2. Parents' Mean Risk Perceptions (%).

Sample	Risk of Getting Skin Cancer[a]	Conditional Risk of Dying from Skin Cancer	Sample Size
All Parents	26.93	12.05	488
All Children	22.46	9.36	488
Mothers	29.17	12.46	368
Fathers	20.08	10.82	120
Daughters	22.31	9.38	242
Sons	22.61	9.33	246
Children aged 3 to 7 years	23.84	10.10	275
Children aged 8 to 12 years	20.68	8.39	213

[a]Initial risk assessment.

Table 3. Use of Sun Protection Products.

Fraction of Time Outdoors that Sun Protection Products Used	Parents	Children
Never	44	15
Less than half	115	80
About half	109	106
More than half	91	106
Always/almost always	129	181

Sun Protection Factor Normally Used	Parents	Children
Less than 15	67	15
15 to less than 30	185	103
30 or higher	192	355

Table 4. Frequency Distribution of Expected Age at Onset.

N=488

Age Range (years)	Parents	Children
Before age 40	45	68
40 - 44	63	42
45 - 49	64	52
50 - 54	111	84
55 - 59	61	66
60 - 64	84	55
65 - 69	41	46
70 - 74	13	49
75 - 79	1	12
Age 80 or later	5	14
Mean age at onset (years)	53	55
Mean age (years)	35	7
Implied mean expected latency period (years)	18	48

	Sample Mean (Std. Dev.) or Proportion		Coefficients (Standard Errors)			
Covariate (Parameter Notation)	Morbidity Risk	Conditional Mortality Risk	Morbidity Risk	Conditional Mortality Risk	Morbidity Risk	Conditional Mortality Risk
Parent's Percentage Risk Reduction	0.289	0.302	0.912	0.717	0.901	0.739
(γ_p^j / σ^j)	(0.200)	(0.200)	(0.272)	(0.267)	(0.274)	(0.267)
Child's Percentage Risk Reduction	0.300	0.299	0.854	1.426	0.843	1.487
(γ_k^j / σ^j)	(0.200)	(0.200)	(0.270)	(0.267)	(0.275)	(0.272)
Cost of Sun Lotion ($/year)	64.518	64.150	-0.011	-0.011	-0.011	-0.011
$(-1/\sigma^j)$	(34.520)	(34.897)	(0.002)	(0.002)	(0.002)	(0.002)
Order (=1 if risk change in column presented last, 0 if first)	0.488	0.512	-0.149	-0.087	-0.151	-0.105
			(0.122)	(0.122)	(0.126)	(0.125)
Family Income ($10,000/year)	5.957				0.028	0.029
	(3.569)				(0.018)	(0.017)
Number of Children in Family	2.078				-0.190	-0.004
	(0.952)				(0.069)	(0.068)
Constant			0.733	0.520	0.981	0.347
$((\gamma_0^j + \alpha^j)/\sigma^j)$			(0.171)	(0.170)	(0.251)	(0.229)
Error Correlation			0.778		0.788	
(ρ)			(0.044)		(0.044)	
Log-Likelihood			-512.553		-505.391	

Table 6. Estimates of γ_k^j / γ_p^j and Altruism Tests.

	Full Sample	Mothers	Fathers	Daughters	Sons	Child Age 3-7	Child Age 8-12
		Estimates of γ_k^j / γ_p^j (Standard Errors) and Tests of Altruism					
Morbidity ratio $(\gamma_k^a / \gamma_p^a)$	0.936	0.927	0.88	0.902	0.96	1.438	0.441
	(0.415)	(0.456)	(0.678)	(0.777)	(0.503)	(0.766)	(0.462)
z-test ratio=1 (p)	0.878	0.873	0.860	0.900	0.937	0.568	0.226
Conditional Mortality ratio $(\gamma_k^b / \gamma_p^b)$	2.005	1.816	3.746	1.512	3.003	5.018	0.661
	(0.853)	(0.837)	(6.133)	(0.702)	(2.688)	(4.962)	(0.417)
z-test ratio=1 (p)	0.240	0.329	0.654	0.465	0.456	0.418	0.416
Wald test, both ratios=1 (p)	0.493	0.608	0.883	0.761	0.750	0.601	0.398
Sample Size	488	368	120	242	246	275	213
LR test, equal parameters between groups (p)		0.975		0.958		0.214	

SkinSaver

Sun Protection Lotion

For Adults & Children

Developed with dermatologists to protect skin from harmful effects of sun exposure.

Skin Cancer
Protection

Ultra
Waterproof

SPF ___

Parsol⊗1789

"Making the outdoors safer
for you and your family."

New SkinSaver® sun protection lotion.

Skin Cancer Protection

✓ **Used as directed in clinical trials, SkinSaver reduced risk of skin cancer by:**

10% for Adults	10% for Children

✓ **Used as directed in clinical trials, SkinSaver had no effect on the risk of dying if skin cancer occurred.**

More Skin Protection

Parsol®1789		SPF_____
Protects against premature skin aging		Protects against sunburn

More Added Features

✷ Ultra long-lasting waterproof formula – One application lasts all day ✷

✷ Non-comedogenic–Won't block pores ✷ Oil-free–Won't feel greasy ✷

✷ Hypoallergenic ✷ PABA-free ✷ Unscented ✷

DIRECTIONS: Apply generously and evenly to all exposed areas of skin at least 15 minutes before sun or water exposure.

ACTIVE INGREDIENTS: Oxybenzone, octocrylene, 2-ethylhexyl salicate, homosalate, avobenzone.

APPENDIX: Supplemental Data and Empirical Results

Table A-1

Hypothetical Sun Protection Product Labels

Label	Percent Change in Morbidity Risk		Percent Change in Mortality Risk	
	Parent	Child	Parent	Child
A	10	10	0	0
B	10	50	0	0
C	50	10	0	0
D	50	50	0	0
E	0	0	10	10
F	0	0	10	50
G	0	0	50	10
H	0	0	50	50

Table A-2. Sample Means by Experimental Design Point.

Label	Morbidity Risk				Conditional Mortality Risk			
	A	B	C	D	E	F	G	H
Percentage risk change for parent	10	10	50	50	10	10	50	50
Percentage risk change for child	10	50	10	50	10	50	10	50
Perceived risk of getting skin cancer for parent	30.26	25.58	26.19	25.44	27.40	25.08	27.59	27.63
Perceived risk of getting skin cancer for child	23.37	22.88	22.18	21.27	23.13	18.90	23.47	24.27
Perceived conditional risk of dying from skin cancer for parent	11.89	11.89	12.05	12.41	11.83	10.66	13.21	12.47
Perceived conditional risk of dying from skin cancer for child	9.30	8.76	9.85	9.59	8.58	8.73	10.43	9.66
Family Income ($10,000/year)	5.67	6.49	5.99	5.66	6.03	6.00	6.14	5.67
Number of Children in Family	2.10	2.10	2.04	2.07	2.23	1.97	2.05	2.07
Parent is female	0.85	0.78	0.68	0.70	0.78	0.78	0.73	0.74
Child is female	0.45	0.53	0.46	0.55	0.46	0.56	0.52	0.45
Child age	7.18	7.12	7.25	6.72	6.95	7.40	6.86	7.07
Sample Size	130	127	114	117	121	120	124	123

Table A-3.   Willingness to Pay to Reduce Skin Cancer Risks: Bivariate Probit Estimates (N=488).

	Mean (s.d.) or Proportion		Coefficients (Standard Errors)			
	Morb. Risk	Cond. Mort. Risk	Morb. Risk	Cond. Mort. Risk	Morb. Risk	Cond. Mort. Risk
Parent's Percentage Risk Reduction	0.289	0.302	0.990	0.711	0.918	0.749
	(0.200)	(0.200)	(0.300)	(0.277)	(0.278)	(0.271)
Child's Percentage Risk Reduction	0.300	0.299	0.849	1.412	0.850	1.384
	(0.200)	(0.200)	(0.279)	(0.288)	(0.271)	(0.272)
Cost of Sun Lotion ($/year)	64.518	64.150	-0.011	-0.011	-0.011	-0.011
	(34.520)	(34.897)	(0.002)	(0.002)	(0.002)	(0.002)
Order (=1 if risk change in column presented last, 0 if first)	0.488	0.512	0.026	0.024	-0.146	-0.104
			(0.022)	(0.021)	(0.123)	(0.123)
Parent Perceived Latency Period	18.092				-0.045	-0.077
	(9.811)				(0.072)	(0.073)
Child Perceived Latency Period	48.148				0.012	-0.028
	(12.239)				(0.060)	(0.059)
Family Income ($10,000/year)	5.957		0.026	0.024		
	(3.569)		(0.022)	(0.021)		
Number of Children in Family	2.078		-0.194	-0.022		
	(0.952)		(0.073)	(0.073)		
Parent is Married	0.830		0.104	-0.019		
			(0.182)	(0.188)		
Parent is College Graduate	0.576		0.081	0.072		
			(0.138)	(0.137)		
Parent Age	35.117		-0.004	-0.004		
	(6.63)		(0.012)	(0.011)		
Parent is Female	0.754		0.271	0.161		
			(0.154)	(0.149)		
Child Age	7.070		0.011	0.051		
	(2.937)		(0.025)	(0.025)		
Child is Female	0.496		0.156	0.061		
			(0.128)	(0.127)		
Close Relative of Parent Diagnosed with Skin Cancer	0.252		0.036	-0.177		
			(0.150)	(0.158)		
Constant			0.063	0.520	0.751	0.815
			(0.144)	(0.170)	(0.303)	(0.296)
Error Correlation			0.791		0.777	
			(0.0443)		(0.0445)	
Log-Likelihood			-500.121		-511.018	

Is An Ounce of Prevention Worth a Pound of Cure?

Ryan Bosworth and Trudy Ann Cameron
Department of Economics
University of Oregon

and

J.R. DeShazo
Department of Public Policy, UCLA*

Corresponding author:

Trudy Ann Cameron
R.F. Mikesell Professor of Environmental and Resource Economics
Department of Economics, 435 PLC
1285 University of Oregon
Eugene, OR 97403-1285
Email: cameron@uoregon.edu
Phone: (541) 346-1242; Fax: (541) 346-1243

_____

Is An Ounce of Prevention Worth a Pound of Cure?

Abstract

We examine how preferences for prevention and treatment policies vary with individual characteristics and policy attributes, which include costs to the individual, the prevalence of the public health problem (numbers of illnesses and deaths), the extent to which each policy reduces illnesses and deaths, the type of health risk (disease) and, for prevention policies, the underlying cause and the time horizon for the policy. Individuals do prefer prevention policies to treatment policies, although at a rate considerably less than the 16 to 1 ratio implied by the "ounce of prevention…" adage. Preferences also differ substantially by the characteristics of the respondent or policy.

1 Introduction

Is an ounce of prevention really worth a pound of cure? Some polices can prevent illnesses by providing a cleaner environment or safer roads. Other policies can allocate resources to help treat those who are already sick or injured. Should we allocate additional resources to help those who are already sick, or should we spend more on measures that will help people avoid illnesses in the first place? Policy makers, at one level or another, must often make these types of tradeoffs when allocating resources for community health improvements.

Previous economic research that directly examines differences in preferences for treatment and prevention policies in a utility-theoretic or willingness-to-pay (WTP) framework is sparse. The existing literature most closely related to our own includes Corso et al. (2002), Hammit and Liu (2004), and Subramanian and Cropper (2000).

Corso et al. (2002) use survey data to assess preferences for treatment policies and prevention policies that provide equivalent mortality risk reductions and find that WTP for treatment policies is much higher. The research reported herein uses a more detailed survey instrument and can exploit a richer set of data on responded characteristics to better understand the systematic differences that tend to make prevention policies preferable to treatment policies.

Hammit and Liu (2004) do not address the difference in prevention and treatment policies. However rather, their research is related to the present study because they investigate in the impact of latency and disease type on WTP. By considering two different latencies and two disease types, Hammit and Liu find that WTP for risk prevention declines with latency and that WTP to avoid a specific cancer risks is moderately larger than WTP to avoid a non-cancer chronic disease with similarly severe symptoms. We also find evidence that WTP declines with

latency and that WTP is relatively higher for policies that prevent cancer risks. However, we find that WTP is relatively lower for policies that provide treatment to cancer victims.

Subramanian and Cropper (2000) investigate the relationship between WTP for public risk reduction policies and the qualitative factors (such as funding fairness or blame for the health risks) associated with those policies. Relative to their study, we provide a more extensive analysis of how WTP for both treatment and prevention policies is related to a broader variety of both quantitative and qualitative factors.

The question of how to value various prevention or treatment policies has been discussed extensively in the Quality Adjusted Life Years (QALY) literature. For example, Gyrd-Hansen (2004) finds that how individuals value health increments depends on whether the question is framed as an individual or social choice. Richardson and Nord (1997) find evidence that individuals feel that the distributional consequences of health programs are important and should be included in the evaluation of any health policy. The difficulty of using private choices to make public policy is clearly illustrated in Ubel et al. 1996). These authors find that subjects in an experimental setting strongly rejected the health rationing choices derived from their own utility responses. Finally, Nord (1994) argues that for the QALY to be a generally empirically meaningful concept, it needs to be interpreted as a measure of social value, rather than of private value.[1]

The findings of this paper, related papers, and the QALY literature have important policy implications. When policy makers allocate resources devoted to public health policies, several strategies are possible. One strategy is to observe individual preferences about tradeoffs between

---

[1] QALYs are most useful for cost-effectiveness analysis of alterative medical therapies. They focus on physical measures of health status and involve the standardization of health decrements relative to a year of perfect health (where death is 0 and perfect health is normalized as 1.) For a brief overview, see Appendix A, available from the authors.

private risk and private income and then use these preferences to make public policy. A policy maker who relies on estimates of the value of a statistical life (VSL) to make public policy would be following this strategy. Another strategy would be to directly estimate the demand (or WTP) for the public health program in question and compare aggregate WTP to the cost of the policy. Similarly, the policy maker may simply hold a referendum on a proposed project. Researchers have noted that using private preferences to make public policy may be problematic. For example, Ubel et al. (1996) provide experimental evidence that in individuals may soundly reject the public policy implications of their own preferences.

The research reported in this paper focuses on directly estimating the demand for public health policies. We feel that this strategy has several important advantages. First, individual preferences for public policies that reduce risk may be very different from the preferences that individuals have for private risk reductions. For example, individuals may be willing to pay for a policy that reduces drinking water contaminants because in reduces risk, provides ecosystem benefits, or because individuals feel that clean drinking water is a "right" that should be enforced by government. Individuals may also have altruistic motives or different notions of fairness that influence WTP for public polices. This issue is especially important when considering polices that affect children, the elderly, or economically disadvantaged groups. Assessing the demand for public health policies directly also allows us to see how heterogeneity in preferences for private risk-reducing polices compares with the heterogeneity in preferences for public polices.

Another advantage of our research strategy is that it allows us to assess what attributes of public health policies individuals view as desirable, as well as how different sociodemographic groups value different kinds of public health policies.

In this study, we use data collected from two analogous surveys of demand for health-related public policies. These surveys were designed to allow us to compare preferences for treatment policies with preferences for prevention policies. The surveys were designed by Trudy Ann Cameron at the University of Oregon Economics Department and J.R. DeShazo at UCLA. The data were collected, either via computer or Web-TV interface, by Knowledge Networks, Inc. The hypothetical treatment and prevention policies presented to respondents follow a randomized design that allows for the investigation of heterogeneity in preferences along several dimensions. The analysis is also enriched by the availability of a wide variety of individual-level sociodemographic variables.

Our analysis uses data from two conjoint stated preference surveys of demand for risk reducing policies that were administered to a nationally representative sample of over 1,500 individuals each. In addition to respondent's answers to the policy questions, we elicited individual-specific measures of the incidence of the perceived private benefits of each policy as well as a measure of attitudes toward government intervention.

The basic empirical framework used for analyzing respondents' stated survey choices is developed in the context of the prevention policy survey in Bosworth, Cameron, and DeShazo (2005) (hereafter BCD). In the present paper, we develop complementary analyses for the analogous treatment policy survey and compare the two types of demands.

While both prevention policies and treatment policies can lead to improved community health outcomes, the presence of systematic differences in consumer preferences for treatment and prevention policies would indicate that resources could perhaps be allocated more efficiently. We seek to establish key differences and similarities across policy types and to provide policy

6

makers with improved information about the potential welfare effects of different types of policy options.

Both the prevention and treatment scenarios presented to respondents vary in elicitation format, as well as in the specific type of illness threat that is addressed, the number of individuals who would benefit from the policies, and the size of the affected community. Basic respondent-level variables include the age, gender, income, education level, and ethnicity of the respondent. We describe the survey design in detail below.

2 Survey Design

In 2003, we conducted two distinct national stated preference surveys. For each, the sample size is approximately 1,500.[2] The two surveys that provide the data used for this paper were designed to elicit demands for policies which are publicly financed and which benefit many individuals (i.e., public goods), rather than privately paid programs with just individual benefits:

(a) The public "prevention" survey concerns policies that reduce contaminants that cause illness (i.e., air pollution, water pollution, food safety problems; see Cameron and DeShazo, 2005a). In terms of broader impacts, we intend these prevention policies to be analogous to the real public policies that lie within the purview of the Environmental Protection Agency and the Department of Agriculture.

(b) The public "treatment" survey concerns public provision of remedial medical interventions to individuals who are ill or injured and which increase their chances of recovery (i.e. devices, therapies, and procedures; see Cameron and DeShazo, 2005b). In terms of broader impacts, these treatment policies are intended to mirror the kinds of real public policies

---

[2] A third survey was also conducted, concerning demands for priced programs that benefit only the purchaser (i.e., private goods). A large pre-test for this third survey, involving over 1000 respondents, was conducted for Canadian consumers.

achieved, for example, through the regulatory decisions of the Food and Drug Administration or the medical research funding decisions of the National Institutes of Health.[3]

For conformability, the survey instruments for the prevention and treatment studies have initial and concluding modules that are very similar. Where they differ is in the key policy choice scenarios. Each policy is described as preventing or treating a named illness or injury. The illnesses in the prevention survey are attributed to a particular exposure pathway (i.e., air, water, food). The effectiveness of these policies is described in terms of the numbers of illnesses prevented (or successfully treated) and the number of deaths prevented. For the individual's community to enjoy the policy, he or she must pay costs in the form of higher taxes (expressed both per month and per year). Each of the five choice sets consists of two explicit policies plus the option to choose neither. We planned these two surveys so that it would be straightforward to pool their data, to test whether the subsets of corresponding utility parameters are identical, and to impose common preference parameters as warranted.

Details about the two surveys:

Module 1 (Introduction) - In the first modules of both the prevention and the treatment surveys, respondents are asked:

   (a) whether they have themselves suffered from each of a range of health threats, or whether family members or friends have experienced these problems.

   (b) to think about their family health histories, and to assess their own degree of risk from each of seven major classes of health threats.

---

[3] Of course, much research and development for medicines is also undertaken privately by pharmaceuticals firms, but these products must still be approved by the FDA.

(c) whether there is room to reduce their health risks by improving their lifestyle or habits,

(d) whether such improvements would reduce their risks of each of a range of health threats

The treatment survey asks the respondent to rate the likelihood that they would recover from each of a list of illnesses if they experienced it, given the quality of their current health care plan.

Both surveys then ask respondents to put aside their personal health concerns and to rate the prevalence of each class of illness and injury in their community (with "community" defined for them explicitly as a randomly assigned number of people living around them).

Module 2 (Tutorial) - The second module of each survey introduces the ideas of public prevention policies and public treatment policies, according to the topic of the survey, and begins to train the respondent how to interpret the summaries of policy attributes that will eventually be incorporated into compact choice tables. Eight pages of the prevention survey (and eleven pages of the treatment survey) are devoted to the tutorial process, where the information to be summarized in each row of the upcoming choice sets is unfolded one row at a time, with careful and clear explanations. These tutorial pages also include comprehension-testing questions to confirm that the respondent understands key attributes of the choices.

Module 3 (Choice Scenarios) - The third module of both surveys contains the five-different stated choice exercises. Each choice, with its preamble and debriefing questions, occupies a set of four survey pages.

In the prevention survey: The complex choice table is preceded by a page that first describes each policy in words, such as "Policy B reduces types of pesticides in foods that cause adult leukemia. New growing techniques and standards would reduce food contaminants that cause leukemia in adults." On the next page, the respondent studies the complete set of attributes of the two alternatives and makes a choice (which can include selection of "Neither policy"). See Figure 2 for an example of a prevention policy choice set.

The key attributes of the hypothetical prevention policies presented to respondents in our prevention survey include the number of cases (illnesses) prevented, the number of deaths avoided, the duration of the policy and the cost of the policy to the respondent. These prevention policies also vary in terms of the underlying cause to which the health effects are attributed: (e.g. an environmental cause (water contaminants, air pollution, pesticides in food); or a non-environmental cause (traffic accidents)). Prevention policies also vary by the specific type of illness or injury that is addressed. These include: cancer (general), colon/bladder cancer, leukemia, asthma, heart disease, heart attack, lung cancer, stroke, respiratory disease, and traffic accident injuries.

In the treatment survey: Each choice exercise also involves a set of four survey pages. The first page again describes the two policies in more detail. For example, "Policy A treats children, adults, and seniors who have leukemia. Those helped will be 25% children, 25% adults, and 50% seniors (i.e. 25/25/50 mix). Then the choice table is presented, containing its summaries of each program. See Figure 3 for an example of this type of choice set.

The duration of the policy and the policy cost are also key attributes of the policies described in the treatment survey. However, rather than the number of avoided illnesses or

deaths, the treatment policies include the number of increased recoveries as well as the number of avoided deaths. Treatment policies vary in terms of the demographic group that would most benefit from the policy (men, women, children, adults, seniors, or some combination of these groups) as well as the specific health threat addressed. For the treatment policies the list includes: prostate cancer, breast cancer, colon/bladder cancer, leukemia, lung cancer, asthma, heart attack, heart disease, stroke, respiratory disease, traffic injuries, and skin cancer.

In both surveys, to follow up on the choice exercise, the respondent is then asked how difficult it was for them to make up their mind on the previous screen with the choice table, and is then asked to reply directly to the question "To what extent would each policy directly benefit you or your family?" This question was asked about each of the two policies in the choice set just considered. Finally, any respondent who selected the "Neither policy" alternative was given an opportunity to check which reasons explain their choice. Some of the available answers constitute reasons that reveal choice-scenario rejection on the part of the respondent (e.g. disbelief that the policy would achieve what was advertised).[4]

Module 4 (Follow-up) – This module of each survey asks a number of auxiliary questions. Among these, the most relevant one for this paper is the question that invites respondents to both the prevention and treatment surveys to rate how involved they feel their government should be in regulation environmental, health, and safety hazards.[5]

The final three pages of each survey instrument are devoted to a hypothetical choice about how to take some lottery winnings, either as a lump sum now, or as a series of payments

---

[4] These answers can be used to limit the estimating sample, if no other economically admissible reason for choosing "Neither Policy" is selected.

[5] See Cameron and DeShazo (2005a) and (2005b) for more about Module 4.

spread out over several years. These choices are used to estimate individual-specific discount

rates. An understanding of these individual discount rates is important to our analysis of the

policy choices in both the prevention and the treatment studies, since the prevention and

treatment policies under consideration in the choice sets in each survey are described as having

different durations.

3   Theoretical Framework

This section explains a simple framework that will allow us to analyze, simultaneously,

preferences for prevention policies and treatment policies. As summarized in Figure 1,

individuals can be in one of three "health" states: healthy, sick, or deceased. Prevention and

treatment policies can both help to decrease the rate of flow from the "sick" to the "deceased"

state. Prevention policies work by decreasing the flow from the "healthy" state to the "sick" state,

while treatment policies work by increasing the flow from the "sick" state to the "healthy" state.

As part of our analysis, we test for systematic differences, across both respondents and health

threats, in individuals' implied preferences over how these flow changes are achieved. We also

provide estimates of the relative values of changes in various flows.

Figure 1
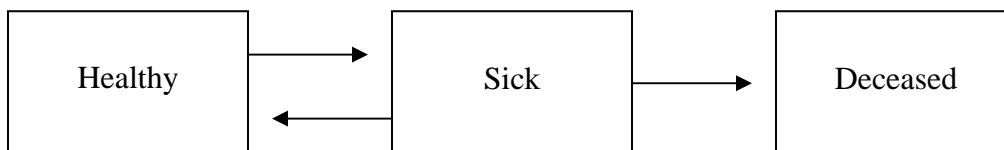
For our most basic specifications, we let the utility of a policy (treatment or prevention)

depend on the number of avoided illnesses (or increased recoveries), the number of avoided

deaths, and the duration of the policy (the length of time the policy is in effect). The individual's

income can also be expected to influence utility. Thus, individual i's indirect utility from policy j can be represented as:

$$V_{ji} = \beta(Y_i) + \delta_1 f(Avoided\ Illnesses_{ji})$$
$$+ \delta_2 g(Avoided\ Deaths_{ji}) \qquad (1)$$
$$+ \delta_3 h(Duration_{ji})$$

Where Y_i is income, β is the marginal utility of income, δ_1 is the marginal utility of an increase in f, δ_2 is the marginal utility of an increase in g, and δ_3 is the marginal (dis)utility of an increase in h.[6] δ_3 captures preferences over the length of time the policy lasts, where δ_3 can be interpreted as the marginal disutility experienced when the benefits of the policy are spread more thinly across time.[7]

We also employ the dummy variable POL_j --equal to 0 if the alternative is "neither policy" (i.e., the status quo) and equal to 1 if it is one of the policy options. The coefficient on this dummy variable, θ, serves a function similar to that of an intercept shifter in a regression model. It captures the average effect on person i's utility of all other unobserved factors, associated with any affirmative policy in the choice set, for which we do not explicitly control in the random-utility model. θ merely shifts the entire utility level and is interpreted as the average effect of unspecified factors on utility of any policy j *relative* to the status quo (Train 1986, pp. 21-27). Allowing θ to vary systematically with individual- or policy-specific attributes, as we will do, increases flexibility in estimation without sacrificing the utility-theoretic foundations of the model.

---

[6] An obvious objection to this simple linear-in-attributes specification is the implicit assumption that utility is additively separable in these generic functions of avoided illnesses and deaths. In empirical work documented in BCD (2005), we relax this assumption by including an interaction term between our functions of avoided illnesses and avoided deaths. To minimize the complexity of our combined prevention and treatment specification, we omit the interaction term in this paper. Of course, more elaborate specifications can be entertained.

[7] BCD (2005) develops a structural utility-theoretic model with constant exponential discounting employed explicitly. Since the ad-hoc model presented here provides a better fit, and is more comparable to previous research, we use it in the empirical portions of this paper.

Each choice set consists of two possible policies and a status quo alternative. We employ a random-utility model that permits analysis with a multiple-conditional logit specification for econometric estimation. To allow for a range of flexible estimation options, we assume at this point only that $f(0) = 0$, $g(0) = 0$, $h(0) = 0$ and that f, g, and h are increasing in their arguments. The utility level provided by policy j to individual i is thus:

$$
\begin{aligned}
V_{ji} = \beta(Y_i - c_{ij}) &+ \delta_1 f(Avoided\ Illnesses_{ji}) \\
&+ \delta_2 g(Avoided\ Deaths_{ji}) \\
&+ \delta_3 h(Duration_{ji}) + \theta POL_j + \eta_{ji}
\end{aligned}
\tag{2}
$$

Where c_{ji} is the annual cost of the policy and η_{ji} is the unobserved random component of total utility. Total indirect utility over the time period of the policy, if the status quo option (neither policy) is chosen, is given by:

$$
V_{ni} = \beta(Y_i) + \eta_{ni}
\tag{3}
$$

Since we assume that $f(0)=0$, $g(0)=0$, and $h(0)=0$, it is convenient to normalize on the level of indirect utility derived under the status quo. The perceived indirect utility difference that we assume drives the stated choices of our respondents is:

$$
\begin{aligned}
\Delta V_{ji} = \beta(-c_{ij}) &+ \delta_1 f(Avoided\ Illnesses_{ji}) \\
&+ \delta_2 g(Avoided\ Deaths_{ji}) \\
&+ \delta_3 h(Duration_{ji}) + \theta POL_j + \varepsilon_{ji}^*
\end{aligned}
\tag{4}
$$

where the η_{ji} are distributed extreme value and $\varepsilon_{ji}^* = \eta_{ji} - \eta_{ni}$.

For this homogeneous-preferences case, it should be noted that the parameters β, δ_1, δ_2, and δ_3 represent the marginal indirect utilities that individuals associate with the attributes of the policy while the parameter θ represents overall increment to utility provided by any policy, regardless of the other attributes. We can allow each of these marginal utility

parameters to vary depending on whether the policy is a treatment or prevention policy and can identify whether differences in preferences are attributable to different characteristics of the policies or to a general difference in the value placed on any type of policy, independent of its particular attributes.

There is no *a priori* expectation that the error dispersion in the choice model for prevention policies will be identical to the error dispersion in the choice model for treatment policies, although this is a testable hypothesis. If we wish to test the equivalence of the marginal utility parameters across the two samples, it will be necessary to allow for distinct error variances. (See Cameron et al. (2002)). We scale the level of indirect utility for the prevention policies and treatment policies by κ_p and κ_t, respectively. Let $1(Treatment)$ be a dummy variable equal to 1 for treatment policy choices and equal to 0 for prevention policy choices. Thus, the indirect utility differences for the prevention policies and treatment policies are:

$$\Delta V_{ji}\Big|_{1(Treatment)=0} = \frac{\beta_p}{\kappa_p}(-c_{ij}) + \frac{\delta_{1p}}{\kappa_p}f(Avoided\ Illnesses_{ji})$$
$$+ \frac{\delta_{2p}}{\kappa_p}g(Avoided\ Deaths_{ji}) \qquad (5)$$
$$+ \frac{\delta_{3p}}{\kappa_p}h(Duration_{ji}) + \frac{\theta_p}{\kappa_p}POL_j + \frac{\varepsilon_{ji}}{\kappa_p}$$

$$\Delta V_{ji}\Big|_{1(Treatment)=1} = \frac{\beta_t}{\kappa_t}(-c_{ij}) + \frac{\delta_{1t}}{\kappa_t}f(Avoided\ Illnesses_{ji})$$
$$+ \frac{\delta_{2t}}{\kappa_t}g(Avoided\ Deaths_{ji}) \qquad (6)$$
$$+ \frac{\delta_{3t}}{\kappa_t}h(Duration_{ji}) + \frac{\theta_t}{\kappa_t}POL_j + \frac{\varepsilon_{ji}}{\kappa_t}$$

Given that the scale of utility is arbitrary, we normalize by assuming $\kappa_p = 1$ for the prevention data set. The parameter κ_t is freely estimated and is interpreted as the ratio of dispersion of the unobserved portion of utility in the treatment sample to the dispersion of the unobserved portion of utility in the prevention sample.

We wish to investigate whether the parameters β_p, δ_{1p}, δ_{2p}, δ_{3p}, and θ_p are systematically different from β_t, δ_{1t}, δ_{2t}, δ_{3t}, and θ_t, based on inferences from respondents' choices among prevention policies and among treatment policies (from separate samples). By introducing the dummy variable **1**(Treatment) as a shifter we can permit the marginal utilities of avoided illness, avoided deaths, and policy duration vary systematically. With the pooled sample, we can test for statistically significant differences in preferences. Of course, we can (and do) allow all parameters (including κ) to vary with other individual- or policy-specific characteristics as well.

3.1 Willingness to Pay

In the deterministic case, formulas for total WTP and marginal WTP are straightforward. Point estimates of total WTP can be calculated by solving for the annual payment that would make the individual just indifferent between (a) paying for the policy and receiving the benefits, and (b) not paying for the policy and not receiving the benefits. Suppose we ignore the symmetric and mean zero error term and the variance-covariance matrix for the maximum likelihood estimates of the unknown preference parameters. We can set the utility difference in equation (4) equal to zero and solve for c_{ij}^{*} in terms of the parameter point estimates and the data. Total WTP for policy j is thus:

$$c_{ji}^* = \frac{\delta_1 f(\textit{Avoided Illnesses}_{ji}) + \delta_2 g(\textit{Avoided Deaths}_{ji}) + \delta_3 h(\textit{Duration}_{ji}) + \theta POL_j}{\beta} \quad (7)$$

Marginal WTP (MWTP)--WTP for incremental changes in one of the attributes of the policy--is calculated by taking the derivative of total WTP with respect to that attribute. For example, MWTP for one-unit increase in $g(\textit{Avoided Deaths}_{ji})$ is simply:

$$\frac{\partial c_{ji}}{\partial g(\textit{Avoided Deaths}_{ji})} = \frac{\delta_2}{\beta} \quad (8)$$

In our empirical work, we use a shifted log specification for the functions f, g, and h. The MWTP in equation (8) above is therefore roughly interpreted as marginal willingness to pay for a 1% increase in avoided deaths.[8]

4   Results

There are five different threads to our empirical results. Section 3.1 discusses our most basic specifications; Section 3.2 describes effects related to the size of the affected population. Section 3.3 details results relating to the socio-economic status of the respondent. Section 3.4 discusses differences in preferences for cancer and non-cancer policies. Finally, section 3.5 reports the results of models designed to investigate preference heterogeneity according to policy attributes.[9]

4.1   Basic Specifications

For our most basic specifications, reported in Table 2, we follow the standard practice in the choice literature and specify a utility function that is linear and additively separable in some

---

[8] Crude confidence bounds of fitted WTP and MWTP, reflecting estimation precision, can always be calculated by sampling from the joint (asymptotically normal) distribution of the maximum likelihood parameters and building up a sampling distribution for each calculated quantity. Of course, since total WTP is a function of policy attributes (see equation (7)), this sampling distribution will differ across policies.

[9] In a separate paper using only the "prevention" survey (BCD 2005), we submit our inferences to numerous robustness and validity checks. We also assess scope effects, order effects, sample selection biases and, through our survey design, attempt to mitigate hypothetical bias associated with incentive incompatibility.

function of the fundamental attributes of the policy. These fundamental attributes include the number of avoided illnesses or increased recoveries, the number of avoided deaths and the duration of the policy. In models where we pool the data from the two samples, we test whether or not the marginal utilities associated with illnesses, deaths, and the duration of the policy can be constrained to be the same across treatment and prevention policies by allowing the coefficients on these fundamental attributes (as well as the generic policy dummy) to vary systematically with a dummy variable that is equal to one if the policy is a treatment policy. We also allow the error variance to differ across policy type. [10,11,12]

In Table 2, the negative and statistically significant coefficient on the interaction term between the Log(*Death Reductions*) variable and the treatment dummy, in models 3 and 4, suggests that respondents place a higher value on deaths avoided via prevention policies than treatment policies. However, the other utility parameters (including the marginal utility of avoided illnesses) are not statistically different for treatment and prevention policies.

The estimates for our parsimonious model 4 in Table 2 imply that (yearly) marginal WTP for a 1% increase in avoided deaths via a prevention policy is about \$238, compared to about \$142 for the same 1% increase in avoided deaths via a treatment policy. These estimates are, unsurprisingly, almost identical to the estimates obtained for models 1 and 2 for the two separate samples. The estimates in model 1 (the prevention sample) indicate MWTP of \$245 for a 1% increase in avoided deaths, while model 2 (the treatment sample) indicates MWTP of \$138.

---

[10] For estimation purposes, we constrain κ to be positive by estimating the logarithm of this parameter. The estimates of κ in the tables below need to be exponentiated to conform to the model presented in section 2.
[11] Models that pool the prevention and treatment samples (or allow for heteroscedasticity in other contexts) are estimated via a heteroscedastic conditional logit optimization routine programmed by the authors using the software package Matlab. Models that do not involve heteroscedasticity are estimated using the packaged conditional logit routine in Stata software. The Matlab code is validated for special cases where either type of software can be used.
[12] We use a shifted log specification to maintain the assumption that *f(0)=0* and *g(0)=0.* For example, we use log(Avoided Illnesses+1) rather than log(Avoided Illnesses)

MWTP estimates for a 1% increase in avoided illnesses, in all models in Table 2, are about $70. We also note that the variance of the errors associated with choices concerning treatment policies appears to be larger by a factor of about $e^{0.3639}$=1.439.

Loosely speaking, the question posed in the title of this paper may be answered as follows: In terms of the estimated marginal utility avoided deaths, an "ounce" of prevention appears to be worth only about two "ounces" of cure in the sense that individuals appear to be willing to pay about twice as much for an incremental (1%) improvement in the number of deaths avoided via a prevention policy than they are for the same incremental improvement via a treatment policy. However, it should be noted that this difference applies only to avoided deaths. This result stands in contrast to that of Corso et al. (2002) who find that respondents to general allocation questions are willing to pay much more for treatment programs than for prevention programs.

There appears to be no statistical difference between the marginal amount individuals are willing to pay to avoid illnesses and the amount they are willing to pay to increase the number of recoveries. In terms of the diagram in Figure 1, individuals are willing to pay about twice as much to decrease the flow from the "Sick" state to the "Deceased" state via prevention policies as they are willing to pay to achieve the same net result via treatment policies. However, they are willing to pay about the same amount to increase the flow from the "Sick" state to the "Healthy" state as they are to decrease the flow from the "Healthy" state to the "Sick" state.

We also note that in all models in Table 2, the estimated marginal utility associated the duration of the policy is statistically significantly negative, indicating that individuals generally prefer policy of shorter duration (holding the total number of avoided illnesses and deaths constant). This result is consistent with positive discounting and is similar to that found by Alberini et al. (2004) who find that WTP to reduce mortality risk declines with latency. Ariely

and Lowenstien (2000) show that in most cases individuals underweight the importance policy latency, but that models (such as the research reported herein) that carefully and explicitly describe policy attributes can increase the likelihood that respondents discount future benefits. Cropper, et al. (1994) also finds that individuals generally values lives saved in the future less than lives saved in the present.

4.2   Population Size Effects

While Table 2 provides an initial answer to the question that headlines this paper, there are a number of additional insights that can be gleaned from more-general models built upon the same framework. For example, most previous studies that ask respondents to value policies that save a given number of lives do not ask respondents to consider the population size of the affected community. The policies presented to respondents in our study offer potential reductions in the number of illnesses and/or deaths in the community where the respondent lives. All policies considered by a given respondent are described as affecting "their community", where the community is described as a specified number of people living around the respondent. This asserted community size is varied randomly across respondents. The results related to population size heterogeneity reported in Table 3 do not exhibit the strong statistical significance characteristic of the other results reported herein. This is unsurprising, however, given the fact that population size is not a line-item attribute in the survey and that variation in asserted population size occurs only across respondents, rather than policies or choice sets.

We endeavor in this study to evaluate willingness-to-pay for health improvements from a public-goods perspective, so we must consider the size of the affected population. Note that this issue does not arise in studies that attempt to estimate only private trade-offs between health and

income. To make this issue clear, consider a simple example: Suppose that the leaders of a community of 100,000 people are considering two policies (policy A and policy B) that are each expected to reduce the number of deaths in the community. Policy A is expected to save the lives of 100 individuals in the community by reducing the risk level of each individual in the community. Policy B is also expected to save 100 lives, but the risk reductions from Policy B will accrue only to the inhabitants of the western side of the community. (i.e. only 50,000 people will see their risk level reduced.)

Even if the two policies cost the same, there is no *a priori* reason why a person (or a community) should be indifferent between the two programs. The decision maker would be choosing between providing a relatively large risk reduction to a smaller number of people, or providing a relatively small risk reduction to a larger group of people. In our example, we would expect that the 50,000 people on the west side of the community would prefer policy B to policy A (if they are selfish), while the rest of a selfish population would presumably prefer policy A to policy B. However, individual notions of fairness or altruistic preferences may lead to valuations that differ from the purely selfish outcome. If individuals are concerned only about their own health and income, we might expect that they would be willing to pay more for a program that affects a smaller population, *ceteris paribus*.

Table 3 thus presents the results of more-general models that allow the estimated utility parameters in our model to vary with the size of the population that will be affected by the policy. Models 1 and 2 again show results for analogous models estimated on our separate samples. In the less-restrictive pooled specification of model 3, we constrain the basic utility parameters (the ones that Table 2 suggests can be constrained) to be the same across treatment and prevention policies, but allow them to vary systematically with the size of the affected population. In the

prevention sample, the marginal utility of avoiding illnesses appears to be lower when the population size is larger. We also note that the coefficient on the interaction with population size and the policy dummy appears to be statistically significantly negative in models 1 and 4, indicating that individuals are less likely to choose either offered policy over the status quo option when the affected population size is larger. Both of these effects suggest selfish behavior. This result is consistent with the idea that social discount rates may be smaller than private discount rates: a larger population size probably causes the specified health improvements to be viewed as less of a private good and more of a public good. In other words, if it is not your life that is saved, it doesn't matter as much *when* that life is saved. However, this tendency is not apparent in the treatment sample.

4.3   Sociodemographic Effects

We now investigate how preferences for prevention policies appear to differ from preferences for treatment policies according to the sociodemographic characteristics of the respondent. Tables 4, 5, and 6 present these results. To keep the dimensionality of the parameter space manageable, we allow only the coefficient on the policy dummy to vary with the sociodemographic variables. Recall that the coefficient on the policy dummy captures how individuals feel about any policy, relative to the no-policy status quo. Coefficients on the interaction terms in Table 4 can be roughly interpreted as capturing the effect of change in a given variable on the latent propensity to choose either of the two offered policy options over the status quo.

Previous work that investigates how WTP for health benefits varies with the sociodemographic characteristics of the individual includes Alberini et al. (2002), and DeShazo

and Cameron (2005), Alberini et al. (2002) find that WTP declines with age, but only after age 70. DeShazo and Cameron (2005), however, find that WTP follows an inverted U-shaped profile. Kartman et al. (1996) find that income is positively related to WTP to reduce the risk of angina pectoris attacks. It should be noted, however, that these authors investigate WTP for private risk reductions rather than the public choices considered here.

The treatment sample indicates lower WTP for females. This may reflect lower incomes and higher marginal utility of income for women (not estimated here) or may reflect different propensities to avail themselves of private health care services and diagnostic procedures. In fact, the magnitude of this effect is relatively large in terms of estimated WTP. The estimates in model 2 in Table 4 indicate that females are willing to pay about $248 less per year for a typical treatment policy than males. However, there is no indication that females are less likely than males to choose either prevention policy over the status quo in the prevention sample.

In contrast to the inverted U-shaped age profile found by DeShazo and Cameron (2005), the separate prevention and treatment samples (as well as the pooled model) indicate that willingness to pay has a U-shaped age profile. This profile reaches an estimated minimum at about age 60 in the prevention sample and at about age 71 in the treatment sample. The curvature of the age profile is also significantly less for the treatment sample.

The models in models 1 and 2 indicate that the income of the respondent, as a proxy for general socioeconomic status, has *opposite* effects on the estimated WTP of the respondent in the two samples. In particular, higher income individuals are *less* willing to pay for prevention policies, while they are *more* willing to pay for treatment policies. These effects, while statistically significant, are relatively small in magnitude: the estimates in model 1 indicate that a $10,000 increase in annual income is associated with an estimated decrease of $0.34 in annual

WTP for a typical prevention policy. Similarly, the estimates in model 2 indicate that a $10,000 increase in annual income is associated with an estimated increase of $0.44 in annual WTP for a typical treatment policy. One plausible explanation for this difference may be the availability of substitutes. Recall that the prevention policies work in one of four ways: cleaner air, cleaner water, fewer pesticides in food, and safer roads. A high income individual can more easily move to a cleaner location, drink bottled or filtered water, eat organic produce, and purchase safer automobiles. However, there are relatively fewer substitutes for prevention policies that lower-income people can exploit.

The results for years of education suggest that more highly educated people are more likely to support prevention policies, but not treatment policies. The estimates in model 1 suggest that, for prevention policies, one additional year of education is associated with an estimated increase of $128 in annual WTP for a typical prevention policy.

Non-white individuals are generally more likely to support both prevention and treatment public policies over the status quo, and more likely to support treatment than prevention policies. The estimates from models 1 and 2 suggest that non-white individuals are willing to pay an estimated $299 per-year more than non-white individuals for prevention policies and $660 more for treatment policies.

Attitudes toward government intervention have a lot to do with individuals' receptivity to publicly supported health policies. Tables 5 and 6 present separate results for the prevention and treatment samples, and demonstrate the impact of the additional variable *Government Preference*. After the choice scenarios, we presented individuals with the following question: "People have different ideas about what their government should be doing. How involved do you feel the government should be in regulating environmental, health and safety hazards?" Individuals were

24

invited to indicate their preferred level of government involvement along a continuum ranging

from minimally involved (0) to heavily involved (7). While this variable is merely ordinal, we

limit the complexity of our estimating specification by treating it as an approximately continuous

variable.

The results in Tables 5 and 6 make the statistical importance of this (endogenous)

variable clear. The maximized value of the log-likelihood function is higher (in both the

treatment and prevention samples) when the variable *Government Preference* is included as a

single shifter on the θ parameter (model 2) than when the entire suite of other sociodemographic

variables is included (model 1). Moreover, the third models of Tables 5 and 6 demonstrate that

there is almost no impact on the statistical significance of the other sociodemographic variables

when the *Government Preference* variable is included.

We conclude from this analysis that although the basic sociodemographic characteristics

of the respondent are important in determining choices, an individual's perception of the proper

role of government is relatively more predictive of their stated choices across proposed public

policies.

4.4   Cancer vs. Non-Cancer Policies

Previous research (e.g. Hammitt and Liu (2004)) has suggested that individuals may be

willing to pay more to reduce cancer risks than non-cancer risks, independent of the severity of

the symptoms of either type of disease.[13] Cancers may simply instill greater fear than other

diseases. We address this interesting question by assessing whether or not individuals are

(broadly speaking) more likely to support policies that address cancer risks than other non-cancer

---

[13] Other researchers, including Tsuge et al. (2005), and Magat et al. (1996) find that it may not be necessary to adjust
VSL estimates for cancer.

risks. Our prevention policies survey asks about several different diseases, including cancers (in general) lung cancer, colon/bladder cancer, and leukemia. The treatment policy sample is asked about colon/bladder cancer, leukemia, lung cancer, prostate cancer, breast cancer and skin cancer. We define an indicator variable, "Cancer," to be equal to 1 if the policy provides prevention or treatment with respect to a major cancer.[14] Table 7 presents results that utilize this variable to differentiate between preferences for cancer versus non-cancer policies.

The results in Table 7 suggest that individuals are *more* likely to support a cancer *prevention* policy than other types of policies, but *less* likely to support cancer *treatment* policies. Since many types of cancers are still viewed as incurable, these findings seem plausible.

WTP calculations based on the estimates in Table 7 suggest that, in the prevention sample, individuals are willing to pay about $310 more (annually) for a policy that avoids deaths and illnesses from a major cancer than from other (non-cancer) illnesses or injuries. In the treatment sample, however, estimates suggest that individuals are willing to pay about $130 less per year for policies that address major cancer risks than for other types of policies. The pooled sample provides similar estimates: increase WTP of about $280 for prevention of cancer, but about $160 less for the treatment of cancer.

4.5   Heterogeneity by Policy Attributes

Tables 8 and 9 present parameter estimates for models that allow the coefficient on the policy dummy to vary systematically with additional attributes of each policy. We find evidence that individuals have statistically distinguishable preferences for some types of policies.

---

[14] We exclude skin cancer from the list of "major" cancers because skin cancer is generally perceived as a less serious health threat than the other cancers considered in our survey.

Previous Research: Authors that investigate how WTP for public health benefits varies with the source of the risk include Vassanadumrongdee and Matsuoka (2005), Carlsson et al. (2004), and Chilton et al. (2002). Vassanadumrongdee and Matsuoka (2005) find that WTP for reductions in the risk of disease from air pollution and the risk of traffic accident are comparable while Chilton et al. (2002) find that the perception of risk influences WTP values for reducing the risk of rail accidents. Carlsson et al. (2004) find that individuals in their sample are willing to pay more to improve air travel safety than taxi travel safety. Subramanian and Cropper (2000) find that the number of lives saved, as well as psychological risk characteristics are important determinants of allocation decisions, and Krupnick and Cropper (1992) find that individuals who have had friends or family with chronic lung disease are willing to pay more to reduce the risk of chronic lung disease. Jacobsson et al. (2005) find that the altruistic component of WTP is greater for more severe diseases. Wittenberg et al. (2003) find that their respondents "were 10 to 17 times more likely to allocate liver transplants or asthma treatment to patients they deemed not responsible for their illnesses than to patients they deemed responsible for their conditions".

This study appears to represent the most comprehensive comparative analysis of systematic variation in WTP for public health policies to date. The policies in our survey vary, as reported above, in terms of basic attributes such as the number of lives saved of the cost of the policy. The policies presented to respondents also vary in terms of the source of the risk, the disease or health threat that is addressed, and the population sub-group that is affected, allowing respondents to consider a wide range of substitute policies when making allocation decisions. In Tables 8 and 9, we report the results of models that allow key utility parameters to vary with a variety of additional policy attributes.

Table 8 allows the coefficient on the policy dummy to vary by the type of disease. We have chosen heart disease as the baseline disease (the omitted category) because it is one of the most common causes of death and there are effective methods for both the prevention and treatment of heart disease. The estimated coefficients on the policy-dummy interaction terms in Table 6 are interpreted relative to the base case of heart disease.[15]

In the prevention sample, we see that individuals are more likely to support public policies that prevent cancer (general), leukemia in children, and asthma in children than they are to support heart disease policies. However, individuals are less likely to support prevention policies that address leukemia in general, stroke, asthma in general, and traffic injuries.

In the treatment sample, there are no public policies that are statistically significantly more likely to be supported than those for heart disease. However, public policies to reduce colon/bladder cancer, leukemia, stroke, respiratory disease, asthma, asthma in children, lung cancer, injuries, prostate cancer, and skin cancer are all less likely to be chosen relative to public heart disease treatment policies.

Finally, Table 9 shows results for models that utilize additional sources of heterogeneity in the type of health risk that are unique to either the prevention or the treatment surveys. As in the choice set example in Figure 2, the scenarios concerning the public prevention policies vary in terms of the underlying cause of the particular illness or injury. These causes include: air pollution, drinking water contaminants, pesticides in foods, and traffic accidents. The "prevention" model in Table 9 suggests that individuals prefer prevention policies that reduce air pollution, drinking water contaminants and pesticides in foods, (relative to policies that reduce the likelihood of injuries via traffic accidents).

---

[15] We report in Tables 8 just the prevention policy and treatment policy results, separately. The pooled model has a very large parameter space and offers few additional insights.

The "treatment" model in Table 9 presents results for a specification which reflects the fact that some treatment policies are targeted at specific socio-demographic groups. For example, breast cancer treatment policies primarily benefit women and prostate cancer treatment policies primarily benefit men. The results in the "treatment" model of Table 9 suggest, perhaps unsurprisingly, that females are statistically significantly more likely than males to support breast cancer treatment policies, while they are less likely than males to support prostate cancer policies.

Other types of policies may be targeted primarily at children, adults, or seniors. Notice in the choice set example in `Figure 3 that the choice scenarios presented to respondents make the targeted beneficiary group explicit. When policies are designed to benefit more than one group, the percentages of the benefits accruing to each group are included explicitly in the description of the choice. For example, policy A in the Figure 3 (the treatment choice set example) treats children, adults, and seniors who have leukemia. The percentage mix is given as 25/25/50, indicating that 25% of the benefits would accrue to children, 25% to adults, and 50% to seniors.[16] We construct the continuous variables *Percent Children* and *Percent Senior* and allow the coefficient on the policy dummy to vary systematically with these variables. We also utilize the variables "*Female*", "*Age65+*", and "*Kids*" to distinguish how preferences differ for treatment policies that affect particular groups. "*Female*" is equal to 1 if the respondent is female (and 0 otherwise), Likewise, "*Age65+*" is equal to 1 if the respondent is age 65 or older, and "*Kids*" is equal to 1 if the respondent lives in a household with any children under the age of 18.

The results in Table 9 suggest that respondents with children in the household are statistically significantly more likely to support policies that benefit children while policies that

---

[16] The tutorial portion of the treatment survey, which precedes these choice scenarios, explains the interpretation of these proportions.

benefit seniors are less likely to be supported. Interestingly, this apparent lack of support for policies that benefit seniors is also shared by seniors themselves.

5   Conclusion

Policymakers face many tradeoffs when allocating funds for public risk reduction and health improvement policies. Some policies can help prevent adverse health states while other policies can allocate resources to help treat those who are already sick or injured. We find that preferences for prevention and treatment policies differ in several important ways.

Individuals appear to have a preference for prevention policies over treatment policies. This preference appears to be driven by a higher marginal value placed on lives saved via prevention policies. We find that individuals are willing to pay about twice as much to avoid deaths via prevention than they are to avoid deaths via treatment.

We also find that the size of the affected population affects preferences for both treatment and prevention policies in ways that are generally consistent with selfish behavior. In particular, individuals are less likely to support prevention policies when the affected population size is larger. The size of the affected population has a much less pronounced effect on preferences for treatment policies.

We find evidence of significant heterogeneity in WTP for prevention and treatment policies according to differences in socio-demographic characteristics. We find that WTP has a U-shaped age profile that reaches a minimum at about age 60 for both types of policies. High income individuals are *more* likely to support treatment policies, while they are *less* likely to support prevention policies. This seemingly strange result may be the result of a wider array of preventative/risk mitigating options available to wealthier individuals.

We also note that more highly educated people are more likely to support prevention policies than less educated people, while there is no systematic heterogeneity in preferences for treatment policies by education level. Females are less likely to support treatment policies, while non-white (non-Caucasian) individuals are more likely to support both prevention and treatment policies.

Respondents in our sample are *more* likely to support prevention policies that address cancer risks than non-cancer risks, but are *less* likely to support major cancer treatment policies than policies that treat other major illnesses or injuries. Respondents in both samples are less likely to support policies that address stroke, leukemia, and asthma than policies that address heart disease. We also find that females are more likely to support breast cancer treatment policies and less likely to support prostate cancer treatment policies. Individuals with children are more likely to support policies that benefit children, but seniors are not more likely to support policies that benefit seniors.

We identify several areas of heterogeneity in preferences by individual and policy attributes and find that respondents are more likely to choose policies that directly affect themselves and/or their family members. We also find that individual perceptions of the proper role of government are significant in explaining whether or not individuals support policy changes over the status quo.

References

Alberini, A., Cropper, M., Krupnick, A., Simon, N. (2002). Does the value of a statistical life vary with age and health status? Evidence from the United States and Canada. Resources for the future discussion paper 02-19. Washington, D.C.

Alberini, A., Cropper, M., Krupnick, A., Simon, N. (2004). Willingness to pay for mortality risk reductions: Does latency matter? Resources for the future discussion paper 04-13. Washington, D.C.

Andreoni, J. (1990). "Impure altruism and donations to public-goods - a theory of warm-glow giving", *Economic Journal* 100, 464-477.

Arana, J.E., Leon, C.J. (2002). "Willingness to pay for health risk reduction in the context of altruism", *Health Economics* 11, 623-635.

Ariely, D., Loewenstein, G. (2000). "When does duration matter in judgment and decision making?" *Journal of Experimental Psychology-General* 129, 508-523.

Bosworth, R.C., Cameron, T.A., DeShazo, J.R. (2005). Advances in evaluating the demand for risk prevention policies. University of Oregon, Department of Economics Working Paper. Eugene, OR.

Cameron, Trudy Ann, and J.R. DeShazo (2005a) "Demand for Public Health-Risk Reduction Policies: The Prevention Survey" http://darkwing.uoregon.edu/~cameron/vsl/public_prevention_framed.pdf

Cameron, Trudy Ann, and J.R. DeShazo (2005b) "Demand for Public Health-Risk Reduction Policies: The Treatment Survey" http://darkwing.uoregon.edu/~cameron/vsl/public_treatment_framed.pdf

Cameron, T.A., Poe, G.L., Ethier, R.G., Schulze, W.D. (2002). "Alternative non-market value-elicitation methods: Are the underlying preferences the same?" *Journal of Environmental Economics and Management* 44, 391-425.

Carlsson, F., Johansson-Stenman, O., Martinsson, P. (2004). "Is transport safety more valuable in the air?" *Journal of Risk and Uncertainty* 28, 147-163.

Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Spencer, A. (2002). "Public perceptions of risk and preference-based values of safety", *Journal of Risk and Uncertainty* 25, 211-232.

Corso, P.S., Hammitt, J.K., Graham, J.D., Dicker, R.C., Goldie, S.J. (2002). "Assessing preferences for prevention versus treatment using willingness to pay", *Medical Decision Making* 22, S92-S101.

Cropper, M.L., Aydede, S.K., Portney, P.R. (1994). "Preferences for life saving programs - how the public discounts time and age", *Journal of Risk and Uncertainty* 8, 243-265.

Cummings, R.G., Taylor, L.O. (1998). "Does realism matter in contingent valuation studies?" *Land Economics* 74, 203-215.

DeShazo, J.R., Cameron, T.A. (2005). Two types of age effects in the demand for reductions in mortality risks with differing latencies. University of Oregon, Department of Economics Working Paper. Eugene, OR.

Fetherstonhaugh, D., P. Slovic, S. M. Johnson, and J. Friedrich. (1997). "Insensitivity to the value of human life: A study of psychophysical numbing", *Journal of Risk and Uncertainty* 14, 283-300.

Frederick, S., G. Loewenstein, and T. O'Donoghue. (2002). "Time discounting and time preference: A critical review", *Journal of Economic Literature* XL, 351-401.

Hammitt, J.K., Liu, J.T. (2004). "Effects of disease type and latency on the value of mortality

risk", *Journal of Risk and Uncertainty* 28, 73-95.

Horowitz, J.K., Carson, R.T. (1991). "A classification tree for predicting consumer preferences for risk reduction", *American Journal of Agricultural Economics* 73, 1416-1421.

Horowitz, J.K., Carson, R.T. (1993). "Base-line risk and preference for reductions in risk-to-life", *Risk Analysis* 13, 457-462.

Jacobsson, F., Carstensen, J., Borgquist, L. (2005). "Caring externalities in health economic evaluation: How are they related to severity of illness?" *Health Policy* 73, 172-182.

Johannesson, M., Johansson, P.O., Oconnor, R.M. (1996). "The value of private safety versus the value of public safety", *Journal of Risk and Uncertainty* 13, 263-275.

Kartman, B., Andersson, F., Johannesson, M. (1996). "Willingness to pay for reductions in angina pectoris attacks." *Medical Decision Making* 16, 248-253.

Krupnick, A.J., Cropper, M.L. (1992). "The effect of information on health risk valuations", *Journal of Risk and Uncertainty* 5, 29-48.

Magat, W.A., Viscusi, W.K., Huber, J. (1996). "A reference lottery metric for valuing health", *Management Science* 42, 1118-1130.

Nord, E. (1994). "The QALY - a measure of social value rather than individual utility", *Health Economics* 3, 89-93.

Richardson, J., Nord, E. (1997). "The importance of perspective in the measurement of quality-adjusted life years", *Medical Decision Making* 17, 33-41.

Subramanian, U., Cropper, M. (2000). "Public choices between life saving programs: The tradeoff between qualitative factors and lives saved", *Journal of Risk and Uncertainty* 21, 117-149.

Train, K. (1986). *Qualitative choice analysis*. Cambridge: The MIT Press.

Tsuge, T., Kishimoto, A., Takeuchi, K. (2005). "A choice experiment approach to the valuation of mortality", *Journal of Risk and Uncertainty* 31, 73-95.

Ubel, P.A., Loewenstein, G., Scanlon, D., Kamlet, M. (1996). "Individual utilities are inconsistent with rationing choices: A partial explanation of why Oregon's cost-effectiveness list failed", *Medical Decision Making* 16, 108-116.

Ubel, P.A., Loewenstein, G., Scanlon, D., Kamlet, M. (1998). "Value measurement in cost-utility analysis: Explaining the discrepancy between rating scale and person trade-off elicitations", *Health Policy* 43, 33-44.

Vassanadumrongdee, S., Matsuoka, S. (2005). "Risk perceptions and value of a statistical life for air pollution and traffic accidents: Evidence from Bangkok, Thailand", *Journal of Risk and Uncertainty* 30, 261-287.

Viscusi, W.K., Aldy, J.E. (2003). "The value of a statistical life: A critical review of market estimates throughout the world", *Journal of Risk and Uncertainty* 27, 5-76.

Wittenberg, E., Goldie, S.J., Fischhoff, B., Graham, J.D. (2003). "Rationing decisions and individual responsibility for illness: Are all lives equal?" *Medical Decision Making* 23, 194-211.

Figure 2 - Sample Prevention Choice Set

These two policies would be implemented for the 100,000 people living around you. Would you be most willing to pay for Policy A, Policy B, or neither of them?

	Policy A	Policy B
	reduces pesticides in foods that cause colon and bladder cancer	reduces air pollutants that cause heart attacks
Policy in effect	over 5 years	over 10 years
Cases prevented	100 fewer cases	200 fewer cases
Deaths prevented	10 fewer deaths over 5 years	5 fewer deaths over 10 years
Cost to you	$70 per month (= $840 per year for 5 years)	$6 per month (= $72 per year for 10 years)

Your choice

☐ Policy A
reduces pesticides in foods that cause colon and bladder cancer

☐ Policy B
reduces air pollutants that cause heart attacks

☐ Neither Policy

Next Question

Figure 3 - Sample Treatment Choice Set

Recall that these two policies will be implemented for the 100,000 people living around you. Below we describe how many of these people get sick and die, with and without these policies.

Would you be most willing to pay for Policy A, Policy B, or neither of them?

	Policy A	**Policy B**
	treats children, adults, and seniors (25/25/50 mix) who have leukemia	treats seniors who have heart disease
How many Policy will affect, and when	700 will get sick over 30 years	10,000 will get sick over 4 years
Increased Recoveries	25 more full recoveries	50 more full recoveries
Deaths prevented	5 fewer deaths over 30 years	5,000 fewer deaths over 4 years
Cost to you	$6 per month (= $72 per year for 30 years)	$35 per month (= $420 per year for 4 years)

Your choice

☐ Policy A
treats children, adults, and seniors (25/25/50 mix) who have leukemia

☐ Policy B
treats seniors who have heart disease

☐ Neither Policy

[Next Question]

Table 1: Basic Policy-Specific Variables (orthogonal)[a]

Variable	Mean	St. Dev.	Min	Max	Description
Yearly Cost					
Prevention	498	351	60	1200	Yearly cost of policy
Treatment	498	346	60	1200	Yearly cost of policy
Illness Reductions					
Prevention	862	1584	0	5000	Cases avoided
Treatment	841	1550	0	5000	Increased recoveries
Death Reductions					
Prevention	101	464	0	5000	Deaths avoided
Treatment	539	1240	0	5000	Deaths avoided
Duration					
Prevention	13.8	9.7	2	30	Length of policy (years)
Treatment	13.9	9.7	2	30	Length of policy (years)
Population Size					
Prevention	0.246	0.359	0.001	1	Size of affected population (millions)
Treatment	0.758	0.358	0.001	1	Size of affected population (millions)

[a]Except for exclusions based on implausible combinations

Table 2: Basic Specifications[b]

Parameter	Variable	Separate Samples		Pooled Sample	
		(1) Prevention Sample	(2) Treatment Sample	(3) Less Restricted	(4) More Restricted
β	-(Yearly Cost/10,000)	5.680	4.9854	6.1162	6.1091
		(9.72)***	(8.13)***	(10.90)***	(12.09)***
δ_1	Log(Illness Reductions)	0.04046	0.0358	0.04156	0.04305
		(5.77)***	(5.05)***	(5.92)***	(7.42)***
	… · **1**(Treatment)	--	--	.001797	--
				(0.48)	
δ_2	Log(Death Reductions)	0.1394	0.0692	0.1455	0.15206
		(11.16)***	(7.20)***	(11.72)***	(12.76)***
	… · **1**(Treatment)	--	--	-.05804	-0.07738
				(-3.26)***	(-6.15)***
δ_3	Log(Duration)	-0.1688	-0.1198	-0.1586	-0.17048
		(-7.39)***	(-5.14)***	(-6.92)***	(-8.72)***
	… · **1**(Treatment)	--	--	-.03868	--
				(-0.70)	
θ	Policy Dummy	-0.2371	-0.2657	-0.2649	-0.27051
		(-3.30)***	(-3.48)***	(-3.67)***	(-4.51)***
	… · **1**(Treatment)			-.04018	--
				(-0.27)	
$ln(\kappa)$	Heteroscedasticity Parameter	0	--	0	0
	… · **1**(Treatment)	--	0	.3659	0.35857
				(2.04)**	(4.01)***
Maximized log-likelihood		-8035.19	-7539.49	-15576.35	-15577.87
Maximized log-likelihood overall		-15574.68		-15576.35	-15577.87
Total sample size (choices)		7556	7033	14589	14589
Total sample size (respondents)		1531	1423	2954	2954

[b]All specifications use shifted log format for Log(X) variable. For example, Log(Death Reductions) is actually Log(Death Reductions +1)

[T]est stat for restrictions in Model 3: 3.34 Critical value: 3.84: Fail to reject restriction

Test stat for restrictions in Model 4: 3.04 Critical value: 7.81: Fail to reject restrictions

Table 3: Population Size Effects

Parameter	Variable	Separate Samples		Pooled Sample	
		(1) Prevention Sample	(2) Treatment Sample	(3) Less Restricted	(4) More Restricted
β	-(Yearly Cost/10,000)	6.2373 (8.51)***	3.5863 (1.57)	5.7920 (8.14)***	6.7188 (10.13)***
	… · Population Size[a]	0.0008 (0.00)	2.0745 (0.27)	3.5046 (0.70)	--
δ_1	Log(Illness Reductions)	0.0545 (5.99)***	-0.0277 (-0.51)	0.0452 (5.05)***	0.0486 (6.92)***
	… · Population Size	-0.0527 (-1.95)*	0.0885 (0.74)	0.0147 (0.50)	--
δ_{2p}	Log(Death Reductions)	0.1324 (8.40)***		0.1407 (8.85)***	0.1414 (9.62)***
	… · Population Size	0.1067 (0.11)		0.0668 (0.55)	0.0648 (1.27)
δ_{2t}	Log(Death Reductions)	--	0.0919 (2.66)***	0.0415 (1.47)	0.0352 (1.37)
	… · Population Size	--	-0.0271 (-0.41)	0.1111 (1.13)	0.1055 (2.32)**
δ_3	Log(Duration)	-0.2107 (-7.37)***	-0.1083 (-1.37)	-0.2011 (-7.24)***	-0.1901 (-7.56)***
	… · Population Size	0.1381 (1.39)	-0.0268 (-0.14)	-0.0199 (-0.17)	--
θ	Policy Dummy	-0.0967 (-1.06)	-0.4665 (-1.56)	-0.1444 (-1.61)	-0.1396 (-1.65)*
	… · Population Size	-0.5216 (-2.45)**	0.2623 (0.82)	-0.5030 (-1.25)	-0.4708 (-1.83)*
$ln(\kappa)$	Heteroscedasticity Parameter	--	--	--	--
	… · Population Size	0.4114 (0.67)	0.0597 (0.05)	0.5356 (0.98)	0.4407 (2.07)**
	… · **1**(Treatment)	--	--	0.1315 (1.27)	0.1320 (1.30)
Maximized log-likelihood		-8025.39	-7529.39	-15569.80	-15573.09
Maximized log-likelihood overall		15554.78		15569.80	-15573.09
Total sample size (choices)		7556	7033	14589	14589
Total sample size (respondents)		1531	1423	2954	2954

[a] Affected population size measured in millions.

Table 4: Sociodemographic Effects

Parameter	Variable	(1) Prevention	(2) Treatment	(3) Pooled
β	-(Yearly Cost/10,000)	5.7124	5.0104	5.8665
		(9.75)***	(8.16)***	(11.03)***
δ_1	Log(Illness Reductions)	0.0419	0.0355	0.0424
		(5.95)***	(4.99)***	(7.32)***
δ_{2p}	Log(Death Reductions)	0.1396	--	0.1406
		(11.14)***		(11.59)***
δ_{2t}	Log(Death Reductions)	--	0.0702	0.0840
			(7.29)***	(6.64)***
δ_3	Log(Duration)	-0.1688	-0.1204	-0.1601
		(-7.38)***	(-5.16)***	(-8.11)***
θ	Policy Dummy	-0.4793	-0.0264	-0.2737
		(-1.62)	(-0.09)	(-1.19)
	…· **1**(Female)	0.0203	-0.1244	0.0198
		(0.43)	(-2.53)**	(0.42)
	…· **1**(Female) ·**1**(Treatment)	--	--	-0.1705
				(-2.18)**
	…· (Age/100)	-2.4105	-1.8901	-3.2836
		(-2.32)**	(-1.89)*	(-3.72)***
	…· (Age/100) ·**1**(Treatment)	--	--	2.0492
				(2.20)**
	…· (Age^2/10,000)	2.0078	1.6335	2.8272
		(2.02)**	(1.71)*	(3.29)***
	…·(Age^2/10,000)·**1**(Treatment)	--	--	-1.8463
				(-1.90)*
	…· Income/10,000	-1.9524	2.2224	-1.8301
		(-2.57)**	(2.90)***	(-2.42)**
	…·Income/10,000·**1**(Treatment)	--	--	4.4788
				(3.62)***
	…·Educ.Years/10	0.7071	0.1046	0.6941
		(6.90)***	(0.98)	(7.06)***
	…·Educ.Years/10·**1**(Treatment)	--	--	-0.5557
				(-3.78)***
	…·**1**(Non-White)	0.17135	0.3770	0.1696
		(2.94)***	(6.00)***	(2.92)***
	…·**1**(Non-White)·**1**(Treatment)	--	--	0.2867
				(2.66)***
$ln(\kappa)$	Heteroscedasticity Parameter	0	--	--
	… · **1**(Treatment)	--	0	0.1931
				(1.62)
	Age effect minimized at	60.7	70.8	41.1(Prev) 128.3 (Treat)
Maximized log-likelihood		-7998.81	-7509.59	-15509.46

Table 5: Sociodemographic Effects (Prevention Sample)

Parameter	Variable	(1) SES Only	(2) Govt. Only	(3) SES and Govt.
β	-(Yearly Cost/10,000)	5.7175	5.9084	5.9274
		(9.75)***	(10.00)***	(10.01)***
δ_1	Log(Illness Reductions)	0.0419	0.0420	0.0432
		(5.95)***	(5.92)***	(6.06)***
δ_{2p}	Log(Death Reductions)	0.1396	0.1438	0.1438
		(11.14)***	(11.40)***	(11.37)***
δ_3	Log(Duration)	-0.1688	-0.1730	-0.1728
		(-7.38)***	(-7.51)***	(-7.50)***
θ	Policy Dummy	-0.4793	-1.3082	-1.4867
		(-1.62)	(-12.42)***	(-4.77)***
	…· Government Preference	--	0.2073	0.2022
			(14.09)***	(13.67)***
	…· **1**(Female)	0.0204	--	0.0103
		(0.43)		(0.21)
	…· (Age/100)	-2.4154	--	-1.9811
		(-2.32)**		(-1.87)*
	…· (Age^2/10,000)	2.0071	--	1.5260
		(2.02)**		(1.51)
	…· Income/10,000	-1.9524	--	-1.8089
		(-2.57)**		(2.34)**
	…·Educ.Years/10	0.7071	--	0.0619
		(6.90)***		(5.95)***
	…·**1**(Non-White)	0.1714	--	0.1329
		(2.94)***		(2.23)***
Maximized log-likelihood		-7998.81	-7875.33	-7847.22

Table 6: Sociodemographic Effects (Treatment Sample)

Parameter	Variable	(1) SES Only	(2) Govt. Only	(3) SES and Govt.
β	-(Yearly Cost/10,000)	5.0104	4.9810	4.9977
		(8.16)***	(8.07)***	(8.08)***
δ_1	Log(Illness Reductions)	0.0356	0.0381	0.0378
		(4.99)***	(5.32)***	(5.26)***
δ_{2p}	Log(Death Reductions)	0.0703	0.0695	0.0703
		(7.29)***	(7.17)***	(7.24)***
δ_3	Log(Duration)	-0.1204	-0.1208	-0.1215
		(-5.06)***	(-5.15)***	(-5.18)***
θ	Policy Dummy	-0.0264	-1.0415	-0.6918
		(-0.09)	(-9.76)***	(-2.31)***
	…· Government Preference	--	0.1524	0.1463
			(10.36)***	(9.86)***
	…· **1**(Female)	-0.1244	--	-0.1235
		(-2.53)**		(-2.49)***
	…· (Age/100)	-1.8962	--	-2.0106
		(-1.89)*		(-1.98)**
	…· (Age^2/10,000)	1.6330	--	1.7432
		(1.71)*		(1.81)*
	…· Income/10,000	2.222	--	2.2750
		(2.90)***		(2.94)***
	…·Educ.Years/10	0.1046	--	0.0710
		(0.98)		(0.66)
	…·**1**(Non-White)	0.3770	--	0.3312
		(6.00)***		(5.18)***
Maximized log-likelihood		-7509.60	-7432.74	-7407.98

Table 7: Cancer v. Non-Cancer Policies

Parameter	Variable	(1) Prevention Sample	(2) Treatment Sample	(3) Pooled Sample
β	-(Yearly Cost/10,000)	5.6529 (9.67)***	4.9757 (8.12)***	5.6253 (10.46)***
δ_1	Log(Illness Reductions)	0.0404 (5.77)***	0.0357 (5.03)***	0.04067 (7.13)***
δ_{2p}	Log(Death Reductions)	0.1391 (11.11)***	--	0.1363 (11.27)***
δ_{2t}	Log(Death Reductions)	---	0.0691 (7.19)***	0.0808 (6.45)***
δ_3	Log(Duration)	-0.1707 (-7.46)***	-0.1201 (-5.15)***	-0.1557 (-7.86)***
θ	Policy Dummy	-0.3035 (-4.13)***	-0.2377 (-3.05)***	-0.3325 (-5.01)***
	… · **1**(Treatment)	--	--	0.0984 (0.98)
	… · MajorCancer	0.1748 (4.57)***	-0.0652 (-1.75)*	0.1805 (4.73)***
	…· MajorCancer·(Treatment)	--	--	-0.2619 (-4.50)***
$ln(\kappa)$	Heteroscedasticity Parameter	0	--	0
	… · **1**(Treatment)	--	0	0.1207 (0.97)
Maximized log-likelihood		-8024.78	-7537.96	-15563.38
Sample size (choices)		7556	7033	14589
Maximized log-likelihood overall		-15562.74		-15563.38
Total sample size (choices)		7556	7033	14589

LR-test of restrictions in pooled model: $\chi^2_{0.05}(4) \approx 9.48$, $\chi^2 = 1.28$, fail to reject restricted model.

Table 8: Heterogeneity by Disease Type

Parameter	Variable	(1) Prevention	(2) Treatment
β	-(Yearly Cost/10,000)	5.6739	5.1971
		(9.62)***	(8.42)***
δ_1	Log(Illness Reductions)	0.0398	0.0350
		(5.65)***	(4.90)***
δ_2	Log(Death Reductions)	0.1439	0.0705
		(11.41)***	(7.28)***
δ_3	Log(Duration)	-0.1781	-0.1180
		(-7.71)***	(-5.02)***
θ	Policy Dummy	-0.1508	0.0388
		(-1.59)	(0.40)
	… · Heart Disease	--	--
	… · Heart Attack	-1.59	-0.0093
		(-1.16)	(-0.11)
	… · Cancer (General)	0.2582	--
		(2.91)***	
	… · Colon/Bladder Cancer	0.0833	-0.2862
		(0.92)	(-3.26)***
	… · Leukemia	-0.4663	-0.4785
		(-4.88)***	(-5.16)***
	… · Leukemia in Children	0.2250	0.0197
		(2.57)**	(0.11)
	… · Stroke	-0.3380	-0.3764
		(-3.59)***	(-4.28)***
	… · Respiratory Disease	-0.0612	-0.1531
		(-0.67)	(-1.70)*
	… · Resp. Dis. in Children	--	-0.1774
			(-1.03)
	… ·Asthma	-0.5373	-0.3942
		(-5.55)***	(-4.03)***
	… ·Asthma in Children	0.1691	-0.3036
		(1.93)*	(-2.19)**
	… ·Lung Cancer	-0.0674	-0.5230
		(-0.74)	(-5.78)***
	… ·Traffic Injuries	-0.1877	--
		(-2.32)**	
	… ·Injuries	--	-0.3486
			(-3.73)***
	… ·Injuries to Children	--	0.2580
			(1.50)
	… ·Prostate Cancer	--	-0.4677
			(-5.20)***
	… ·Breast Cancer	--	-0.0578
			(-0.68)
	… ·Skin Cancer	--	-0.8526
			(-8.94)***
Maximized Log-likelihood		-7930.80	-7453.51
Sample Size (Choices)		7556	7033

Table 9: Heterogeneity by Other Policy Attributes

Parameter	Variable	(1) Prevention	(2) Treatment
β	-(Yearly Cost/10,000)	5.6777	5.1482
		(9.72)***	(8.35)***
δ_{1p}	Log(Illness Reductions)	0.0405	--
		(5.78)***	
δ_{1t}	Log(Illness Reductions)	--	0.0359
			(5.04)***
δ_2	Log(Death Reductions)	0.1390	0.0714
		(11.11)***	(7.38)***
δ_3	Log(Duration)	-0.1693	-0.11896
		(-7.40)***	(-5.07)***
θ	Policy Dummy[a]	-0.3491	-0.2135
		(-4.08)***	(-2.63)***
Cause of ailment			
	…· Air Pollution	0.0981	--
		(1.75)*	
	…· Water Contaminants	0.1181	--
		(1.80)*	
	…· Pesticides in Foods	0.2317	--
		(3.55)***	
Gender-specific illnesses; respondent gender			
	…· Breast Cancer	--	0.0865
			(0.92)
	…· Breast Cancer·**1**(Female)	--	0.3655
			(2.94)***
	…· Prostate Cancer	--	0.1620
			(1.72)*
	…· Prostate Cancer·**1**(Female)	--	-0.5533
			(-4.03)***
"Affected group" choices			
	…· Percent Children	--	0.1220
			(1.29)
	…· Percent Children·**1**(Age65+)	--	-0.2553
			(-1.50)
	…· Percent Children·**1**(Kids)	--	0.6027
			(4.24)***
	…· Percent Seniors	--	-0.2295
			(-4.50)***
	…· Percent Seniors·**1**(Age65+)	--	0.0042
			(0.05)
Maximized Log-likelihood		-8028.5356	-7475.99
Sample Size (Choices)		7556	7033

[a] Prevention: omitted category= traffic accidents;
 Treatment: omitted category=all other illnesses or injuries

Session II: Issues With Morbidity Valuation

Altruism and Environmental Risks to Health of Parents and Children by Mark Dickie and Shelby Gerking

Is An Ounce of Prevention Worth a Pound of Cure? By Ryan Bosworth, Trudy Ann Cameron, and J.R. DeShazo

Mark Dickie and Shelby Gerking's paper, Altruism and Environmental Risks, is very interesting and well written. It is thorough and was a pleasure to read. The authors test a model of altruistic family behavior using a sun screen that will protect against both the risk of getting skin cancer, as well as dying from the cancer conditional on a positive diagnosis. They employ a stated preference survey using adults in Mississippi. Adults are asked their perceived risks of contracting and dying from skin cancer for both themselves and their child, and then they are asked the willingness to pay (WTP) for a sun screen that will reduce these risks by 10 percent or 50 percent, which are randomly assigned. The results are used to test the existence of altruism in the family.

Altruism is an important concept to consider in economic analysis. Primarily our concerns rest with the impact of altruism on valuation. In its simplest form, if parents, or any individual for that matter, behave in an altruistic manner, then individual values for a risk reduction will be compromised to the extent that they incorporate more than just the individual's WTP. By summing individual values we would then risk double-counting or over-estimating the total value for a risk reduction.

The paper could be more informative in this regard by including some discussion of the different types of altruism. Paternalistic altruism exists when an individual has concern for another's welfare, but is not necessarily concerned about the costs imposed on that individual. In other words, the paternalistic altruist does not incorporate the other's utility function into their own decision-making. Non-paternalistic altruism exists when an individual cares about both the benefits and costs imposed on another. That is, the non-paternalistic altruist fully accounts for the other's utility when making decisions. It would be useful to have a discussion of the different types of altruism and how they relate to this study, as well as valuation results.

Some additional questions that arose when reading the paper that could have implications for the application of these results to policy include:
    How do the results change when there are multiple children in the household? Do parents adjust their WTP to account for the additional children?

How do you account for two-parent versus single-parent households? If each parent in a two-parent household is altruistic how does this affect the values for the child?

How do you account for other individuals who are altruistic towards children, such as grandparents?

Overall, this is a well-written and interesting paper that sheds light on an important issue for benefits analysis.

Ryan Bosworth, Trudy Cameron, and J.R. DeShazo also have an interesting and well-written paper, "Is An Ounce of Prevention Worth a Pound of Cure?" They investigate preference for treatment versus prevention policies over a wide variety of policy attributes, types of illnesses and accidents, and respondent characteristics. There is a substantial amount of information in a short paper. Their results show that individuals are willing to pay almost double for prevention of death than for treatment of an illness that can cause death. For example, WTP for prevention of a death is $245, whereas WTP to treat an illness that causes death is $138. This is not surprising. There is disutility associated with entering the diseased state and therefore individuals are willing to pay to avoid entering that state. The authors also find that people are willing to pay equivalent amounts to treat and prevent illnesses, at about $70 for both.

The largest contribution of this paper to policy is in terms of determining how a policy maker may allocate resources. These results suggest that people would rather prevent than treat outcomes. Again, this is not surprising and it would be useful for the paper to explore more of why this might be the case. My sense is that it is related to either the uncertainty associated with outcomes, or the stigma, or both.

In terms of uncertainty, people are WTP to avoid uncertain outcomes, particularly those that result in death. Individuals would rather prevent cancer, than be in the state of having cancer and facing the possibility of death and having to back out from that state.

These results are consistent with the approach we have found to be the case in the manufacturing sector. Twenty years ago the Pollution Abatement Costs and Expenditures (PACE) survey primarily addressed costs associated with treating pollution, say installing scrubbers on a stack to treat emissions, or filters at water discharge areas to treat water before release. Today, we are in a pollution prevention paradigm. The treatment options have been addressed and we now focus on preventing emissions before they are created. Much of the expenditures at manufacturing facilities that we see through the PACE results support this notion.

It is also possible that stigma is driving these results. People would rather not enter a disease or illness state that may have a stigma associated with it. Hence, they are willing to pay more to avoid the stigma of being a survivor. It would be useful to explore these ideas further in the paper.

Other questions that would be useful to address include:

What are the implied VSL or morbidity values that result from this study?

What is the impact of the complex question design on results?

Overall, the paper is interesting and provides a useful discussion of how individuals value treatment versus prevention programs.

No documents are available regarding Kevin Boyle's discussion comments.

Summary of the Q&A Discussion Following Session II

Perry Beider, (Congressional Budget Office)
Commenting on the presentation of the Bosworth/Cameron/DeShazo paper and referring specifically to the finding that "someone ideologically opposed to government intervention would support a certain program once there was enough personal direct benefit perceived from it," Mr. Beider asked if the researchers observed that it went the other way also. In other words, was it observed that people who were generally in favor of government intervention *did not* support policies if there was too little perceived personal benefit?

Trudy Cameron, (University of Oregon)
Dr. Cameron responded that "it is sort of treated symmetrically—if it works in one direction, then it works in the opposite direction also, just by the structure of the model."

J.R. DeShazo, (UCLA)
Dr. DeShazo continued the response, adding: "But the effect isn't quite as large—the ideology effect dominates. It's true that people strongly ideologically in favor of government intervention are responsive to the size of the private benefits, but much less so than at the other end of the continuum."
_____

Bryan Hubbell, (U.S. EPA, OAQPS)
Also addressing the Bosworth/Cameron/DeShazo paper, Dr. Hubbell commented on the finding presented toward the end of the paper that when people were asked whether they prefer policies that help seniors or not, all of them said "no," including the seniors. He commented that "we just throw out the term *seniors* as if that's a well-defined term," and he added that he is curious to know whether the researchers worked to uncover an age breakpoint for this phenomenon. He clarified by asking, "What age does a policy have to affect before people will say that they'd rather not have that policy?—Is it 50? 55? 60? 65?—and is there any kind of declining support ratio at that point?"

Trudy Cameron
Dr. Cameron responded that the issue raised is an item of discussion on the Wednesday agenda. She added, "For private preferences we have some very detailed and elaborate analysis of age effects that are much richer than the simple quadratic thing that tends to dominate most of the prior literature. In the public choices study, which was discussed today, the distinction among beneficiaries is just defined in three groups—seniors, adults, or children. It's left to the individuals to interpret whether they are a senior or not."

J.R. DeShazo
Picking up on the response, Dr. DeShazo added, "Actually, I think for the respondents we did define the age intervals—60 or 65 was the cutoff."

Trudy Cameron
Dr. Cameron clarified, "But that's the beneficiaries—we have very detailed information about the respondent's age, of course, so that can be much richer."

_____

Douglass Shaw, (Texas A&M University)
Addressing the authors of both papers, Dr. Shaw asked, "What's the welfare measure?—what is it *really*?" He said that in Dr. Cameron's journal paper it is a "pretty careful derivation of an option price, which is kind of what we *think* it should be."

Dr. Shaw also asked, "When you do the subjective risk estimates, are you going after just the baseline risks, or are you also getting the subjectives on the risk changes?—and in either one of the designs, did you look to see if things are adding up?" He expounded that particularly in the Dickie/Gerking study there should be an obvious implication when doing a conditional probability. Acknowledging that the authors said they can do compound probabilities, which Dr. Shaw classified as "a very unusual result in the literature," he asked whether they did anything simple also.

J.R. DeShazo
Seeking clarification of the question, Dr. DeShazo asked, "Do you mean data analysis-wise or with the respondents or . . . ?"

Shaw
Dr. Shaw stated, "On the latter one, you're sort of saying that the results support that you can do compound probabilities, so obviously there's a law of probability between a conditional and an unconditional probability, so did you kind of just ask them to do a little experiment in the survey where you could verify that in fact they got that?"

DeShazo
Dr. DeShazo responded, "When I said you could do compound probabilities, I'm not sure I meant that if you asked them the unconditional probability and then did the multiplication on their conditional and their unconditional morbidity risk, would you actually get exactly the same number. I think they were making tradeoffs between those two risks that were consistent with the model, and I think they could distinguish between the two risks and not be confused between them. But we didn't ask them "what do you think the unconditional mortality risk is?" which you could use then to test whether they were really doing the math right. So, I don't know that."

Dr. DeShazo continued, "The perceived risk is all baseline; we asked them "what do you think the risk is?" The risk changes are exogenously assigned in the experimental design—they just come packaged in the sunscreen." He added that for the welfare measure they were looking at anti-willingness to pay for risk changes that would occur later in life.

Unidentified Participant

The response to Dr. Shaw was clarified by explaining that the probabilities were presented one at a time. For an example, "first the respondents were asked about the probability of getting skin cancer, and after they wrestled their way through that question, then they were asked about the chance they would die, given that they had it. So, we don't really have any results that say people can juggle two probabilities at the same time—and I'll bet they can't do it, just as you alluded."

_____

New Questioner
"How do you go about measuring violations of rationality when people are answering your survey questions?—is it a violation of transitivity assumption? If you do those measures, what do you do with the results? Do you throw out people who are clearly violating rationality?

J.R. DeShazo
Commenting that it was a very good question, Dr. DeShazo replied, "You can and we have looked at violations of rationality. We've also looked at how much attention people spend absorbing the information that we've given them, and we have altered our sample based on some minimum level of attentiveness that we felt they needed." He went on to say that there is always the sticky issue of how much information is enough and how much is too much. He added that he is "deeply concerned about the declining cognitive efficiency of individuals when they're given too much information." Saying that "we are all always given too much information—and we sort through it," he identified one of the tasks for researchers presenting information to individuals is to ask, "Have we left out something that is important?" Dr. DeShazo said he believes that if you give individuals enough familiarity with the attributes that make up a program, they'll decide for themselves which attributes are most and least important, and the proof (or disproof) of that will show up in your statistical analysis.

Dr. DeShazo added that in addition to looking at time on task his research team asks respondents, "How difficult was that choice?" In closing, he said that "in the context of evaluating their risk judgments, we can actually look at whether or not they make consistent decisions—we give them quizzes, basically, in the private version of the survey."

Trudy Cameron
Continuing the response to the questioner, Dr. Cameron stated, "J.R. mentioned this notion of how much attention people give to different aspects of a particular survey design. J.R. with Herman Fermo has some pretty rigorous work that came out in 2002 looking at how the structure of the randomized design affects the amount of noise, the choice inconsistencies that people make." She added that work that they're doing now, in conjunction with another student at Oregon University, Dan Burkhart, "has to do with an

actual sort of optimal allocation of attention problem." Dr. Cameron asserted that this deepens the model a bit by "acknowledging that what you think you're estimating as a marginal utility in a choice model as a consequence of standard multiple-choice specification is actually the product of some fractional attention, which may be very small or very large, times the true underlying marginal utility that you would be estimating if they were paying full attention—and that's producing some very interesting results. That's just some fundamental broader research that will have some bearing on these data as well."

_____

Lauraine Chestnut, (Stratus Consulting, Inc.)
Saying that this might be getting back to Douglass Shaw's question about welfare measure, Ms. Chestnut addressed this question to Drs. DeShazo and Cameron: "When you're asking questions about public policies that affect the person and everybody else in the community, how do you interpret that relative to the private valuation numbers that we tend to want for benefit/cost analysis. So, the example of the responses for seniors— and we've seen this in some other studies that ask these questions about public policy— what does that mean for valuation purposes?"

Trudy Cameron
Dr. Cameron responded, "Going back to respond to Douglass's question, which I didn't get a chance to: In the private choices survey, the model is highly structural and has to do with discounted expected utility maximization getting to an option price. But, for individual choices with respect to their own budgets and their own preferences for stuff that happens to them, it's a little easier to do that. This may account for why we haven't directly addressed much of the public choice stuff before—it's harder to come up with a solid, theoretical model about how people should think about these public goods. So, by its nature the public choices study with its two different surveys is very much more exploratory. Perhaps the term *descriptive* would be better—we're sort of identifying the stylized facts that need to be addressed in any further theorizing rather than starting with a rigid model. Bosworth has a more structural specification with respect to discounting— that's stuff he's working on now and finishing up for his third essay—but we figured we'd start with just the high points of the actual description of people's choices."

J.R. DeShazo
Acknowledging that he is relatively young to this field, having been actively involved in VSL literature only for 3 or 4 years, Dr. DeShazo said he finds it "a hard sell that we should be using these private good estimates for public policies because preferences over aspects of the public policy are *so different*. We could try to explain two different things—their actual support, their behavior—and it seems to me that if you're interested in their actual behavior with respect to these policies, you have to give them these attributes of public policies that don't hold or don't exist for private programs. Also, that behavior presumably reveals something about their perceived welfare from the public policy. I think the challenge is really on those that want to use the private estimates,

because, to me, they seem like very different utility functions with very different arguments in them."

_____

Reed Johnson, (RTI)
Mr. Johnson said that he had actually "looked forward to a lively discussion on the IOM Report, but I guess neither of the discussants were asked to comment on that. I'm afraid that we have on our hands another NOAA Blue Ribbon Panel Report that's going to be cited for the next 15 or 20 years, long after the evidence base that was used to make the recommendations has become obsolete. I'm a little concerned that the panel was constituted in a way that sort of biased it in favor of conventional ways of thinking about health utility that don't really line up very well with the way most people in this room think about utility—and I'd like to thank Alan Krupnick for his valiant efforts to try to keep the process a little bit more honest in that respect."

He continued, "There are a couple of aspects of the recommendations that I find troubling, in addition to Nathalie's points. For example, the quality recommendation is that the quality should be elicited for a general population sample. I work a lot with patient surveys and patient preferences and with some general population surveys. For many of the particular outcomes of interest, it is difficult for people who have never experienced that outcome to give meaningful values. . . . The general result is that patients experience much less of a utility loss than the general population assumes that they experience, partly because of adaptation and partly because people just imagine that something is going to be a lot worse than the experience actually turns out to be."

Mr. Johnson added that the report includes a recommendation for more research, but he said he thinks "the recommendation on gathering more data is stronger than the recommendation for improving methods," and he said he would have liked to have seen a much stronger advocacy for providing "measures of health utility that are both theoretically correct and empirically robust."

In conclusion, Mr. Johnson stated that he feels the publication of the report is an opportunity for groups like this to become engaged in trying to understand not only what obligation EPA is going to have in terms of doing their analysis but also what we can do to help encourage "more nuance of interpretation and more flexibility in use of methods."

Someone
"I thank Reed for that compliment. There were a lot of people on the committee that worked hard to do what we did. What we were trying to do is to create separation between measures of utility that we use in this literature, the economic valuation literature, and the measures of quality-adjusted life years and so on that are used in this other literature. . . . Hopefully that will serve the policy process and also serve our profession."

_____

END OF SESSION II Q&A

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

PROCEEDINGS OF SESSION III: PANEL DISCUSSION ON THE USE OF THE INTERNET IN VALUATION SURVEYS

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

Table of Contents

Session III: Panel Discussion on the Use of the Internet in Valuation Surveys

Session Moderator: William Wheeler, U.S. EPA, National Center for Environmental Research

Panel Participants:

- Nathalie Simon, U.S. EPA, National Center for Environmental Economics
- J.R. DeShazo, University of California–Los Angeles
- Shelby Gerking, University of Central Florida
- Alan Krupnick, Resources for the Future
- Jon Krosnick, Stanford University
- Brian Harris-Kojetin, Office of Management and Budget

Questions and Discussion

Washington, DC
April 10-12, 2006

Session III: Panel Discussion on the Use of the Internet in Valuation Surveys

Nathalie Simon, U.S. EPA, National Center for Environmental Economics

Will has asked me to sort of set things up for the panel discussion, so I'll talk through it a little bit and then present the charge questions. The way I see it, there are three kinds of internet surveys. There are those in which you recruit individuals into the sample using standard random probability sampling—and then you ask people to actually complete the survey over the internet, using a link that you provide. Then there are internet surveys using standing panels, and there are two kinds of standing panels: those in which individuals self-select into the panel and those in which the panel is created using sampling techniques.

It seems to me that there are benefits to all those types of web-based surveys. Generally speaking, once the survey is administered, you tend to have quicker turnaround on the results. In addition, you often have lower costs with these types of surveys and lower respondent burden. You can have greater accuracy as well—there's no interviewer bias or data-entry mistakes to worry about. Generally, individuals are entering the data the way they want to, and then they submit the results to you.

There is also greater flexibility in how the information is presented. You can have complicated skip patterns programmed directly into the survey instrument—you can have extensive use of graphics and color, which would be expensive or difficult to do using other modes. You can also have more interactive questions and can basically tailor the survey to individuals as they're going through it. Especially in the case of the standing panels you have the availability of some unique information that has been collected prior to the survey being administered.

You can also get information on time for question, and you can have extensive variable-tracking information if you need it. In some cases, you also have the possibility of using a voice-over, which can be very helpful in getting people to understand the questions that are being asked and to take the time to listen to the questions as well as reading them.

Of course there are a number of problems associated with web surveys as well. With the panel-based surveys you can often have low response rates. In fact, those of you who are users of Knowledge Networks, if you start looking at the response rate from the time that people are initially contacted to join the panel, the response rate is rather low. Given this low response rate, non-response bias then becomes an issue. In other venues and at other conferences, I've also heard a concern expressed that panels run the risk of creating

"expert survey takers"—I believe Reed [F. Reed Johnson] referred to them as "trained seals" at one point. That is a concern, as well. If you're going to the same individual repeatedly with surveys, do you create this expert survey taker?

There are other issues as well, especially with those surveys in which individuals are self-selecting into the panel. Really, these result in little more than convenience samples. It's often difficult to tell whether you're getting more than one individual from a household and things like that. You may have problems with actually downloading the information from the internet, and you may run into technology constraints as well.

Regardless of these problems, we are intrigued by the benefits associated with these web-based surveys, especially given that telephone surveys are becoming more and more difficult to do and mail surveys are also somewhat difficult—these modes pose problems when we're dealing with complicated questions that do involve complicated skip patterns or where we would like to use more complicated graphics.

As an agency though, to my knowledge, I think we've managed to use web-based surveys only on a very limited basis, and generally these have been surveys that have been done for research purposes. You'll hear more about one of these tomorrow, "Eliciting Risk Tradeoffs for Valuing Fatal Cancer Risks." This was work done by Chris Dockins, Melonie Sullivan, who is no longer with our office, and George Van Houtven, who will be presenting the paper tomorrow. Two other surveys looked at willingness to pay for water improvements, one designed by Kip Viscusi looking at eliciting willingness to pay estimates for improvements to fresh water, and then a survey looking at coastal water improvements.

But, again, these were surveys that were either couched in terms of pure research or testing of survey instruments—so, they were pilot surveys. We have yet, at least to my knowledge, to get approval for a web-based survey that would feed directly into a policy analysis for one of our rules and regulations.

Faced with the problems associated with web surveys, but trying to balance those against the benefits that we could exploit, I wonder whether there is a way to actually address some of these issues. One of the most important ones, perhaps, is this issue of non-response bias. It seems to me that non-response bias is perhaps more of an issue if you are dealing with low response rates, but it seems that you have a potential for non-response bias regardless of what the response rate is, unless of course you're dealing with a survey where you have 100 percent compliance. So, it seems to me that one question is "How do we address that?"—How do we go about trying to improve the representativeness of the sample or "How do we test for sample representativeness?"

Thinking about all these things, we had asked our panelists here to think about several questions as they were looking back over their own research and what they've done in this field. [referring to a slide] We have the questions for the panelists up on the monitor here. Basically, we've asked people to think about their experience with using the internet as a survey mode and to think about the choice of the survey mode in their

research and to consider the tradeoffs between convenience, cost, and bias and to comment on the key issues. Specifically, we've asked them to address these questions:

- What special issues must be considered when using the internet as a survey mode?
- Are there special circumstances where it makes sense to use the internet for stated-preference surveys?
- What conditions or circumstances does the internet provide and under what circumstances should the mode be avoided?
- What specific follow-up analysis or testing should be conducted when using the internet?

_____

Brian Harris-Kojetin, Office of Management and Budget

I'm going to take a very "10,000-foot-level perspective" here and then focus in a little bit and touch on some of the things that Nathalie mentioned, but I suspect that others in the panel will focus even more deeply into the specific issues. For those of you who have ever wondered, "Why does OMB review our surveys?" it is because we are required to by law. The Paperwork Reduction Act requires that any information collection that is sponsored or conducted by a federal agency go through this review, and the purpose of this is to improve the quality and practical utility of information that is gathered by federal agencies.

I want to make you aware of some new guidance that we recently issued in January of this year. It's entitled "Questions and Answers When Designing Surveys for Information Collection," and it covers a broad swath of things, but I'll just provide a brief overview of that. Just so you know, the intended audience for this guidance is very broad. It's intended to be used by people implementing the Paperwork Reduction Act—Chief Information Officers, Program Managers, survey folks, who are out there in the front lines doing this—and it covers a wide variety of different topics—everything from what do you have to do in terms of some basic process issues in terms of submitting the information collection requests or "OMB clearance packages," as they're more popularly known to what kinds of different issues you need to address and explain and justify and document here. I'm going to focus on a few of these that are related to some of the things you're interested in here in terms of internet surveys.

Specifically, we have questions on when should agencies consider designing a survey. Obviously, a survey is just one method in a social scientist's arsenal—it's appropriate some kinds of questions and issues and not so much for others. We have a whole section on sampling covering probability samples, coverage issues, sampling frames, . . .We bring in the issue that Nathalie raised, too, in terms of non-probability samples—that there are these internet panels out there that are essentially convenience samples. Even though these panels often boast of their numbers, which can reach over a million people, what you have is still an entirely self-selected convenience sample of "1.2 million" people who had nothing better to do than stumble across a web site and say, "Sure."

I'll also touch on several points about the mode of data collection . . . also the Government Paperwork Elimination Act (GPEA), which OMB is also in charge of implementing. This required agencies to allow citizens electronic options for reporting to the federal government. Although this law was not really written with surveys in mind, it can be applied that way—it's more for people who are applying for benefits or things like that or for businesses conducting transactions with the government. . . .

There's also a short section here in the guidance on stated preference methods. For those of you familiar with OMB Circular A-4, there's nothing really new here.

Generally, agencies are being encouraged to do a lot of electronic reporting, but there are some important stipulations in GPEA—agencies are encouraged to do this, "as practical." So, if you're doing a very small-scale survey or if you're doing anything with fewer than 5,000 respondents you don't even really need to consider it—or if it's otherwise just not practical or cost-effective.

Most federal agencies that use the internet for their surveys use it as one option in a multi-mode survey. Looking across agencies, it's being used more and more for establishment surveys or business surveys or surveys of organizations or institutions like hospitals or schools, and sometimes it is being used as the sole mode for those or for some specialized survey, such as a web site satisfaction survey. It's being used as the sole mode pretty much exclusively in cases where your target population or some sub-population of that has nearly universal web access. Not all businesses have that, but in certain industry sectors you can really count on that. There are several post-secondary school surveys that are now based exclusively on web collection.

One other thing that I want to point out is that web surveys are sometimes touted as convenience, and with some of the things I've seen from agencies it's not clear if they're thinking about the respondent or themselves. It can be very convenient for the *researcher* sometimes to use a web survey, but not so much for the respondent. The worst case scenario that I've seen in this regard is where the agency sends out a request to a respondent saying, "Please do our survey on the web. What you can do is download this, print out the PDF file, go fill it out, and then get on the web and put all the information in." This is *not* more convenient for the respondents. Why not just mail them the survey and let them mail it back? Why do they have to go through these extra steps?

In terms of cost, web surveys are often portrayed as being less costly. This is true under some circumstances, especially if it's a very simple survey that doesn't require much complex programming or testing. We have a lot of government surveys that very quickly become very complicated

In terms of bias and error reduction, we're looking for agencies to take these things into account and talk about how they are dealing with them in terms of why they've chosen the mode or modes that they're using.

In terms of choosing an internet survey as the preferred mode or one to be avoided, in reviewing packages we're basically looking for a good understanding and justification of how the agency is balancing some of the advantages and disadvantages—and Nathalie mentioned a number of these. Email reminders are certainly cheap and convenient for prompting respondents, especially if you can include that hyperlink in the email message that will take them right there. That has advantages over sending them a postcard with a long address that they have to type in. You do get faster data collection without delays in receiving the data. For instance, respondents can't tell you, "Oh, I mailed that last week,"—you know whether it's completed or not. Using visual aids and sometimes even multimedia is another advantage, as is the ability to build in some of these edit-and-consistency checks.

Disadvantages: Again, reflecting some of those issues mentioned earlier in terms of coverage and non-response and measurement error. What is the sampling frame?—Where did this sample come from?—Can you actually draw a random sample from your target population?—How well are you covering your target population? There are issues of response rates, in general—again, when they are used as a sole mode, web surveys tend to have lower response rates than other modes. That said, they are more often used as a mixed mode. Respondents have to be computer-literate and have access. There are hardware and software differences that can affect your presentation. Finally, there are some respondent concerns about confidentiality when giving information over a web site.

In terms of follow-up analysis or testing, I want to make two points. One is that pre-testing is just as important [as follow-up]—have the questions been tested to determine whether they are functioning as intended? When you're putting this on a web instrument, you need to do the usability testing as well. As far as follow-up analyses to assess a potential non-response bias—we all recognize that non-response rates *don't* indicate non-response error—they're an indicator for the *potential* for non-response bias. We expect that surveys collecting influential information should achieve high response rates, and agencies need to consider how what they are doing is going to give them data of the quality that they need. Our guidance, as many of you are probably aware, says that if an agency is getting a response rate of less than 80 percent, they need to plan a non-response bias analysis. There's a variety of ways of doing that—I think some of the people [fellow panelists] are going to talk about some specific examples here. Bob Groves and Mike Brick have taught a course now several times at a few federal agencies as well as to the general public—this is in the joint program in survey methodology—on Practical Tools for Non-Response Bias Analyses.

_____

Shelby Gerking, University of Central Florida

I want to report on some joint work that Mark Dickie and I have done using web-based surveys in valuation studies. We have some experience, at least, working with internet panels. We've worked with the CentERpanel at Tilberg University in the Netherlands and looked at willingness to pay for greater protection of seals there. That was back in the early part of this decade. Also using CentERpanel, we've looked at willingness to

pay for reduced risk of pancreatic cancer. Using Knowledge Networks, we've looked at blue-collar workers' willingness to pay for on-the-job-safety improvements. That was another study done earlier in this decade. More recently we've looked at parents' willingness to pay for reduced skin cancer risk to themselves and to young children ages 3 – 12. This last study is the one that I want to base my remarks on now, because it serves as a side-by-side comparison to the computer-assisted study that Mark Dickie reported on earlier.

The Knowledge Networks, or KN, Skin Cancer Survey in 2005 was transmitted to about 1200 panelists, and we, in one way or another, got down to 644 panelists with a child between the ages of 3 and 12 years who actually did complete the survey that was provided. The panelists completed the survey at home—there was about a 3-month period for Knowledge Networks to design, pre-test, and field the survey and for respondents to return a usable data set. It was a very smooth process, with very good, helpful people to work with.

The comparison is with the Hattiesburg Skin Cancer Survey from 2002 that Mark Dickie reported on. The survey was virtually identical, though not exactly identical, to the Knowledge Networks survey. It consisted of a sample of 612 parents with at least one child between the ages of 3 and 12 years. As Mark indicated, that survey was obtained by random-digit dialing of Hattiesburg area residents, and it took about a hundred calls from the poor students there to generate one completed survey. There were lots of hang-ups and lots of reasons why people might say "No," but there was also an eligibility problem, of course, because people had to have at least one biological child living at home between the ages of 3 and 12 years—that accounts for a lot of the extra phone calls. Respondents came to the University of Southern Mississippi campus and took this survey in a computer lab there, so rather than just being able to off-load the survey to the good folks at Knowledge Networks, we needed a lot of students and oversight to make sure that we at least knew what was going on in this computer lab.

As to the cost, using Knowledge Networks cost us $82 per completed survey. This excludes the investigator time needed to develop the survey—in other words, the clock starts running when you hand the survey to Knowledge Networks. It includes all pre-test costs and all of Knowledge Networks costs and all university indirect costs. With the Hattiesburg survey, it cost about $123 per completed survey. Again, that excludes the cost of investigator time used for survey development. Although it's not exactly the same, I tried to make the comparison as much apples to apples as I could. Anyway, the oversight that you need with one of these computer-assisted surveys is significant, and I valued Mark's time and my time on that job at about 9 cents per hour. The Hattiesburg cost includes the $25 participation fees provided to those who came to the computer lab and took the survey, and it includes all pre-test expenses, programming costs, labor and telephone charges, and university indirect costs. So, the Hattiesburg survey came at about a 50 percent cost premium.

Data Quality:
As far as sample composition, the Hattiesburg sample was more highly educated than the KN sample, and this is what you would expect, given that random-digit dialing was used to recruit the survey. The sample was more highly educated than you would have expected, given the census data for the Hattiesburg area. The Knowledge Networks survey was more representative of the United States population, but I would call attention to the fact that we're not really sure who completed all the surveys. When we were debriefing pre-test participants, out of eight such persons that we spoke with (Knowledge Networks had arranged the calls and was on the line also), we found out that one of them was not the person who had completed the survey—it was that person's spouse, instead. How widespread this problem is I have no idea—I'm not trying to condemn the Knowledge Networks survey on the basis of one observation.

The average survey completion time for the Hattiesburg survey was 26 minutes, and we had projected a completion time of 25 – 30 minutes, based on our own experience taking the survey and the time it took pre-test respondents. Twenty-three percent completed the survey in 20 minutes or less—you also want to know how many people just ripped right through it and probably didn't pay too much attention to what they were doing. In the KN survey, it took 1178 minutes for those respondents to complete the survey. One interpretation is that these people obviously work much more carefully than they do in Hattiesburg, but there are other interpretations as well that could be offered. One is that if you're at home and you're doing this on the internet, you're free to look at the survey. That's when the clock starts running, and then you say, "Yes, I see what this is—it looks very interesting—I think I'll do it in three days." That's possible. Another possibility is that you look at the survey and begin to do it but you decide to come back later to finish it. Then when you return, you have to pick up where you left off and reconstruct your train of thought. Seventeen percent of the surveys returned were "resumed interviews"— this is how Knowledge Networks refers to a survey that exceeds 100 minutes. Actually, I would classify a resumed interview as any that took from 30 minutes on, but this is how Knowledge Networks furnishes the data. Thirty-nine percent of KN respondents completed the survey in 20 minutes or less.

Another issue is the level of respondent engagement—the question distractions and interruptions come in. Looking at the KN survey, you begin to look at that average completion time, and you begin to think a lot about distractions and interruptions. Imagine someone trying to complete the survey and the cat is climbing up the drapes, the dog is barking, and the kids are playing with matches, someone's at the door, the telephone's ringing—all these things could be happening at once, who knows? Or, none of those things could be happening and someone just decided on their own that they would rather complete the survey later. Again, who knows? Anyway, the possibility of distractions and interruptions is certainly there.

With the Hattiesburg survey, where the respondents were completing the survey in the university computer lab, about the only possible distraction would be someone teaching calculus across the hall and a respondent might decide that they would rather go learn about the quotient rule. I don't think this happened, though.

A number of people in the KN sample had taken a lot of surveys—presumably they were experienced—that could be good, it could be bad—Reed referred to this sort of thing as the trained seal effect. Who knows? With the Hattiesburg study, it was a fresh sample—they hadn't participated in any previous surveys, at least none that we had done. There was also more item non-response in the KN survey than in the Hattiesburg survey. In the computer-assisted survey we had *practically no* item non-response, whereas in the KN survey there was *a lot*.

Did changes in features of the hypothetical sun lotion that Mark described alter willingness to purchase it in a predictable way? Well, with a change of price, yes. As the price went up, willingness to buy the stuff went down. How about extent of risk reduction? In the Hattiesburg survey, in a between-respondent comparison, we got higher willingness to pay for larger risk reduction, so there's an external scope test there. With the KN survey, again in a between-respondent comparison, we got *significantly lower* willingness to pay for larger risk reductions. What are possible explanations for the difference in outcome? I mentioned the greater education level of the parents in Hattiesburg. Maybe better-educated people are just in a better position to do these surveys than less-educated people. We did a variety of tests to try to detect whether education level had any bearing on the outcome of the extent of risk questions that we asked, and the answer was "no." It was just that the KN respondents, in general, were poorer at this than the Hattiesburg sample.

As far as resumed interviews in the KN sample, if you just took out all the people who took 100 minutes or more to complete the survey, would the basic results change? The answer is "no." Was there a greater level of engagement on the part of the respondents in the Hattiesburg survey? Maybe—I don't know—but it is a concern. One thing I wish we could have generated was some within-respondent evidence as to how people respond to changes in risk.

_____

Alan Krupnick, Resources for the Future

Wow—those are quite problematic responses to that survey of Knowledge Networks, and I don't want this panel to become a referendum or a judgment of Knowledge Networks, but it's probably worth saying why we mention Knowledge Networks so much. There may be people here who don't understand that. The reason that Knowledge Networks is so attractive is because they made an attempt through random-digit dialing to convert people to their panel who were not internet users. They gave them this special technology, webTV. You don't need a computer to take these surveys when you have this technology, so it deals with the problem of non-internet-users.

We (Maureen [Cropper], Nathalie [Simon], Anna [Alberini], and myself) did a national U.S. mortality-based survey in the year 2000 or so for our mortality risk valuation work, which has been reported in a couple of different journals. I wanted to talk a little bit about our experiences, particularly in regard to some of the responses I have after

listening to Shelby's presentation. Then I want to give a little advertisement for what's going to happen at Resources for the Future in October.

So, we've had experience with both Knowledge Networks and Ipsos Reed, which is a Canadian firm that does probability-based internet sampling but doesn't have the webTV technology. First, going through the work on mortality risk valuation, we basically had exactly the same setup, although different locations, as Shelby. In our Canada sample, it was a random-digit-dialed sample of people in Hamilton, Ontario that came to a central location to take the survey on a computer. Then later we did a national sample using Knowledge Networks on webTV or the computer. We got extremely close results on both of those surveys. Many of you in the audience have seen our bar graphs—almost equal responsiveness to the bids, which were basically PPP-corrected, so they were equivalent bids across the two countries. We had significant external scope effects. We had very little item non-response. Maybe this can be explained partly by the fact that we were using the panel in its early days—by the time Shelby got to it, it was rather old.

The one benefit that we saw from Knowledge Networks that you can't get easily from these in-person, self-administered surveys at centralized locations is that you can pick up infirm or immobile people—if they're in your panel or however you get them. That's important to many health surveys, so we thought that was a benefit from our work although I can't prove it. We also looked at the timing issues—these people who take 100 minutes or more, and so on. As Shelby mentions, we didn't find any effects on timing.

So, let me go to our Adirondack survey. This was done by Knowledge Networks in New York, so our sample of people was panelists from New York state, where we estimated the willingness to pay for improvements in the Adirondacks, and it was set up with an external scope test framework. What we did here is we used two different modes—an RDD mail survey and a panel internet survey. We had Knowledge Networks do both of these for us. The survey for the two was as identical as we could make it, given the difference in mode.

So, we did a few things. The first is that we looked at the demographics comparing the two modes to each other and comparing them to the census. We did pretty well. There were some observable differences across various samples, which we corrected using weighted regression. Differences in observables really don't cause any major problems. Then we used a Heckman selection analysis on the panel internet survey using KN's panel data, so we know from the panel who was exposed to the survey and had an opportunity to take the survey but chose not to. We did the analysis with that group and with the group that did take them, and we did find some groups less likely to respond to our survey—women, minorities, and the lower-educated were less likely to respond—but we didn't' find any statistical effect of the unobservable component of response on willingness to pay. Of course, the limitation of this kind of analysis is that we did not look further back in the chain to compare our results to people who chose not to be on the panel. So, that's going all the way back to the beginning of the RDD effort, and we weren't able to do that.

Finally, we compared the frame mode, the RDD mail results for willingness to pay to the panel internet results for willingness to pay, and we found that they were quite similar—there was no statistical difference between those two. For what it's worth, that's what we found.

Finally, I just want to mention what we're going to be doing in October. We'll be hosting an OPEI-funded workshop on the general topic of sampling bias. It's called "Sampling Representativeness: Implications for Administering and Testing Stated-Preference Surveys." We're going to bring in experts—some of the people on the pane here—survey researchers, statisticians, cognitive psychologists, and government officials, including Brian [Harris-Kotejin] and others to help better define the problems and work toward a solution. Our motivation here is this linkage that OMB makes between low response rates and therefore unreliability of the surveys. Our view is that you could have an 80 percent response rate that doesn't guarantee representativeness, or you could have a 10 percent response rate that does. What we need to do is decide what our performance measures are going to be and then what protocols we need to follow—and I know OMB is interested in defining those kinds of protocols—to permit us to take advantage of internet technologies that are out there to get these surveys done at low cost, quickly, and flexibly to give all the advantages that Nathalie mentioned and not give that up on what may be a false goal of lowering non-response rates. What we want to lower is sampling bias, and that's a different thing.

_____

Jon Krosnick, Stanford University

I'm a professor of communication, political science, and psychology at Stanford University, and I'm delighted to have the opportunity to speak with you this afternoon. I make my life, among other things, focusing on survey methodology. Increasingly lately I've found myself obsessed with mode—doing mode studies for a variety of reasons and trying to answer the general question of: What impact does mode choice have on survey outcomes?

As some of you no doubt know, there are lots of different sources of error in surveys. One is coverage error. That is, if we're doing a telephone survey, we'll fail to reach households that have no telephone access at the moment that we call. There is non-response error. That is, people of particular types choose not to participate and therefore bias the sample composition. Interviewers make errors in reading questions and in hearing and recording answers. Respondents make errors in interpreting questions and in doing inadequate memory searches for relevant information—integrating or reporting, as well. When you put all of this together, if produces what we think of these days as "total survey error"—that's sort of the sum of all of these errors. In order to provide the most accurate measurements from a survey, we want to minimize all of these various sorts of error. My focus during my few moments today is on how mode can impact the sum total.

There are various ways to think about how mode choice does have impact. As I've said already, if you decide to do a telephone interview, you have coverage error—period. That doesn't mean your results will be different from the results you would get if you had overcome that coverage error, but it does mean that if you ask people a question like "Do you have working telephone service in your house?" you will not get the right answer because of the method you used to contact people.

But, there are some other cases in which mode differences are less predictable, less expected, and less anticipatable. Let me say from the start here, my discussion is going to focus on probability samples only. As you've heard already, there are internet survey firms offering, at fabulous prices, internet surveys provided from non-probability samples. We *have* done work on non-probability samples, and we find consistently that those samples are less accurate in the data that they produce, sometimes *dramatically* inaccurate. I personally don't take them seriously for the kinds of work that requires generalization to populations, so I'm not going to spend any time talking about that today. What I am going to talk about very briefly [referring to slide] are the four primary "contender" modes these days and the considerations or variables associated with these modes that can help differentiate between them. I'm not going to go into great detail, but we could think through how face-to-face interviews, versus telephone interviews, versus paper-and-pencil questionnaires could differ in the rapport and trust that the respondents feel they have in researchers, in the confidentiality they feel their responses can be assured, the modeling of commitment that a researcher or an interviewer might provide and become contagious with respondents, and so on. There are lots of these different factors and 10 minutes is not adequate time to go through this theoretical analysis.

What I do want to do, though, is very quickly skate you across a set of mode comparisons leading to the ones we care about most on the internet. First of all, comparing face-to-face with telephone interviews, you know that in the late 60's to early 70's when telephone penetration in households became essentially universal, the appeal of the many practicalities of the telephone attracted researchers to that mode, especially the reduced cost. The question that arises is: Was there any price paid by saving that money and moving to the telephone and not having to ship interviewers around the country, being able to supervise them closely, being able to complete surveys much more quickly, and so on. [Dr. Krosnick then showed a slide that listed "all the studies that had been done comparing face-to-face to telephone interviewing before we did our work, and showing all the design flaws that they suffer from that prevent you from being able to make any inferences, unfortunately, from them about the question we care about."]

So, we did a study that used three different national experiments—a data set collected in 1976, another one in 1982, and another one in 2000—conducting the same survey side-by-side, random-digit-dialed telephone nationally as well as face-to-face with area probability samples. I want to just show you, without going into great detail, that for the full samples [again, referring to slide] there was more reporting error in the telephone data than in the face-to-face data across the board. The data show that the real cost of moving to the phone is for the least-educated respondents—they get hit the hardest by the added cognitive burdens of a telephone conversation. In addition, the telephone

respondents complained more often about how long the interview was lasting, they expressed more dissatisfaction with the length of the interview, they said that the interview was "too long" more often, and, amazingly, their interviews were *shorter* than those of the face-to-face respondents. Is it surprising that people feel rushed on the phone?—maybe not.

Interviewers also rated the respondents on the phone as "less interested" in the interview process and "less cooperative" with the response process, and we found that the telephone respondents were more likely to distort answers in socially desirable directions than were the face-to-face respondents, who presumably developed a sense of rapport and trust with their interviewers more effectively. In addition, the telephone respondents said they were more uncomfortable discussing sensitive topics, and the interviewers rated the phone respondents as being "more suspicious" than the face-to-face respondents.

Okay, that was very, very quick, but you get the bottom line, which is that in this contest face-to-face wins.

What about a competition between telephone and paper-and-pencil, as we move closer to the internet case? In this case, this is a study that we did for NASA, funded by the FAA—a study of airline pilots who fly you and me around on commercial airplanes. This was using a survey project called the National Aviation Operations Monitoring System. A field experiment was involved—licensed pilots were interviewed and they were randomly assigned either to be interviewed by telephone or self-administered questionnaires, and they were asked factual questions. We built into the experiment a measure of the accuracy of answers, and what we found was that the telephone provided substantially more accurate responses than the paper-and-pencil questionnaires did. So, in this case when you take the interviewer out and leave respondents on their own, the quality goes down. In general, the respondents forgot events they should have reported more on paper than they did when they were walked through the questionnaire by an interviewer on the telephone.

The respondents answering the paper-and-pencil questionnaire actually realized that their answers were less accurate. When we asked them to rate how accurate the answers were as descriptions of their experiences, they reported significantly lower confidence in the accuracy of their answers. The real story here is this one: Whereas it took 27 minutes on average for the respondents to complete the interview by telephone, it took only 16 minutes for the paper-and-pencil respondents to complete that very same questionnaire. They rushed through the questionnaire; they overlooked events and by failing to report them, compromised the accuracy of the data they provided. As a result, the winner in this little "race" is the telephone.

Now we move, finally, to your favorite topic: telephone versus internet. So, paper-and-pencil and computer modes seem pretty similar—no interviewer involved, just answering questions on your own—maybe we should be worried about this competition, maybe we should be pessimistic. What do the data say? Well, we have two kinds of data [again, referring to slide]. One is a lab experiment, where we brought a group of respondents

into our lab and randomly assigned them either to complete a questionnaire on a computer by themselves in a cubicle or to complete the very same questionnaire over an intercom system, being interviewed orally by an interviewer down the hall. What we found is, depending on which measure of validity we looked at, . . . large majorities of comparisons showed statistically significantly higher validity for the computer than for the oral interview and *no* statistically significant differences suggesting the oral interview was superior to the computer. So, interestingly, we find here that the computer yields more-accurate reports than the oral administration. Furthermore, in the computer case, manipulating the order in which response choices were presented to people had no meaningful impact on those answers—54 percent versus 51 percent. However, on the intercom we found a very pronounced order effect, where we manipulated the order of choices and it produced a big difference in the answers people gave.

Lastly [again, referring to a slide], the pressures toward social desirability were more powerful on the telephone than on the computer. On the computer, White respondents were quite willing to say they were in favor of decreased government help for Black Americans, whereas being interviewed on the intercom the plurality of respondents said they supported *increased* help for Black Americans instead.

So, what are our conclusions? Well, face-to-face beats telephone. Computer beats telephone. Telephone beats paper-and-pencil. So, one possibility is that face-to-face produces better data quality than computer, which produces better data quality than telephone, which produces better data quality than paper-and-pencil. If this were true, it would sort of be the case that you get what you pay for—the more expensive the method, the higher quality the data. . . . We shouldn't over-generalize here, but I guess what I would say is I think there's a lot of promise in the data I've shown you for the potential of the internet mode to produce valid data. The question is: Can it be accomplished effectively?

_____

J.R. DeShazo, UCLA

Given all the discussion about the benefits, I don't think I'm going to cover the benefits. Let me briefly tell you what my experience has been in the context of four surveys and then talk about sample selection correction, because following up on Alan's point, I think what we do want to reduce is sample selection bias. We, entirely through the efforts of Trudy [Cameron], did go back to the random-digit-dial stage and evaluate sample selection bias for both opinions that were expressed and the propensity of being our final samples for the first three surveys that we did through Knowledge Networks.

[goes through a series of slide that describe the surveys they did]

We were very much concerned that our estimates of willingness to pay would not be representative of the U.S. population, and so Trudy began thinking about how to go about correcting for that. . . . One of the problems in random digit dialing is to figure out who chooses not to be recruited by Knowledge Networks, but the problem doesn't stop there.

Here's a summary of the process so you can get an idea of the magnitude of the problem: There's the initial random-digit-dialed contact, at which time individuals can select out of the sample if they're not recruited. They could be recruited by Knowledge Networks and not profile—that is to say, not enter their panel at time "t." Assuming they enter their profile at time "t," they may at time "t + 1" select out of the panel and not be active and thus not be available to us when we draw our sample. Then, of course, the final selection stage occurs if they are not drawn randomly or otherwise by Knowledge Networks as part of our estimated sample. What we wanted to do is explore differences and describe the systematic selection out of our estimated sample as a function of a set of individual characteristics. . . .

One of our surveys gathered data on public opinion with respect to whether the government ought to intervene in environmental health and safety programs. One concern of ours was "did the panel have a liberal bias?" and we thought we could get at this question by focusing on this question about the appropriateness of government intervention. A more fundamental question, given that we are interested in estimating demand and peoples' willingness to pay is: Does this selection process lead to a non-representative sample that is going to express a biased willingness to pay? The second approach goes about estimating marginal selection probabilities, conditional selection probabilities, and then allowing the marginal utility associated with the attributes of the programs we're interested in the peoples' willingness to pay for to depend on the propensity to respond to the survey.

Approximately half a million individuals were contacted by Knowledge Networks or one of their subcontractors. We placed a restriction on our sample—we wanted adults over 24 years of age . . .there were 1600 individuals that were recruited for the sample. The nice thing about the random-digit-dial information that we were able to obtain is that we could match it with census data. This was not easy and it took a huge amount of time. Basically, the way we did it is we used individuals' addresses and their telephone exchange and Trudy developed an algorithm to associate the probability that that individual in either that address or telephone exchange would be associated with a particular census tract. Then she very cleverly developed a set of 15 orthogonal factors plus using data on voting behavior—basically, these propensities to participate or to persist in the sample. This was extremely laborious, so much so that it justified a paper by itself (Cameron and Crawford). Let me say that there are three papers that are available on our attempts at sample selection correction.

These 15 orthogonal factors explain 88 percent of the variation [unintelligible words] characteristics across tracks, so this is a very robust selection model.

Given the limited time, let me just get to the conclusions. For the first analysis on the question of liberal bias, whether or not we were obtaining an average representation of peoples' opinions as to whether or not the government should intervene via environmental health and safety programs—we found that there was basically an insignificant point estimate of bias in the distribution of attitudes toward regulation. So,

there was no appreciable effect that resulted from selection on the response item of interest.

In the second analysis, we did find statistically significant but very, very, very tiny effects on the key parameters across respondents' propensities to persist in the panel, so much so that they were, in the context of our willingness to pay estimates, insignificant—and I'll stop there.

_____

END OF SESSION III

Summary of the Q&A Discussion Following Session III

Mary Evans (University of Tennessee)
"It's my understanding of these panels, such as Knowledge Networks and Harris Interactive, that if you submit a fairly small number of questions they may sort of piggyback your questions onto a larger survey. I'm wondering, first, if that may explain some of the differences in Shelby Gerking's experience and Alan Krupnick's experience with Knowledge Networks in particular. Secondly, I'm wondering if anyone is aware of any studies that look at the effect this kind of piggybacking has on results, whether there's a systematic bias."

1st responder (Gerking or Krupnick)
He responded, "The Knowledge Networks study that we did was not piggybacked on any other" and added that, in fact, it was sufficiently long that Knowledge Networks determined that a time constraint should be imposed on it—they wouldn't piggyback it with another one.

2nd responder (the other one)
He added "and that's the same with ours. Ours was about 30-32 minutes on average, as well, and there was no piggybacking, so that won't explain it."

3rd responder
This person clarified that there are two kinds of piggybacking. One is when your questions go first before other people's questions, in which case there's no impact so who cares? The other possibility is that your questions get added to the end of somebody else's, and this creates two issues. He explained, "One is that your questions are now appearing when respondents are more fatigued. Secondly, prior questions have been on particular topics and have activated thinking in particular directions. There's plenty of literature suggesting that fatigue and the content of prior questions can indeed influence answers to later questions, so there's every reason to believe that that's problematic." He continued on saying, "On the other hand, there's absolutely nothing unique to Knowledge Networks or Harris Interactive in piggybacking, because if you take Alan's survey or any survey that I've done, all of the questions at the end of the questionnaire are sort of piggybacked on all the questions at the beginning of the questionnaire. So, anything that comes late in the questionnaire could be influenced by what came earlier, just as in any other case." He concluded by saying that although it's not unreasonable to ask if there's impact of early questions on late questions, but it's not unique to those firms.

Here, the questioner made an unintelligible follow-on comment.

3rd responder
This person replied, "Absolutely," and he said he would repeat the comment so that everyone could hear it. He summarized the comment, saying "that it would make a difference on the results of willingness to pay for asthma if the prior questions were about cell phones versus whether they were about asthma medications." He went on to say that he doesn't think there's any doubt about that, "and it could very well be true that your

early questions in your questionnaire can influence the later ones regarding asthma, too. In particular, there's one very well documented danger: If you ask early questions on willingness to pay for cleaning up pollution in the ocean, people will feel as if they have less disposable money available by the time they get to the asthma questions. We know of that problem, and that will occur in any questionnaire as a result."

4th responder
"There is a related issue, as well, which is the expectations of respondents when they first begin to take the survey. If they're seasoned panelists, they may be used to taking surveys where the questions are similar to: Would you open an account with thus-and-so bank if the account had these features and we threw in a free pizza? That's one kind of question. Or, to go along with Jon Krosnick's presentation: Would you vote for President Bush if he stood for election today?—Yes, No, Don't know. When you follow such a question with one such as: Now, assume you're an asthmatic—would you pay for this or that type of medication to control these or those kinds of symptoms?—then you're just increasing the level of difficulty for these questions. If somebody was not expecting to see something that difficult, maybe that would be a flag."

Trudy Cameron, (University of Oregon)
Dr. Cameron said she just wanted to acknowledge "the remarkable cooperation" that she and Dr. DeShazo received from Knowledge Networks in doing the non-response study, "going all the way back to the original RDD contacts." She specifically acknowledged the hard work of Mike Dennis and Rick Lee as well as a consultant, Dale Culp. She added, "If I had been them, I would have been very much more nervous about the downside of this enterprise. All of us, collectively, heaved a sigh of relief when things turned out pretty well, . . . but we put them way out on a limb, and we're very grateful they did cooperate in providing that data." Adding that the exercise has been done as much "at arm's length" as possible, she closed by saying that she is "comfortable that what we're finding is the right stuff."

J.R. DeShazo (University of California, Los Angeles)
Dr. DeShazo added, "These are firms—and they'll respond if we tell them what we need and they have enough lead time and planning time. One of the challenges Knowledge Networks had was that they hadn't thought to keep track of all of their random-digit-dialed contacts. They had to go back and recover that and were uncertain as to whether or not they could. So, whether we're expressing professional standards for data quality or responding to OMB, I think that there's a market out there for data collection. If we communicate our needs clearly, we're large enough demanders of the product that they are going to be responsive."

Unidentified speaker
"We use Harris Interactive to get access to their chronic disease panel for surveying patients. I've never tried to do a general population survey with them—I've done a couple with Knowledge Networks. One of the marketing strategies that Harris uses, that I believe they have implemented subsequent to Jon's study six years ago, is a fairly sophisticated propensity weighting scheme, in which every other month they conduct a

random-digit-dial survey and an internet survey from their panel and then attempt to devise a weighting scheme to match not only the demographics but the responses to certain attitude questions, particularly attitude questions that screen well for people who take internet surveys. Jon, are you aware of this sheme? Does it make sense to you? Do you think it's fixing some problems? Knowledge Networks' argument is that we can match the demographics but it doesn't really necessarily match people who are going to join a panel and answer survey questions every week."

Jon Krosnick, (Stanford University)
Saying he was happy to comment on this, Dr. Krosnick responded, "The Harris Interactive propensity weighting scheme is proprietary—they will not describe how they do it to anybody—and they *did* have it in place at the time that we did our 2000 study, which I showed you. We were provided with the proprietary propensity weights, and when we analyzed the Harris data, both with the weights and without the weights, we found that it did not change the substantive results at all—it didn't change the means or the distributions of variables. What it *did* do was increase the standard errors of the estimates. The reason for this is because when we looked at the weights, there were some as large as 20 or 30 and some as small as 0.1 or 0.2. So, the weights are dramatic and they didn't have any real impact on the results that we looked at. As it turns out, Harris will not normally reveal the questions that they use in those parallel surveys to develop the weights, but they actually accidentally sent us the questions. So, having seen the questions, I can tell you that I'm not even slightly surprised that they don't do anything helpful."

Dr. Krosnick continued, "The more recent study we've done, which I haven't mentioned to you, is one in which we compared the same questionnaire administered by random-digit-dial telephone, Knowledge Networks, and six other firms that use volunteer samples, some of which do weighting by quotas on demographics and one of which provided proprietary propensity weights. We found the same thing—the propensity weights didn't change anything, and the volunteer samples were substantially less accurate. So, my results that I showed you earlier and these new results are not focused on demographics. The vast majority of our results comparing the reliability and validity have to do with substantive measures of attitudes, beliefs, behaviors, and so on. In cases where you can compare factual matters—like whether people have a driver's license or not, whether they have a passport or not—and other figures where there are *official* numbers to compare to, the probability samples from telephone and from Knowledge Networks were equivalently accurate and the volunteer samples were notably less accurate."

In clarifying how the panels process a shorter survey, Dr. Krosnick stated, "Their panelists are answering questions every week, so they'll add your question to a survey that's already going to go out anyway. How much does this cost them to add one more question?—nothing—get your $500—fabulous.

Another responder

"I just wanted to mention with respect to the cost figures that were presented before—we might have received a bulk discount. Our experience, just for the benefit of future negotiations, was that the total cost for Knowledge Networks was less than $45 an observation for a 30-minute survey."

Jon Krosnick
"Definitely a bulk discount."

Unidentified questioner
"Does that include university overhead?" When a responder replied, "No, it doesn't," the questioner said, "Okay, that's part of the difference."

Jon Krosnick
Dr. Krosnick added, "There's also a very subtle but interesting issue on overhead for those who care about this. The universities make a distinction between subcontracts and service purchases. If it's a subcontract, you only pay indirects on the first—let's say $20,000; if it's a service purchase, you're paying indirects on everything. You definitely want to negotiate with your university to make it a subcontract so you don't pay more indirects than necessary."

James Hammitt, (Harvard University)
Dr. Hammitt said he wanted to get some of the panelists' perspectives on a question related to the cost issue. He continued, "When I first got involved in internet surveying, it seemed to me that compared with phone surveys the fixed cost of setting it up might be high but the marginal cost per respondent would be very much lower because you don't need the live interviewer. With something like a Knowledge Networks panel, there's obviously a cost to maintain the panel and an opportunity cost to use it up. Is it right that the marginal cost per respondent will tend to be much lower with internet than with phones, for example? It seems to me that that would have implications for how we design surveys, because, as Jon has commented, there's a concern that if you ask people a lot of questions they get tired out and the responses toward the end may not be very good. However, if the marginal cost per respondent is low, we should just have very short surveys of a very large sample, whereas with phone surveys there's so much cost involved in getting somebody on the line who is willing to answer your questions that we tend to go for a longer interview with them."

Jon Krosnick
Dr. Krosnick replied, "I think that's definitely misleading. Basically, when you think about fixed costs of telephone interviewing—you have to hire a staff, you have to train the staff, you have to have supervisors, you have to have facilities and machines and all that—then once you get them in there, if they keep making more phone calls obviously making one extra phone call doesn't require all that much more staff time. Similarly, Knowledge Networks has to invest a bunch of money in recruiting a panel and then equipping the panel and paying them incentives and keeping them all going every week. My guess is that adding another respondent to the panel is actually considerably expensive—you have to make recruiting phone calls and get them signed up and send

them the equipment and all that—and you have only so many people in your internet panel. So, when you say that adding one extra respondent doesn't increase the cost very much, that's sort of true, but the whole fixed cost scheme is pretty burdensome, I think. You might say that you don't have to make a new phone call. Adding that marginal respondent on the internet case isn't that expensive if you weren't going to use them anyway that week, but it's not clear that Knowledge Networks doesn't want to use them anyway."

Kelly Maguire, (U.S. EPA)
Addressing Brian Harris-Kojetin, Dr. Maguire stated that he had mentioned that "many federal governments are moving toward using mixed modes," and she said she was wondering whether any of the panelists have experience with using mixed modes. She added that one of her concerns is that "when you start to use multiple modes within one research study, you introduce other biases that become more problematic than say the non-response bias that you're trying to correct in the first place."

Alan Krupnick (RFF)
Dr. Krupnick responded, "I mentioned in my remarks that we *did* use mixed mode—we used a mail survey and the Knowledge Networks internet survey." He acknowledged that the two surveys were "not exactly the same" due to the "issues you have to confront in switching these modes"—but they were pretty close. He added, "Maybe we were fortunate to have our willingness to pay estimates not be any different across these two modes. If they had been different, then we would have faced the issue of trying to explain why, but we didn't have to do that."

_____

 END OF SESSION III Q & A

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

DISCLAIMER

Table of Contents

Integrating Economic and Behavioral Questions into National Health Surveys

US EPA NCER/NCEE Workshop

11 April 2006

1

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

2

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

3

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?

4

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

5

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

6

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

7

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:

$$q \cdot T \tag{1}$$

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:

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

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:

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

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 ith 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

Parameter	Parameter Name	Eagle (n=193)		SSD (n=210)		IQ (n=208)		RC (n=196)		Combined (n=204)	
		ECOFIRST				HHFIRST				COMBINED	
		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

Description	Variable										
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 χ2=10.3, p=0.001

QALYweight by Endpoint, Kruskal-Wallis χ2=0.6, p=0.4
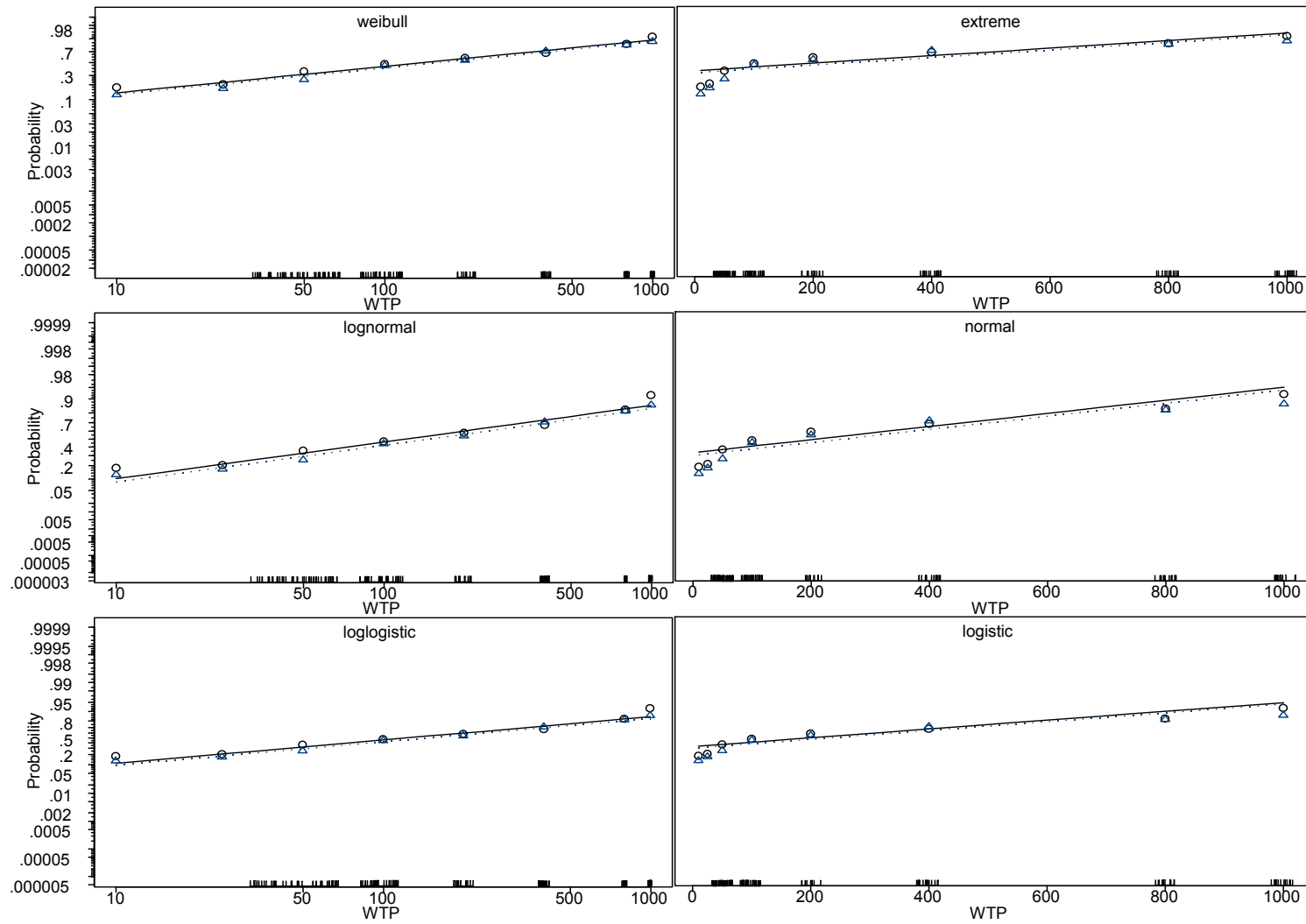
FIGURE 1: Probability Plots for the HHFirst Single Endpoint
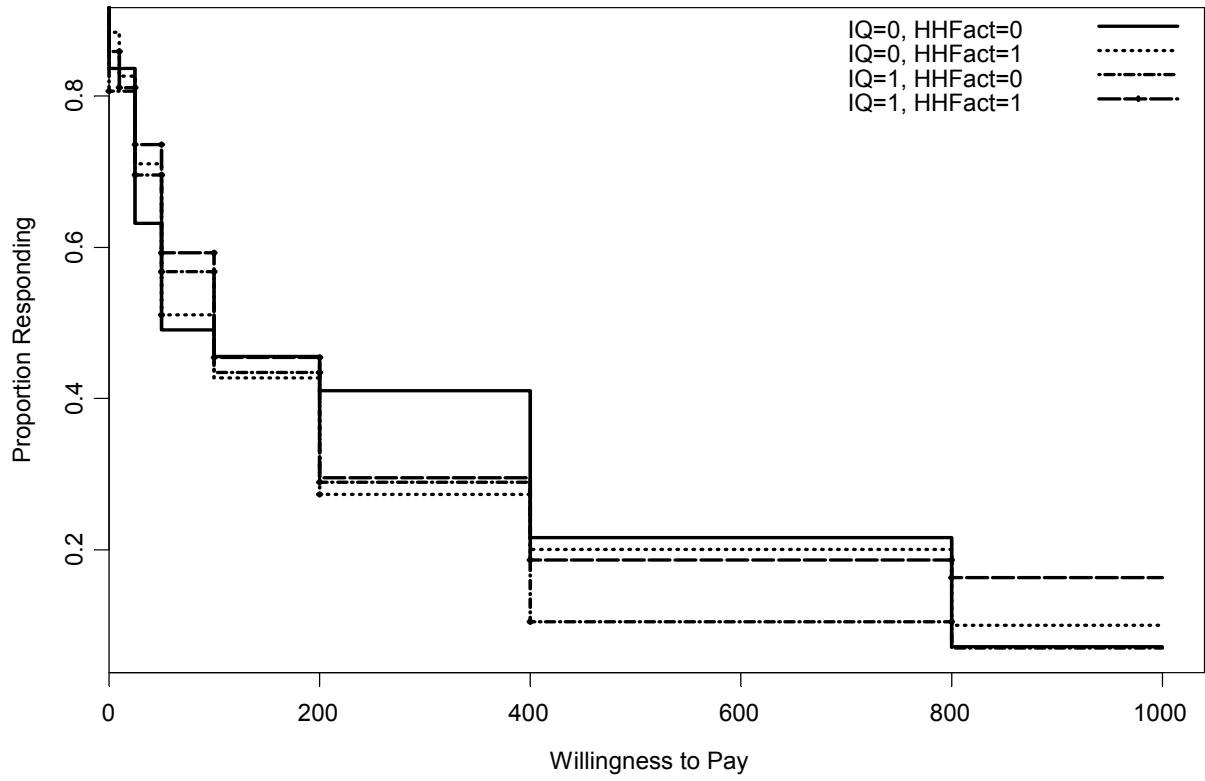
FIGURE 2: Willingness to Pay Across Risk Reductions for Human Health Endpoints

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,

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.

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

s.t.

$$\pmb{p'q} + w^a l^a + w^b l^b \leq y^a + y^b + y^H + (w^a + w^b)T$$

where, in the two adult household case, $i = (a, b, H)$ indexes individuals a, b and the household respectively, $\pmb{p}$ is a vector of prices, $\pmb{w}$ is a vector of wages, $\pmb{y}$ 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, $\pmb{q}^i$ are private goods, $\pmb{Q}$ is a household public good, $\pmb{q} = (\pmb{q}^a, \pmb{q}^b, \pmb{Q})'$ 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."

9

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.

14

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.

15

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 Management	Pooled Income	Separate	Assigned Management	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.

Decisionmaking	Husbands View (%)			Wives View (%)			Couple-Level Agreements (%)			
	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).

Rank	Air Pollution H	Air Pollution W	Climate Change H	Climate Change W	Radon H	Radon W	Lead Paint H	Lead Paint W	Smallpox H	Smallpox W	Smallpox Vaccine H	Smallpox Vaccine W	Anthrax H	Anthrax W	Influenza H	Influenza 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

Figure 2.

Dyadic Adjustment Scale Differences
Husband's Score minus Wife's Score

Figure 3.

Differences in the Number of Correctly Answered Lead Questions
Husband's Score minus Wife's Score
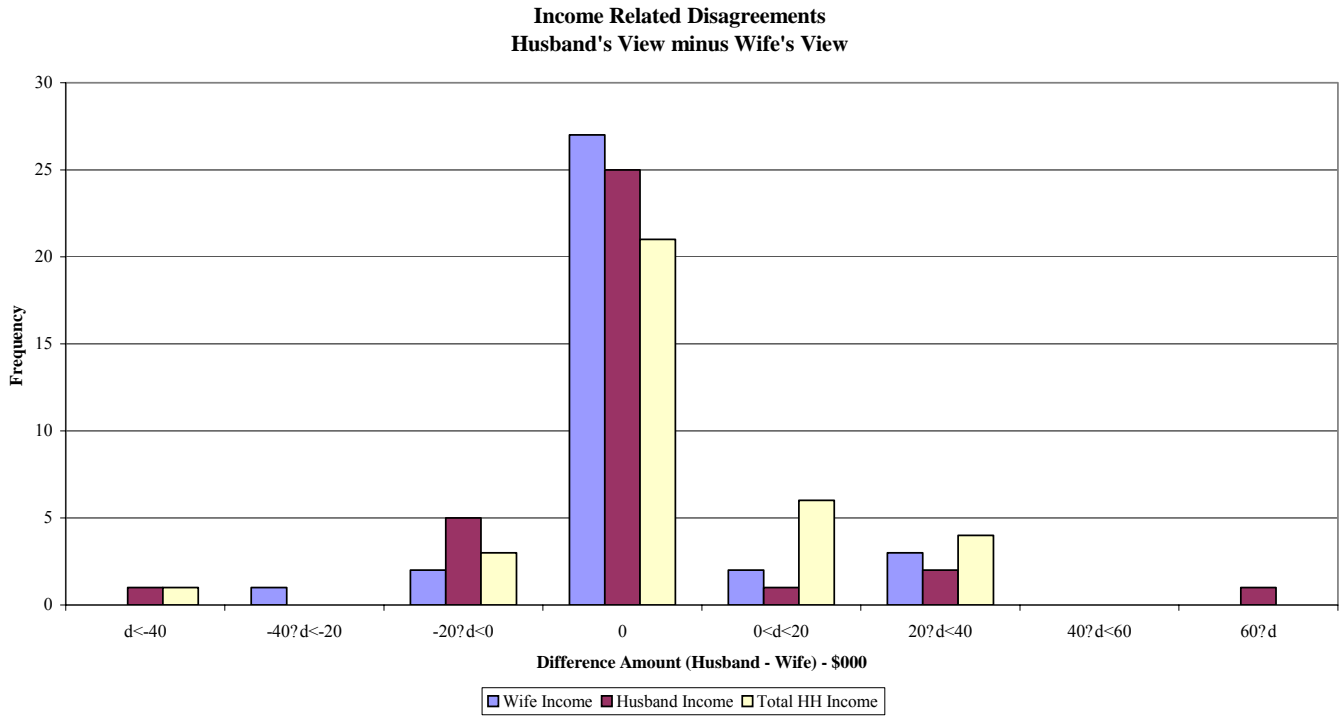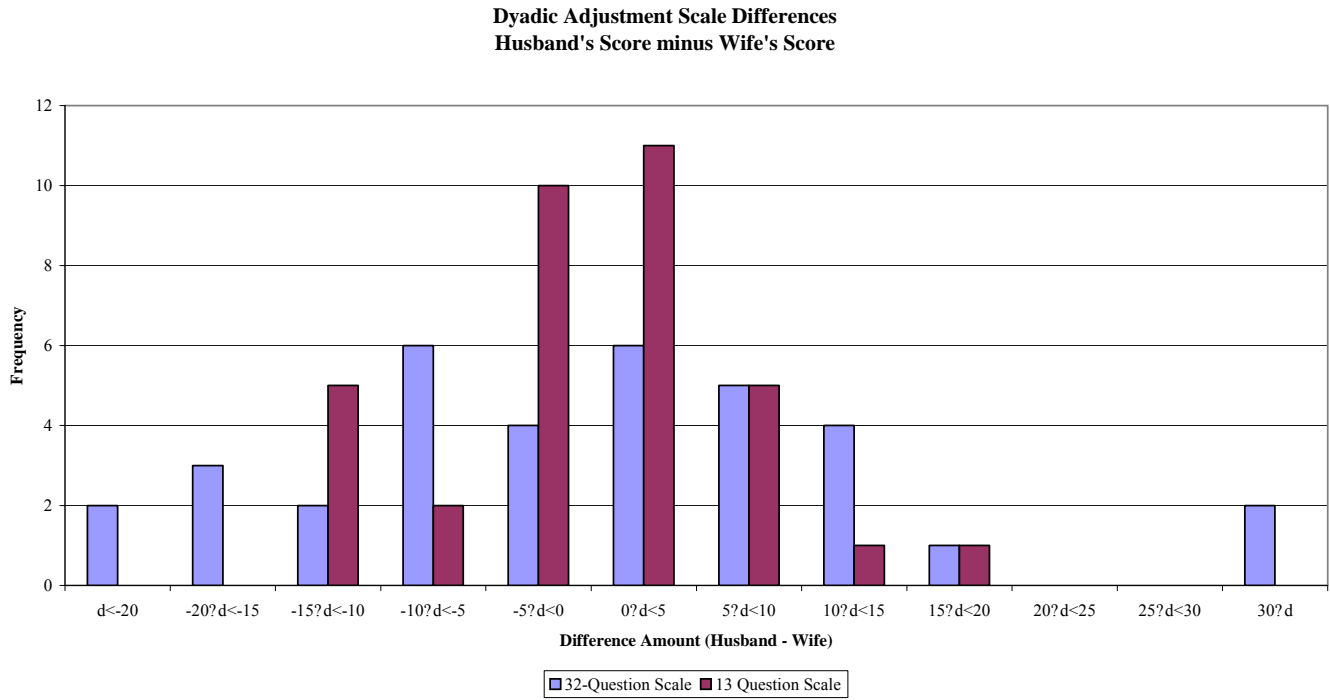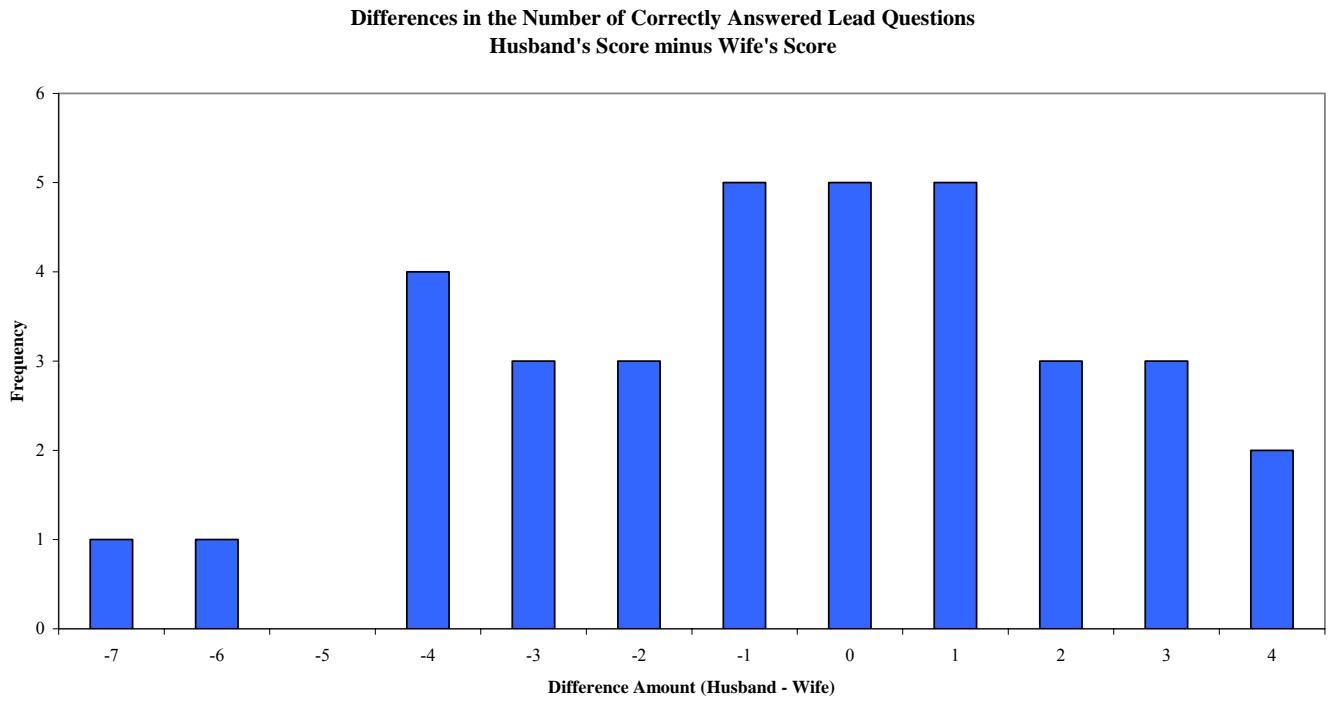
REFERENCES

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

2

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.

<u>Accounting for Latency.</u> 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.

<u>Magnitude of Cancer Premium.</u> 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

9

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

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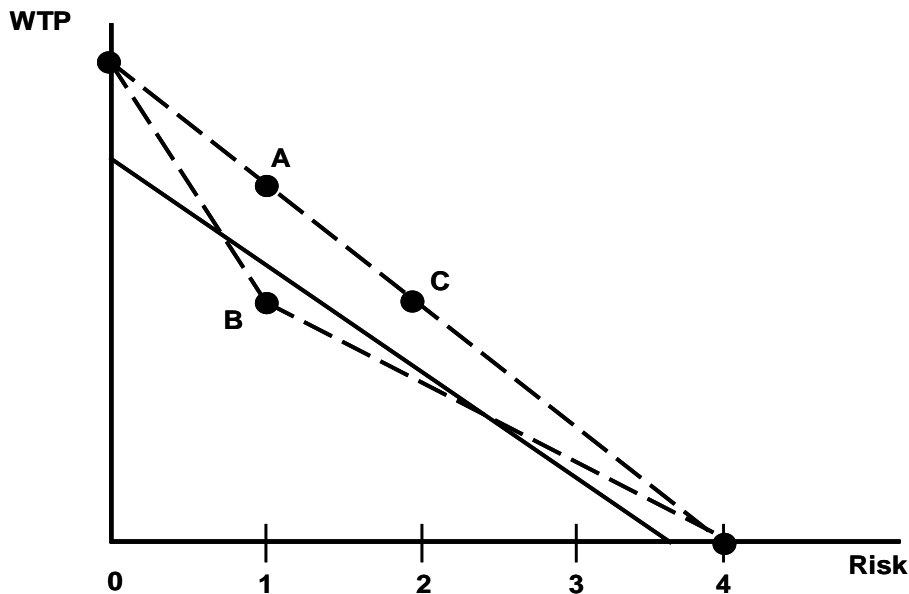
1 will be at some point B. Alternatively, if the risk levels 1-4 are very small probabilities, respondents may find it difficult to discriminate between absolute differences. They may

4

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

7

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

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

PROCEEDINGS OF SESSION V: EMPIRICAL ISSUES ASSOCIATED WITH MORTALITY RISK VALUATION

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

Table of Contents

Valuing Mortality Risk Reductions at EPA

Presentation at EPA Workshop
Morbidity and Mortality:
How Do We Value the Risk of Illness and Death?
April 11, 2006
Kelly Maguire
US EPA
National Center for Environmental Economics

1

Background

- Value of statistical life (VSL) estimate used in monetizing mortality risk reductions
- Same central estimate used since 1999
- Currently revising Economic Guidelines and revisiting VSL guidance
  - Literature grown considerably since default estimate was derived
  - EPA commissioned reports have raised issues with underlying literature
  - Recently published meta-analyses provide new means of combining estimates

2

EPA's current VSL Guidance

- EPA relies on benefits transfer:
  - Point estimate of $6.2 million ($1999) as an estimate of the value of statistical life
  - $6.9 million ($2004)

- Derived from 26 studies:
  - 5 stated preference studies
  - 21 hedonic wage studies
  - Studies date from 1974-1991
  - Values range from $0.7 million to $16.9 million ($1999); $0.8 million to $18.6 million ($2004)
- One value from each study was used to fit a Weibull distribution

3

Alternative Estimates

- $5.5 million ($1999) has been applied recently in air rules
- Central value from the range of values suggested in recent meta-analyses
- Distribution has a confidence interval from $1 to $10 million
  - $1 million is lower end of interquartile range from Mrozek and Taylor (2000)
  - $10 million is upper end of interquartile range from Viscusi and Aldy (2003)

4

Further Guidance

- EPA applies the same estimate to all populations
- EPA applies same estimate to all types of risk
- VSL is adjusted for timing
  - Discounted for risk reductions in future years
  - Inflated to account for growth in real income over time

5

Revisiting the VSL

- Number of new mortality risk valuation studies
- EPA funded three studies to examine various segments of the literature
  - Black, et al. (2002): HW literature
  - Alberini (2004): CV literature
  - Blomquist (2004): AB literature

6

Revisiting the VSL (cont.)

- New meta-analyses of mortality risk valuation literature
  - Mrozek and Taylor (2000): HW only
  - Viscusi and Aldy (2003): HW only
  - Kochi, et al. (forthcoming): SP and HW

7

Process for Revising VSL: Meta-analysis Panel

- SAB-EEAC expressed an interest in learning more about meta-analysis
- Panel of experts met in December 2005
- Goals
  - Discuss the issues and challenges in conducting meta-analysis for mortality risk literature
  - Prepare summary report

8

Process for Revising VSL: Consultation with SAB-EEAC

- Presentation of meta-analysis report by SGEs
- Other issues to raise
  - Population issues
  - Types of studies
  - Relevant measures
  - Covariates
- Guidance on revising VSL estimate

9

Some Options

- Continue to use current estimate
- Adopt results from existing meta-analysis
- Derive new estimate
  - Fit a distribution
  - Conduct a new meta-analysis
  - Other?

10

Eliciting Risk Tradeoffs for Valuing Fatal Cancer Risks

George Van Houtven*
RTI International

Melonie B. Sullivan
Institute for Family Centered Services, Inc.

Chris Dockins
US Environmental Protection Agency

May 2006**

Paper prepared for presentation at

U.S. EPA NCER/NCEE Workshop:
"Morbidity and Mortality: How Do We Value the Risk of Illness and Death"
Washington, DC
April 10-12 2006

Abstract:

Reductions in cancer risks are among the most important and tangible benefits resulting from a variety of environmental, food safety and other public health initiatives; however, relatively little is known about how individuals value reducing cancer risks compared to other types of risks. Most of the existing empirical research on the valuation of mortality risks has focused on accidental (occupational and/or automobile) fatalities. It is often argued, however, that differences between the characteristics of cancer risks and accidental risks may lead to significant differences in how they are valued. In particular, the time lag between exposure to carcinogens and its physical manifestation (i.e., the latency period), as well factors such as the fear, dread, pain and suffering may affect individuals preferences for avoiding cancer risks. To address this issue, we conducted a national survey of adults that elicits their relative preferences for avoiding automobile fatality and fatal cancer risks. We specifically examine how strongly individuals prefer avoiding one type of risk over the other, how this strength of preference is affected by the length of the morbidity and latency periods, and how preferences differ across different types of cancer. Our results indicate that individuals generally have a strong preference for avoiding fatal cancer risks relative to automobile fatality risks; however, as expected, this preference is inversely related to the length of the cancer latency period.

2

Introduction

Reductions in cancer risks are among the most important and tangible benefits resulting from a variety of environmental, food safety, and other public health initiatives. Nevertheless, little is known about how individuals value reducing cancer risks relative to other types of risks. Although a large empirical literature exists that has generated estimates of willingness to pay (WTP) to reduce mortality risk, this literature has focused almost exclusively on accidental (occupational and/or automobile) fatalities. It is often argued, however, that differences between the characteristics of cancer risks and accidental risks may lead to significant differences in how they are valued (as measured by WTP). Specifically, the time lag between exposure to the cancer risk and its physical manifestation (i.e., the latency period) may lower WTP for cancer risk reductions relative to accidental risk reductions. In contrast, the pain and suffering associated with the morbidity period that precedes a cancer death may increase WTP for reducing cancer risks. If a fatal cancer engenders more fear and dread than an accidental fatality, then WTP to reduce cancer risks may be higher.

Although a few studies have examined the empirical effects of risk characteristics on preferences for risk reductions, there has been little attempt to specifically and systematically test for how variation in latency and morbidity associated with cancer affects preferences.[1] To address this research gap, we have designed and implemented a national survey of adults that elicits their relative preferences for avoiding two types of potentially very different mortality risks—risk of automobile fatality and risk of contracting a fatal cancer.

The objective of this study is to use stated preference methods to assess individuals' tradeoffs between the two types of risks. In particular, we estimate how strongly individuals prefer avoiding one type of risk over the other, how this strength of preference is affected by the length of the morbidity and latency periods, and how preferences differ across different types of cancer. In addition to informing the debate on how individuals perceive different types of risks, and how these perceptions may affect preferences to reduce different risks, the results will indicate that additional research on valuing different types of fatal risks is warranted.

The analysis finds that individuals have a strong preference for avoiding cancer risks relative to automobile fatality risks of the same magnitude; however, as expected, this preference decreases as the cancer latency period increases. Individuals' preferences for avoiding future cancer risks are also, as expected, positively related to their chances of surviving until the age of onset of illness. The details of these results are discussed below.

---

[1] The only exception appears to be recent work by Trudy Cameron and J.R. DeShazo described in section 1, below.

3

1. Background

Although some anecdotal evidence exists that preferences for reducing fatal cancer risks may depend on characteristics of the risk, the U.S. Environmental Protection Agency's (EPA's) Science Advisory Board Environmental Economics Advisory Committee (SAB-EEAC) (EPA, 2000) concluded that "there is not sufficient theoretical and empirical basis for … accounting for these differences [in characteristics]." On the specific question of how cancer risk valuation differs from risks of accidents, the SAB-EEAC noted that "the value of reductions in cancer risks should include both the value of the reduced risk of death and the value of reduced risk of the morbidity, fear, and dread that precedes the death incident" but added that "existing studies provide little reliable information as to the magnitude of this premium." The purpose of this study is to begin to provide a basis for this empirical research by exploring specific aspects of cancer risk valuation in a limited sample setting.

The existing empirical literature on mortality valuation focuses largely on safety rather than on cancer-related mortality risks.[2] There are a few exceptions, but even these studies provide relatively little evidence on how specific risk characteristics may affect preferences for reducing fatal cancer risks. For example, Smith and Desvousges (1987), Hammitt (1990), and duVair and Loomis (1993) use contingent valuation (CV) to estimate individuals' WTP to reduce environmental/food safety risks of death; however, in none of these cases were deaths described or necessarily interpreted as cancer risks. In contrast, Carson and Mitchell (2000) used a CV survey, conducted in 1985, to elicit WTP to reduce carcinogenic risks from trihalomethanes in drinking water. The survey did not specify the type of cancer or related health outcomes, nor did it discuss latency effects; therefore, it is difficult to establish whether and how these factors affected respondents' stated preferences. The Carson and Mitchell results were based on a relatively small (n = 237) and very localized sample (Herrin, IL).

Two recent surveys from Asia have begun to examine cancer risks relative to other mortality risks. Hammitt and Liu (2004) employ CV in Taiwan to estimate WTP to reduce risks of cancer and non-cancer illness (liver disease). The design characterizes risks as either acute (beginning within a few months with death to follow 2 or 3 years later) or latent, where symptoms begin about 20 years in the future. Results suggest that WTP for cancer is about one-third larger than WTP to reduce the risk of a comparable chronic disease, but the estimate is not statistically significant at the 10% level. Individuals appear to discount for latency at about 1.5% per year, but the survey does not vary latency periods, morbidity times, or consider accidental fatalities.

---

[2] Some of these studies address risk dimensions associated with cancer, such as latency (see, for example, Krupnick, et al. 2005). There is also a body of literature on how public preferences for risk reduction programs vary by type of illness and other risk attributes (e.g., Subramanian and Cropper, 2000).

Tsuge, Kishimoto, and Takeuchi (2005) use a choice experiment to estimate marginal WTP for reduced mortality risks from cancer, heart disease, and accidents. Results from a Tokyo-area sample (n=400) show a small, but significant preference for reducing a generalized cancer risks relative to heart disease and accidents. This difference is not sensitive to risk magnitude. Results also suggest that respondents exhibited a strong preference for earlier risk reductions, with implicit discount rates estimated at approximately 20%. The survey does not vary morbidity periods or examine specific cancer types.

Cameron and DeShazo (2004) use a choice experiment to evaluate several aspects of health and risk valuation, including cancer, morbidity, and latency effects. Draft results suggest WTP generally diminishes over latency periods, with the specific effect contingent upon age, wealth, and the "illness profile," including length of morbidity and whether or not the effect is ultimately fatal.

Magat, Viscusi, and Huber (MVH) (1996)) used a computer-based survey to explore individuals' tradeoffs between automobile fatality and specific cancer risks. Using a mall intercept recruitment in Greensboro, NC, MVH administered the survey to 727 adults. The survey asked individuals to choose between two hypothetical residential locations that differed only in terms of the risks of automobile death and risk of lymph cancer. By varying the risks in the two locations, MVH estimated the lymph cancer "risk equivalents" for auto death—the risk ratio at which respondents were indifferent between the two locations. They found that for *terminal* lymph cancer the median respondent viewed the two risks as equivalent (risk equivalent = 1), and on average, nonfatal lymph cancer risks were "valued" at roughly two-thirds the rate of auto death risks. Although respondents were provided with information about the consequences of the disease, it is not clear which attributes of the disease or its risk primarily affected individuals' relative preferences for avoiding the two types of risks. In particular, the role of latency periods for cancer risks was not addressed in this study.

This study builds on the work by MVH, using a similar preference elicitation method. However, our survey is specifically designed to examine how individuals' risk equivalence rates between auto death and fatal cancer risks are affected by latency periods, morbidity outcomes, and types of cancer. This type of information is not available from existing research.

3. Conceptual Model

To model risk preferences we use an approach similar to MVH (1996). We assume that respondents make choices to maximize expected lifetime utility, E(U), which is defined in the following way:

$$E(U) = P_D U(D,Y) + P_C U(C,Y) + (1 - P_D - P_C)U(H,Y). \qquad (1)$$

According to this expression, lifetime utility is determined by health outcomes (D, C, or H) and wealth (Y). Individuals are assumed to face probabilities of three mutually exclusive lifetime health profiles. The first is dying in the very short term (e.g., within a year) in an auto accident (D) with probability P_D, the second is contracting an eventually fatal cancer (C) with probability P_C, and the third "normal health" or all other health outcomes (H), with probability $1 - P_C - P_D$[3].

Totally differentiating Eq.(1) and setting dE(U) = 0 and dP_C = 0 results in the following expression for the marginal rate of substitution between income and risk of automobile death, which is also commonly referred to as the value of a statistical life (VSL) (see, for example, Hammitt [2000]).

:

$$VSL = \frac{dY}{dP_D} = \frac{U(H,Y) - U(D,Y)}{\dfrac{\partial E(U)}{\partial Y}} \qquad (2)$$

Alternatively, setting dP_D = 0 results in the following expression for the marginal rate of substitution between income and risk of cancer, which can be interpreted as the value of a statistical cancer case avoided (VSC):

$$VSC = \frac{dY}{dP_C} = \frac{U(H,Y) - U(C,Y)}{\dfrac{\partial E(U)}{\partial Y}} \qquad (3)$$

Combining Eqs. (2) and (3), and assuming that $U(D,Y) = 0$, the relationship between VSL and VSC can be rewritten as

$$MER = \frac{VSC}{VSL} = \left(1 - \frac{U(C,Y)}{U(H,Y)}\right). \qquad (4)$$

We refer to the term in brackets as the "mortality equivalence ratio" (MER) for avoided fatal cancer risks, which translates avoided fatal cancers into equivalent avoided accidental deaths. In other words MER = VSC/VSL. Therefore, if MER is less (greater) than 1 this implies that avoided fatal cancer risks are valued less (more) than avoided immediate mortality risks from car accidents. If, for example, MER equals 2, this implies that an avoided fatal cancer is "equivalent" to 2 avoided car fatalities.

---

[3] Since latent risks of cancer are only relevant if one does not die from immediate automobile fatality risks, P_C in this model should more accurately be replaced by $P_C(1-P_D)$ in Eq.(1); however the second order interaction of the two risks is small enough relative to P_c that excluding the interaction has little effect on the analysis.

The conceptual framework outlined above defines the general relationship between VSL and VSC; however, it does not specifically address how VSC is expected to vary with respect to characteristics of the cancer risks. In particular, it does not address the effects that latency, *t*, may have on the expected utility of the fatal cancer profile, U(C, Y) and therefore on VSC and MER. Including latency in equation (4) we express MER as:

$$MER(t) = \left(1 - \frac{U(C(t),Y)}{U(H,Y)} \right) \tag{5}$$

In this expression, the lifetime utility of the cancer health profile can be expressed as the discounted sum of utilities in future periods.

$$U(C(t),Y) = \left(\sum_{j=0}^{t-1} (s_j)(d_j)u^h(y_j) \right) + (s_t)(d_t)u^c(y_t) \tag{6}$$

where:

s_j = probability of surviving *j* periods into the future from the present (*j*=0)

d_j = time preference factor, discounting utility in period j to the present

y_j = consumption in period *j*

$u^k(y_j)$ = state dependent utility in period *j,* with k=c referring to cancer state and k=h referring to healthy state.

Similarly the lifetime profile for H, which is independent of the latency factor t, can be expressed as:

$$U(H,Y) = \sum_{j=0}^{\infty} (s_j)(d_j)u^h(y_j) \tag{7}$$

Based on this framework it is possible to formulate and examine specific hypotheses regarding the effects of latency period and perceived survival probabilities on preferences for avoiding cancer risks. In particular, as demonstrated in Appendix A,

- an increase in the cancer latency is expected under most circumstances to have a negative effect on MER, so the relative preference for avoiding cancer risks declines[4], and

- for any specified cancer latency t, an increase in the perceived survival probability to that period (s_t) is expected to have a positive effect on MER.

4. Empirical Methods

In our survey, which is described in more detail below, respondents are faced with a choice between two locations, A and B, where the only difference between the locations is the rate (i.e., risk) of fatal cancers (P_C^A vs. P_C^B) and auto deaths (P_D^A vs. P_D^B). Location A has fewer auto deaths and Location B has fewer cancers than the respondent's current location. In effect, they are presented with a pair of lotteries and asked to choose the one they prefer.

This choice is also illustrated in Figure 1, where Location A has fewer auto deaths and Location B has fewer stomach cancers than the respondent's current location. To compare the two options, we define the risk difference ratio (RDR) between A and B as:

$$RDR = \frac{P_D^B - P_D^A}{P_C^A - P_C^B} \tag{8}$$

The RDR therefore represents the slope (in absolute value) of the line between the A and B risk combinations. The respondent is assumed to choose the location that lies on the indifference line that is closer to the origin (i.e. the risk combination that provides the highest expected utility).

In Figure 1, both locations A and B are shown on the same indifference line. Indifference between the two areas (lotteries) implies that these areas offer the same expected utility:

$$P_D^A U(D,Y) + P_C^A U(C,Y) + (1 - P_D^A - P_C^A)U(H,Y) = \\ P_D^B U(D,Y) + P_C^B U(C,Y) + (1 - P_D^B - P_C^B)U(H,Y) \tag{9}$$

Assuming again that $U(D,Y) = 0$, and rearranging terms

---

[4] Using a somewhat different framework, Hammit and Liu (2004) also conclude that under most conditions, individuals' willingness to pay for reducing latent risks will be lower than for reducing current risks by the same amount.

$$MER = 1 - \frac{U(C,Y)}{U(H,Y)} = \left[\frac{P_D^B - P_D^A}{P_C^A - P_C^B} \right]^* = RDR^* . \qquad (10)$$

This equation show that the value of RDR that equates E(U) between two location -- RDR* -- is also equal to MER. In other words, MER represents the negative slope of the indifference curves in Figure 1; therefore, it defines the RDR that is consistent with indifference between the two locations.

By varying the values of P_C^B and P_D^A across respondents, the survey presents location choices that entail different RDRs. By observing how choices vary with respect to this variation in RDR, the survey responses can be used to estimate the average/expected value of MER. Equally important, they can be used to estimate how MER varies according to the characteristics of the cancer risks and the characteristics of respondents.

To model and interpret results using the discrete choice approach, we assume that MER varies in both systematic and stochastic ways across respondents. This assumption is formally expressed as

$$MER_i = \alpha + \beta X_i + \varepsilon_i . \qquad (11)$$

The systematic component of this expression describes MER as a function of X, which is a vector that includes both survey variables and characteristics, as well individual characteristics. The random component (ε_i) captures factors that are unobservable to the analyst and are assumed to vary randomly, identically, and independently across respondents.

In the discrete choice context, one does not observe MER_i for each respondent but rather a latent variable m_i^*, which can be characterized as

$$m_i^* = 0 \text{ if } RDR_i \geq MER_i \qquad (12a)$$

$$m_i^* = 1 \text{ if } RDR_i < MER_i . \qquad (12b)$$

In this case, m_i^* can be represented by a dummy variable, which is equal to 1 if the respondent prefers Location B (the location with fewer cancers) and 0 otherwise. In other words, the lower (higher) the value of RDR_i, the larger (smaller) is the reduction in cancers relative to auto deaths and the more likely that respondent i chooses Location B.

Assuming that ε_i is normally distributed $N(0, F^2)$, a probit model can be used to analyze the discrete choice responses and to estimate coefficients of the MER function (Eq. [4.9]) and F. The results of the probit analysis are discussed in Section 7 below.

A stated preference by individual i for Location B ($PREFERB_i=1$) indicates that she prefers the location offering a reduction in cancer risk over the location offering a reduction in auto death risk. It also implies that $MER_i > RDR_i$. Given the probability distribution of ε_i, the probability of preferring Location B can be expressed as:

$$
\begin{aligned}
&\Pr(PREFERB_i = 1) \\
&= \Pr(MER_i > RDR_i) \\
&= \Pr(\beta X_i - RDR_i > \varepsilon_i)
\end{aligned} \qquad (13)
$$

The last equality holds due to the symmetry of the distribution. By defining $\theta = \varepsilon / \sigma$, we define a standard normal random variable, $\theta \approx N(0,1)$, which implies that

$$
\begin{aligned}
&\Pr(PREFERB_i = 1) \\
&= \Pr\left(\left(\frac{\beta}{\sigma}\right)X_i - \left(\frac{1}{\sigma}\right)RDR_i > \theta_i\right) \qquad (14) \\
&= \Phi\left(\left(\frac{\beta}{\sigma}\right)X_i - \left(\frac{1}{\sigma}\right)RDR_i\right)
\end{aligned}
$$

By varying RDR randomly across individuals in the survey and controlling for factors included in X_i, a probit model can be used to estimate the vector β / σ and the scalar $1 / \sigma$. We refer to the corresponding probit coefficient estimates as the vector $\hat{\alpha}$ and the scalar $\hat{\gamma}$ respectively

Given the assumptions in Eq. (11) and the assumed distribution of the random term, *expected* MER for individual i can be expressed as:

$$
E(MER_i \mid X_i) = \left(\frac{\beta / \sigma}{-1 / \sigma}\right)X_i = \beta X_i \qquad (15)
$$

Therefore, using the probit results, expected MER_i can be estimated by $(-\hat{\alpha} / \hat{\gamma})X_i$.

5. Survey Design

The survey questionnaire was designed to be administered via WebTV to households in the U.S.. It was developed, pretested, and revised in several stages, using input from focus groups and multiple in-person cognitive interviews.

The sample for the survey was drawn from a panel of respondents prerecruited by Knowledge Networks, Inc. (KN). The only specific inclusion criterion was that respondents needed to be at least 18 years of age. The KN panel is based on a nationally representative, list-assisted, random-digit-dial (RDD) sample drawn from all 10-digit telephone numbers in the United States.

The survey instrument presented respondents with information on two hazards: death from a car accident and death from cancer. Respondents were randomly assigned one of three different types of cancer: stomach, liver, or brain cancer. The survey provided information on national averages and ranges of risk of each hazard, as well as information on cancer symptoms, treatment, and side effects. Importantly, the survey further explained that, although death from an automobile accident usually occurs almost immediately, cancers take years to develop before they are diagnosed (the latency period) and that the individual is typically sick for some time before death occurs (the morbidity period). Respondents were asked to assume for the purposes of the survey that the latency period has a length of t years (where respondents are randomly assigned values of t equal to 5, 15, or 25 years) and the morbidity period has a length of m years (with randomly assigned values of 2 or 5 years). Time lines, which are individualized to the respondent's reported age, were used to illustrate the differences in timing of exposure and death from the two hazards. Specifically, the time lines show that auto accidents and death are typically simultaneous occurrences, while demonstrating that exposure to the carcinogen occurs in year 1, diagnosis occurs in year $t + 1$, and death occurs in year $t + m+1$.

Respondents were then presented with a sequence of similar choice scenarios. They were asked to imagine that they have a job that requires them to move to one of two areas (A or B) for a period of 1 year. They must choose between the two areas, which differ only with respect to their exposure to the two hazards. They were asked to assume that their annual baseline risks from the two hazards-- the risk of dying in auto accident and the risk of dying of a specific cancer in their current area of residence – were both represented by 100 deaths per million people. They were then asked to choose between moving to Area A which has fewer auto accident deaths per million than their current location or Area B which has fewer fatal cancer deaths per million than their current location. Respondents were first introduced to the choice task with a few simplified practice questions. They were then presented with a choice scenario where they faced a tradeoff between avoiding risks of fatal auto accidents or avoiding risks of fatal cancers. An example choice scenario is shown in Figure 2.

Several aspects of the questionnaire design were randomly varied across respondents to test for their effects on a respondent's choices regarding risk reductions. These treatments were selected to test for scope effects and question-framing effects. They include the following:

- three different types of fatal cancers (stomach, liver, or brain cancer, each compared to fatal auto death risks);

- three different assumed latency periods for the cancer (5, 15, or 25 years);

- two different assumed morbidity periods for the cancer (2 or 5 years);

- two different formats for "introductory" choice questions, one in which Location A was clearly superior in the introductory scenarios and the other in which Location B was superior (included to test for framing effects in choice responses); and

- five different choice scenarios, each corresponding to a different RDR –

Thus, all together, there are 180 (3x3x2x2x5) different versions of the survey that are randomized across respondents.

Several other design characteristics are noteworthy. First, restricting the time frame to 1 year allows us to focus on risks from 1 year of exposure to the carcinogen and avoids the confounding issue of cumulative exposure to the carcinogen. This implicitly assumes an underlying dose-response model in which a single exposure can cause the cancer, as opposed to a model in which there is no risk of cancer until some threshold of cumulative exposures is reached. Second, emphasizing that the two new areas are exactly the same in every way but the risk exposure controls for the effects of other perceived location characteristics on reported preferences. Third, providing a baseline and maintaining new risk levels at or below the baseline controls for scenario rejection. Finally, after the practice questions and before the choice task, respondents were reminded of their individual time line for cancer exposure, diagnosis, and death.

6. Survey Data

The on-line WebTV survey was sent to a total of 1,351 households participating in the KN panel. To ensure proper functioning of the instrument, a subset of this sample—125 households—was initially contacted via email, and responses were acquired from about half this sample. After reviewing these responses and making minor adjustments to the instrument, email invitations were sent to the remainder of the sample.

By the end of March, 1,010 individuals (each from a different household) had submitted completed surveys to KN—a 73.7 percent invitation response rate. To achieve this rate of response, several of the 1,351 households were sent email and telephone reminders throughout the survey administration period.

To analyze responses to the main choice question, we excluded 136 respondents who "failed" the practice choice question. That is, if respondents did not indicate a preference for the "dominant" location (with fewer auto deaths *and* fewer fatal cancers), even after being given a chance to revise their response, it was assumed that they did not understand or were not willing to accept the choice scenario. An additional 17 respondents were dropped because, when presented with an automated follow-up description of their response to the first (nonpractice) choice question, they did not agree with the description

but also were not willing to revise their answer. We also excluded 69 respondents who did not have a preference for either location. Consequently, the size of the analysis sample is 788 respondents.

To investigate whether there were systematic differences between the analysis sample and (N=788) initial recruitment sample (N=1351), we conducted a probit analysis with respect to demographic characteristics. This analysis revealed that age, race, education, and household size were all significant determinants of whether respondents were included in the analysis sample. However, when this process was included as the first stage of a Heckman sample selection model, with the probit analyses described in Section 7 as the second stage, there was no evidence that the selection process led to biased estimates of the coefficients in the second stage model.

Descriptions and summary statistics for all the variables used in the analysis are provided in Tables 1 and 2 respectively. Overall, over half of the respondent (65 percent) preferred to location with lower cancer risks. The average age of the sample was 45.4 years, ranging from 18 to 93, average income was $50,600, and the average number of years of education was 12.5. Nineteen percent of the sample classified themselves as from a minority group.

The analysis also includes variables describing respondents' experience with and perceptions of cancer and automobile fatality risks. A relatively small percentage of the sample had experienced cancer themselves (CANCYOU, 8 percent) or had a close friend or relative who had experienced the cancer described to them in the survey (CANCFRIEND, 12 percent). A somewhat larger percentage had experienced a serious autoaccident (CARYOU, 18 percent) or had lost a close friend or relative to a car accident (CARFRIEND, 17 percent). On average, respondents believed that they had lower risk of dying of cancer or a car accident than others in their area; however, a large majority indicated that, for the purposes of the survey, they were able to assume that their risks were the same. Twenty five percent of respondents indicated that, in choosing between Locations A and B, they considered the possibility that a cure for cancer might be found.

Finally, to account for how differences in perceived survival probabilities affect preferences for avoiding cancer risks, we included data from the survey where respondents were asked: "How likely do you think it is, in percentage terms, that you will live for another X years or more?" The value for X corresponded to the cancer latency period that was presented to the respondent later in the survey. As expected, the average perceived survival rate was higher for X=5 (87 percent) than for X=15 (75 percent), which was also higher than for X=25 (68 percent). However, contrary to expectations and evidence from life tables, the perceived survival rate declined more rapidly from 5 to 15 years than from 15 to 25 years.

A main objective of the analysis is to evaluate how preferences for avoiding fatal cancer risks relative to auto death risks vary with respect to the relative size of the risk reductions and the length of the cancer latency period. Figure 3 provides a first look at this issue, by graphing the percent of respondents who preferred the location with lower cancer risks in relation to RDR and latency period. As expected, this percent generally declines with the RDR (i.e., larger relative reductions in auto death risks reduce the preference for the lower cancer risk location) and it declines with latency. The statistical significance of these results and their implications for calculating MERs are examined in the next section.

7. Model Results

Based on this framework, we estimated several probit specifications, all using PREFB as the dependent variable. In the simplest model specification, we assumed only random (no systematic) heterogeneity across respondents:

$$MER_i = \beta_0 + \varepsilon_i \qquad (16)$$

This model was estimated using probit specification (1) in Table 3. X_i in this case is simply the constant term, and, using Eq. (15), expected MER is estimated to be 2.3. In other words, without accounting for systematic heterogeneity across respondent characteristics or across survey versions, individuals were estimated to value avoided fatal cancer risks at somewhat more than twice the rate of fatal auto risks.

More complex models, which allow and control for heterogeneity in various ways are reported in specifications (2) through (5) in Table 3. All of these additional specifications control for and measure the effects of latency period on MER. These models consistently find that respondents' choices are significantly affected by differences in the latency period. As expected, individuals' preferences for avoiding fatal cancer risks (relative to automobile risks) decrease as the length of the cancer latency period increases. These specific results are described and discussed in more detail below.

Specifications (2) through (5) also control for the type of cancer (STOMACHC and BRAINC), the duration of cancer morbidity (MORB5), and the framing of introductory "practice" questions (INTROFORMAT), all of which were varied randomly across respondents. The brain cancer coefficient is consistently negative and statistically significant, whereas the coefficient for stomach cancer is never statistically significant. These results suggest that individuals have a significant preference for avoiding stomach and liver cancer risks compared to brain cancer risks.

Differences in the duration of cancer morbidity (MORB5) never have a significant effect on stated preferences in any of the model specifiations. The lack of an observed

morbidity duration effect on preferences may be because any negative effect of increasing the length of illness prior to death is offset by a corresponding delay in the time of death (for a given latency period, which in the survey is defined from the current period to the time of diagnosis). Alternatively, the difference between two and five years of morbidity may not have been large enough to influence respondents' choices.

In contrast, the framing of the introductory questions does have a significant effect on respondent choices in all of the model specifications. Respondents who received the format in which Location A (Location B) was clearly superior in the introductory scenarios were also less (more) likely to prefer Location B in the choice question involving a tradeoff between cancer and automobile risk reductions. Therefore, although these questions were included to help respondents understand the choice framework, they also appear to have created somewhat of a starting point bias for respondents..

Specification (3) also includes several demographic characteristics such as age, health status, education, and race, as well as the variables characterizing respondents' experience with and perceptions of cancer and automobile fatality risks. Of these variables, the only ones that have consistently significant (at a 10% level or less) effect on stated preferences are household income and whether they live in an MSA. Individuals in higher income households are less likely to prefer reducing cancer risks, whereas urban residents are more likely to do so. The effect of age on the relative preferences for avoiding cancers is negative, but it is not statistically significant in any specifications. To the extent that age affects preferences through perceived survival probabilities, these effects are explored in specification (4) and are discussed in more detail below.

Individuals' experience with and perceptions of cancer and automobile risks are also explored in specifications (2) to (4). Although individuals were asked to assume, in answering the choice questions, that their own risks were the same as others in their area, perceptions of higher than average cancer risks for themselves made them more likely to prefer the area with lower cancer risks. Similarly, perceptions of higher than average automobile risks for themselves made them more likely to prefer the area with lower automobile risks. These results suggest the respondents may have implicitly adjusted the risk reductions presented to them to fit their own circumstances. Individuals who had experienced cancer themselves were less likely to prefer reducing future cancer risks. This effect is not statistically significant, but it may reflect some adaptation to the illness. Also, respondents who had close friends or relatives die from cancer or automobile accidents were more and less likely, respectively, to prefer avoiding these risks. These effects are not statistically significant either, but they may be a sign of individuals' heightened fear or dread of these outcomes through indirect personal experience. Finally, individuals who had considered the possibility of a cancer cure were significantly less likely to prefer avoiding latent cancer risks. This finding is consistent with individuals

expecting to derive higher utility from a future cancer health state, if the cancer has a lower chance of being fatal.

To evaluate the effects of latency on preferences, specifications (2) and (3) estimate separate coefficients for the 15 year and 25 year latency period dummies ($\hat{\alpha}_1$ for LAT15 and $\hat{\alpha}_2$ for LAT25), with the 5 year latency period as the reference condition. Both coefficients are negative, significantly different from zero, and significantly different from one another. Therefore, as expected, longer latency periods reduce the relative preference for avoiding cancer risks. To specifically explore the effect of latency on MER, we define latency-specific MER as MER(t), such that.

$$MER(t)_i = \beta X(t)_i + \varepsilon_i \qquad (17)$$

Using this definition and equation (15) and setting all variables in X set at their sample means (except for LAT 15 and LAT 25), we estimate separate expected MERs for the 5, 15, and 25 year latency periods[5]. The predicted values range from 3.23 for the 5 year latency to 1.54 for the 25 year latency.

If MER declines linearly with respect to latency, adapting equation (17) we then have:

$$MER(t)_i = MER(0)_i - \phi t + \varepsilon_i \qquad (18)$$

Testing the linearity restriction in specifications (2) and (3) is therefore equivalent to testing whether $\hat{\alpha}_2 - \hat{\alpha}_1 = \hat{\alpha}_1$. Applying a Wald test to the estimated coefficients, we found that in both cases the linearity restriction cannot be rejected (at a 5% level of significance).

In specification (4), we impose the linearity assumption by replacing the latency dummy variables with a continuous variable (LATENCY = t). With this model and equation (18), it is also possible to extrapolate the results and estimate:

- E[MER(0)]—the implied expected MER if latency were zero and the onset of cancers, like auto deaths, were immediate[6] and
- $t*$—the length of the latency period that would be required to make expected MER equal 1 (i.e., to make individuals indifferent between reducing fatal cancer and auto death risks).

To estimate E[MER(0)] we set LATENCY=0 and the other explanatory variables in X at the sample mean. The results, which are reported Table 3 indicate that on average, avoided fatal cancer risks *without* latency would be valued at over three times avoided automobile death risks.

To estimate t*, we define the following condition

---

[5] For the MER calculations, the values of CANCERRISK and CARRISK were set at 3 – equal to the same risk as the average individual in their area – rather than at the sample means, which were somewhat smaller.

[6] In the case of cancer, death would still be delayed by the duration of morbidity.

$$E[MER(t^*)_i] = \beta X(0)_i - \phi t^* = 1 \qquad (19)$$

and solve for t*. Using specification (4) and the mean sample characteristics, we estimated t* to be roughly 32 years. In other words, latency periods for cancer risks would need to be on average over 30 years to make individuals indifferent between reducing fatal cancer and auto death risks.

The final specification in Table 3 was included to specifically examine how individuals' perceived survival probabilities (SURVIVERATE) modified the effect of latency period on their choices. Because these survival probabilities are specific to the latency period presented to each respondent, they are interacted with their corresponding latency dummy variables in specification (4). As expected, the coefficients on the survival probabilities are all positive and significant. These results suggest that individuals prefer to avoid cancer risks in X years if they are more likely to be alive in X years. The size of these three coefficients are also ordered as expected, decreasing in magnitude as the latency period increases from 5 to 15 to 25 years. The difference between 5 and 15 years is not statistically significant, but the difference between 25 years and the two shorter latency periods is significant (at a 0.05 level) in both regressions. Therefore, even after controlling for differences in perceived survival probabilities, latency still has a significant effect on individuals' preferences for avoiding future risks.

8. Summary and Conclusions

Environmental protection programs, as well as food safety and many other public health programs, often benefit society by reducing cancer risks. Because many avoided cancers are expected to be fatal, these health benefits are often measured in terms of "statistical lives saved," and they are typically valued using available estimates of VSL. One of the drawbacks of this benefits assessment approach is that few of these VSL estimates, which reflect individuals' WTP to reduce mortality risks, have been specifically designed to capture preferences for avoided cancer fatalities. In most cases, these estimates have been derived in the context of immediate and or accidental deaths.

There are at least two reasons why VSL estimates based on risks of immediate accidental deaths may not be appropriate for valuing avoided fatal cancer risks. The first is that individuals may view cancer deaths as being qualitatively different from accidental deaths, perhaps associating particular dread or fear with cancers. The second reason is that cancer risks are often likely to involve extended latency periods between the time of exposure and the observable effects of illness.

The purpose of this study has therefore been to directly explore differences in individuals' preferences regarding fatal accidental and fatal cancer risks. First, when

17

directly comparing risk reductions of the same magnitude, is there evidence of a "cancer premium"? That is, do individuals systematically prefer avoiding cancer risks and, if so, by how much? Second, to what extent does cancer latency modify differences in preferences for the two types of risks?

To address these issues we administered a web-based preference elicitation survey to a general population sample of adults in the US. The focal point of the survey was a choice task that asked respondents to choose locations that offered either lower automobile fatality or cancer risks. The relative risk reductions, as well as the characteristics of the cancers, were varied randomly across respondents.

The main findings of the survey are that individuals made choices that revealed (1) a significant cancer premium and (2) a cancer premium that declined with the length of the cancer latency period. On average, to make individuals indifferent between avoiding the two types of risks, they required risk reductions for fatal cancers that were two to three times larger than for fatal automobile risks. Preferences for avoiding cancer risks were also significantly reduced by longer latency periods; however, the survey results indicate that latency periods greater than 30 years were generally required to offset the effects of a cancer premium.

Our analysis also finds that the effect of latency periods on preferences is itself affected by individuals' perceived survival probabilities. The lower the chance of survival for a given latency period, the less individuals preferred avoiding cancers with that latency. Our results indicate that perceived survival probabilities (less than 100 percent) are one reason that individuals discount future cancer risks; however, this discounting persists to some extent even after accounting for survival.

Our results suggest that using current estimates of VSL based mainly on data from accidental death risks may not be appropriate when evaluating the benefits of avoided cancer risks. Unless cancer latency periods exceed 30 years, these VSL estimates are likely to understate the true benefits of reduced cancer risks. Further research using different preference elicitation and measurement approaches is needed to confirm these findings; however, they provide more evidence that policy analyses would benefit from VSL estimates that are better tailored to the risk reductions contexts in which they are applied.

References

Cameron, T.A. and J.R. DeShazo. 2004. "An Empirical Model of Demand for Future Health States when Valuing Risk-Mitigating Programs." University of Oregon, Economics Department Working Paper 2004-11. (http://ideas.repec.org/p/ore/uoecwp/2004-11.html)

Carson, R.T., and R.C. Mitchell. 2000. "Public Preference toward Environmental Risks: The Case of Trihalomethanes." University of California San Diego, Discussion Paper.

du Vair, P and J. Loomis. 1993. "Household's Valuation of Alternative Levels of Hazardous Waste Risk Reductions: An Application of the Referendum Format Contingent Valuation Method." *Journal of Environmental Management*. 39(2): 143-155.

Hammitt, J. K. 2000. "Valuing Mortality Risk: Theory and Practice." *Environmental Science and Technology* 34(8): 1396-1400.

Hammitt, J. K. 1990. "Risk Perceptions and Food Choice: an Exploratory Analysis of Organic- versus Conventional-produce Buyers." *Risk Analysis*, 10(3):367-74.

Hammitt, J. K. and J. Liu. 2004. "Effects of Disease Type and Latency on the Value of Mortality Risk." *Journal of Risk and Uncertainty* 28(1): 73-95.

Krupnick, A., A. Alberini, N. Simon, and M. Cropper. 2005. "Willingness to Pay for Mortality Risk Reductions: Does Latency Matter?" Resources for the Future Discussion Paper (www.rff.org/documents/rff-dp-04-13.pdf)

Magat, W. A., W. K. Viscusi, and J. Huber. 1996. "A Reference Lottery Metric for Valuing Health." *Management Science* 42(8):1118-1130.

Smith, V. K. and W. H. Desvousges. 1987. "An Empirical Analysis of the Economic Value of Risk Changes." *The Journal of Political Economy*, 95(1):89—114.

Subramanian, U., Cropper, M. 2000. "Public choices between life saving programs: The Tradeoff Between Qualitative Factors and Lives Saved," *Journal of Risk and Uncertainty* 21(1): 117-149.

Tsuge, T., A. Kishimoto, and K. Takeuchi. 2005. "A Choice Experiment Approach to the Valuation of Mortality." *Journal of Risk and Uncertainty* 31(1): 73-95.

U.S. Environmental Protection Agency (EPA). 2000. *"An SAB Report on EPA's White Paper Valuing the Benefits of Fatal Cancer Risk Reduction."* Letter to the EPA Administrator from the EPA Science Advisory Board, July 27, 2000. EPA-SAB-EEAC-00-013.

Figure 1. Preference Map for Two Categories of Risk

Number of
Stomach
Cancers per
Million

Current
Location

100

Location
A

Location B

Number of
Automobile Deaths
per Million

100
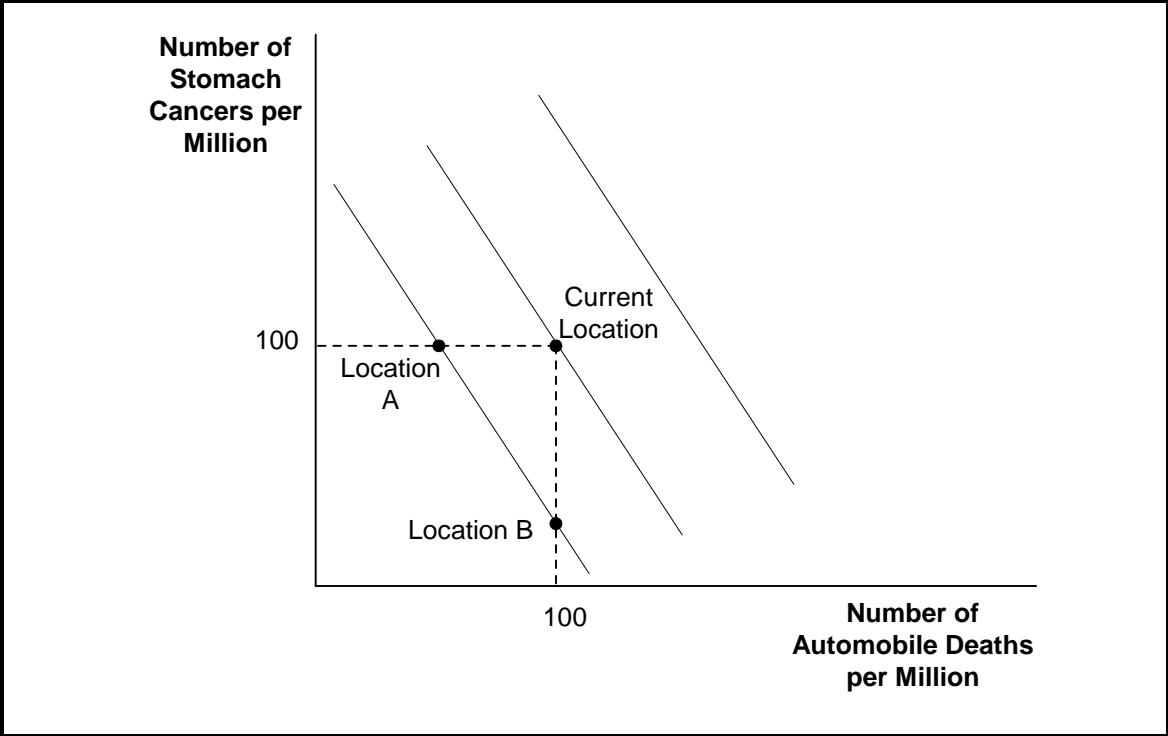
Figure 2. Example Choice Task in the Risk Tradeoff Survey

The table below summarizes the <u>only</u> differences between Location A and Location B.

	Location A	Location B
Car accident deaths (per year)	50 per million people	100 per million people
Fatal stomach cancers (caused per year)	100 per million people	50 per million people

If you had to move to one of these locations, which one would you prefer?

Location A	Location B
◉	○
No Preference Between Location A and Location B	
○	

Figure 3. Preferences for Avoiding Fatal Cancer Risks Relative to Auto Death Risks
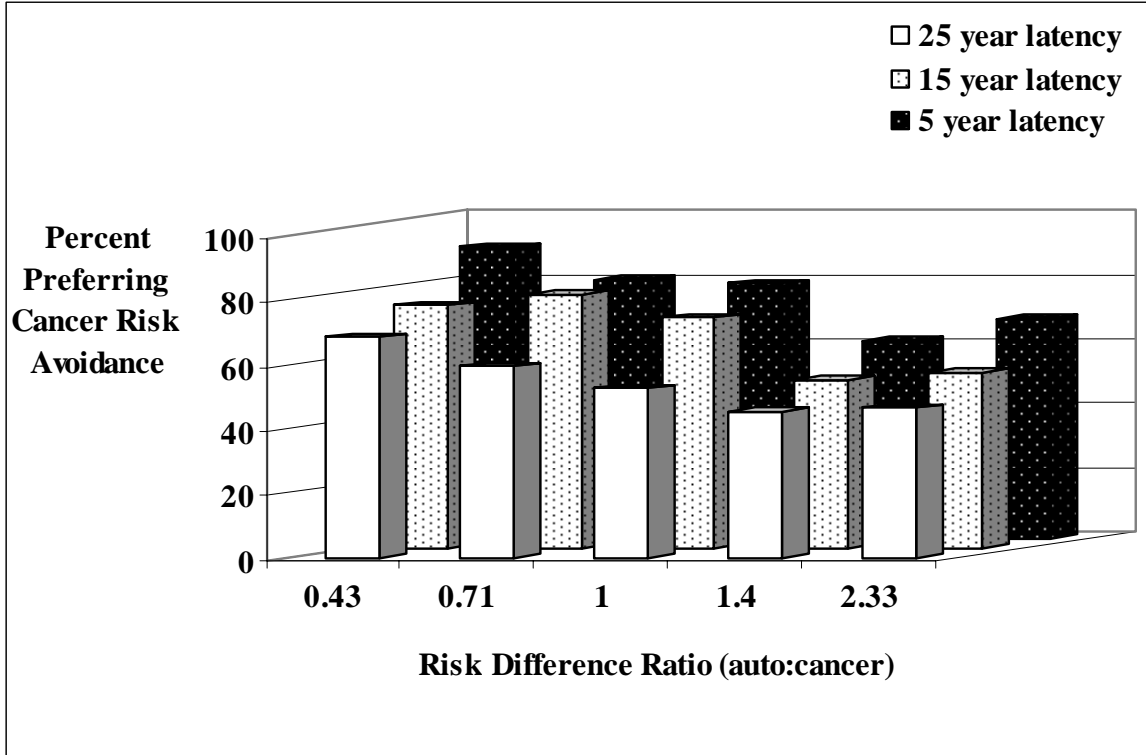
Table 1. Descriptions of Analysis Variables

Variable Name	Description
PREFERB	= 1 if choose "Prefer Location B"
RDR	= Risk difference ratio presented in choice table
LAT15	= 1 if 15-year latency period
LAT25	= 1 if 25 year latency period
LATENCY	= Latency period (5, 15 or 25)
SURVIVERATE	Self-reported (perceived) probability of surviving for duration of latency period
LAT5SURV	= Interaction between 5-year latency period and perceived chance of survival during latency period (LAT5*SURVRATE)
LAT15SURV	= Interaction between 15-year latency period and perceived chance of survival during latency period (LAT15*SURVRATE)
LAT25SURV	= Interaction between 25-year latency period and perceived chance of survival during latency period (LAT25*SURVRATE)
MORB5	= 1 if 5-year morbidity period (= 0 if 2-year morbidity period)
INTROFORMAT	= 1 if Location A dominates Location B in introductory choice questions
STOMACHC	= 1 if stomach cancer version
BRAINC	= 1 if brain cancer version
AGE	= respondent's age
HEALTHNOW	= Self assessment of respondent's current health status
GENDER	= 1 if male
MINORITY	= 1 if race non-white
EDUC	Number of years of education
HHINCOME	Household income ($'000)
HHSIZE	Household size
MSA	= 1 if respondent lives in an MSA
CANCERYOU	= 1 if respondent ever had cancer
CANCFRIEND	= 1 if friend or relative had experienced the cancer described in the survey
CANCERCURE	= 1 if considered possibility of cure for cancer during latency period
CANCERRISK	Self-rated fatal cancer risk compared to average (1= much lower, 5 = much higher)
CARYOU	= 1 if hospitalized because of a car accident
CARFRIEND	= 1 if friend or relative died in car accident in last 10 years
CARRISK	Self-rated fatal car risk compared to average (1= much lower, 5 = much higher)

Table 2. Summary Statistics for Analysis Variables

Variable	N	Mean	SD	Min	p25	p50	p75	Max
PREFERB	788	0.65	0.48	0	0	1	1	1
RDR	788	1.16	0.65	0.43	0.71	1	1.4	2.33
LAT15	788	0.32	0.47	0	0	0	1	1
LAT25	788	0.36	0.48	0	0	0	1	1
LATENCY	788	15.36	8.21	5	5	15	25	25
SURVIVERATE	788	0.77	0.27	0	0.6	0.9	1	1
MORB5	788	0.50	0.50	0	0	1	1	1
INTROFORMAT	788	0.49	0.50	0	0	0	1	1
STOMACHC	788	0.34	0.47	0	0	0	1	1
BRAINC	788	0.33	0.47	0	0	0	1	1
AGE	788	45.39	16.97	18	31	44	58	93
HEALTHNOW	787	2.54	0.92	1	2	3	3	5
GENDER	788	0.48	0.50	0	0	0	1	1
MINORITY	788	0.19	0.40	0	0	0	0	1
EDUC	788	12.45	3.22	6	12	14	14	16
HHINCOME	788	50.56	36.49	2.5	22.5	45	67.5	187.5
HHSIZE	788	2.61	1.24	1	2	2	3	8
MSA	788	0.84	0.37	0	1	1	1	1
CANCERYOU	788	0.08	0.27	0	0	0	0	1
CANCFRIEND	782	0.12	0.33	0	0	0	0	1
CANCERCURE	783	0.25	0.43	0	0	0	0	1
CANCERRISK	782	2.56	0.84	1	2	3	3	5
CARYOU	788	0.18	0.38	0	0	0	0	1
CARFRIEND	785	0.17	0.38	0	0	0	0	1
CARRISK	783	2.47	0.95	1	2	3	3	5

Table 3. Analysis of Risk Tradeoffs: Probit Results for Location Choice

	Dependent Variable = PREFERB									
	(1)		(2)		(3)		(4)		(5)	
Variable	Coef.	z-statistics	Coef.	z-statistics	Coef.	z-statistics	Coef.	z-statistics	Coef.	z-statistics
CONSTANT	0.816	8.56	1.262	8.28	1.788	4.01	1.947	4.31	0.302	0.57
RDR	-0.355	-5.03	-0.350	-4.89	-0.374	-4.92	-0.373	-4.91	-0.382	-4.98
LAT15			-0.316	-2.60	-0.300	-2.34				
LAT25			-0.634	-5.39	-0.632	-5.06				
LATENCY							-0.032	-5.10		
LAT5SURV									1.239	4.99
LAT15SURV									1.061	4.09
LAT25SURV									0.710	2.57
MORB5			0.031	0.33	-0.042	-0.43	-0.042	-0.42	-0.039	-0.39
INTROFORMAT			-0.173	-1.83	-0.207	-2.10	-0.208	-2.11	-0.234	-2.36
STOMACHC			0.111	0.95	0.096	0.78	0.097	0.80	0.071	0.57
BRAINC			-0.215	-1.86	-0.235	-1.97	-0.235	-1.96	-0.253	-2.10
AGE					-0.005	-1.57	-0.005	-1.56	0.002	0.53
HEALTHNOW					-0.046	-0.82	-0.046	-0.80	0.004	0.07
GENDER					-0.053	-0.54	-0.053	-0.54	-0.022	-0.22
MINORITY					-0.100	-0.79	-0.101	-0.79	-0.106	-0.82
EDUC					-0.008	-0.47	-0.008	-0.47	-0.013	-0.83
HHINCOME					-0.002	-1.79	-0.002	-1.79	-0.003	-2.15
HHSIZE					0.006	0.14	0.006	0.14	0.008	0.20
MSA					0.251	1.89	0.250	1.89	0.238	1.83
CANCERYOU					-0.296	-1.49	-0.296	-1.48	-0.192	-0.94
CANCFRIEND					0.249	1.53	0.250	1.55	0.268	1.67
CANCERCURE					-0.401	-3.60	-0.400	-3.58	-0.409	-3.63
CANCERRISK					0.239	3.73	0.240	3.75	0.240	3.71
CARYOU					-0.027	-0.21	-0.027	-0.21	-0.033	-0.26
CARFRIEND					-0.170	-1.27	-0.170	-1.28	-0.152	-1.15
CARRRISK					-0.207	-3.69	-0.206	-3.69	-0.189	-3.33
Number of obs	788		788		775		775		775	
Calculated Values										
E[MER]	2.30									
E[MER(0)]							3.67			
E[MER(5)]			3.31		3.23		3.24			
E[MER(15)]			2.41		2.43		2.39			
E[MER(25)]			1.50		1.54		1.54			
t*							31.40			

Appendix: Proofs of latency and survival probability effects on MER

A.1 The effect of cancer latency on MER

As shown in equation (5), the effect on MER of increasing latency (t) depends on how it affects the lifetime utility of the cancer profile. Expanding on equation (6), the effect on $U(C(t), Y)$ of increasing latency by one period, from t to t+1, can be written as

$$
\begin{aligned}
U(C(t+1),Y) & -U(C(t),Y) \\
&= \left(\sum_{i=1}^{t-1}(s_i)(d_i)u^h(y_i) \right) + (s_t)(d_t)u^h(y_t) + (s_{t+1})(d_{t+1})u^c(y_{t+1}) \\
&\quad - \left(\sum_{i=1}^{t-1}(s_i)(d_i)u^h(y_i) \right) - (s_t)(d_t)u^c(y_t) \\
&= (s_t)(d_t)\left[u^h(y_i) - u^c(y_t) \right] + (s_{t+1})(d_{t+1})u^c(y_{t+1})
\end{aligned}
\tag{A.1}
$$

For simplicity, the duration of each time period, as indexed by i, is the same as the duration of cancer morbidity (i.e., between diagnosis and death). The first term in equation A.1 will be positive as long as the utility of a period in normal health, $u^h(y)$, is greater than with cancer, $u^c(y)$. Even if the utility of a period (t or t+1) with cancer is negative, the second term in this expression will also be less in absolute value terms than the first, as long as $u^h(y)$ is positive (and y_{t+1} is not substantially less than y_t). Consequently, an increase in latency should increase $U(C(t),Y)$ and decrease MER. This result is consistent with the intuition that extending cancer latency will reduce aversion to cancer risks.

A.2 The effect of survival probability on MER

In contrast to a change in cancer latency, an increase in survival probability, s_t, will affect MER through both the cancer and the normal health utility profiles. To examine the effect of increasing s_t on the MER, we must examine its effect on the ratio $U(C.Y)/U(H.Y)$. To do this we first define the following expressions:

$$A = \sum_{i=0}^{t-1}(s_i)(d_i)u^h(y_i)$$

$$B = (d_t)u^c(y_t)$$

$$C = \sum_{j=t}^{\infty}(s_{jt})(d_j)u^h(y_j) \tag{A.2}$$

Where $\dfrac{U(C(t),Y)}{U(H,Y)} = \dfrac{A + s_t B}{A + s_t C}$

s_{ij} = probability of surviving to period j conditional on surviving to period t

Differentiating the lifetime utility ratio R with respect to s_t, we get:

$$\frac{\partial R}{\partial s_t} = \frac{A(B-C)}{(A+s_tC)^2} < 0 \qquad \text{if } A>0 \text{ and } B<C \qquad (A.3)$$

These results imply that as long as the healthy state provides positive utility, and this utility is greater than the utility in the cancer state, then, *for a given latency t*, increasing the probability of survival to t, will decrease U(C(t),Y) and increase MER. In other words, for a given cancer latency, increasing the probability of survival for that period will increase aversion to cancer risks.

DISCUSSION PAPER

Adjusting the Value of a Statistical Life for Age and Cohort Effects

Joseph E. Aldy and W. Kip Viscusi

RESOURCES
FOR THE FUTURE

Adjusting the Value of a Statistical Life for Age and Cohort Effects

Joseph E. Aldy and W. Kip Viscusi

Abstract

To resolve the theoretical ambiguity in the effect of age on the value of statistical life (VSL), this article uses a novel, age-dependent fatal risk measure to estimate age-specific hedonic wage regressions. VSL exhibits an inverted-U shaped relationship with age. In the year 2000 cross-section, workers' VSL rises from $3.2 million (ages 18–24), to $9.9 million (35–44), and declines to $3.8 million (55–62). Controlling for birth-year cohort effects in a minimum distance estimator yields a peak VSL of $7.8 million at age 46 and flattens the VSL-age relationship. The value of statistical life-year also follows an inverted-U shape with age.

Key Words: value of statistical life, job risks, hedonic wage regression, VSLY

JEL Classification Numbers: J17, I12

Contents

Adjusting the Value of a Statistical Life for Age and Cohort Effects

Joseph E. Aldy and W. Kip Viscusi*

A strident controversy with respect to the value of life has been whether the benefit of reducing risks to the old are less than for younger age groups. In particular, should there be a so-called "senior discount" when assessing the value of reduced risks to life? This question has drawn the attention of policymakers in a number of countries. In 2000, Canada employed a value of statistical life (VSL) for the over-65 population that is 25 percent lower than the VSL for the under-65 population (Hara and Associates 2000). In 2001, the European Commission recommended that member countries use a VSL that declines with age (European Commission 2001). In 2003, the U.S. Environmental Protection Agency (EPA), which has traditionally employed a constant value of a statistical life to monetize mortality risk reductions irrespective of the age of the affected population, conducted analyses of the Clear Skies initiative that included a "senior discount."[1] This effort to apply such a discount in its Clear Skies initiative analyses generated a political firestorm and ultimately led to abandonment of any age adjustments in benefit values assigned by the Agency.[2]

Intuitively one might expect that older individuals may value reducing risks to their lives less because they have shorter remaining life expectancy. The commodity they are buying through risk reduction efforts is less than for younger people. Carrying this logic to its extreme,

---

[1] In the "senior discount" analyses, the EPA provided two alternatives to account for age. One approach was based on a standard value of a statistical life-year approach that explicitly accounts for life expectancy. The second approach assumed that individuals over age 70 had a value of statistical life equal to 63 percent of the value for those under 70.

[2] For a sense of the political reaction and EPA's decision to discontinue the use of an age-based value of statistical life, refer to "EPA Drops Age-Based Cost Studies," *New York Times*, May 8, 2003; "EPA to Stop 'Death Discount' to Value New Regulations," *Wall Street Journal*, May 8, 2003; and "Under Fire, EPA Drops the 'Senior Death Discount,'" *Washington Post*, May 13, 2003.

the VSL would peak at birth and decline steadily thereafter. For models in which consumption is constant over the life cycle, Jones-Lee (1989) showed that the VSL should decrease with age. Whether consumption will in fact be constant over time depends critically on the presence of perfect capital and insurance markets.

Numerous theoretical studies have shown that the age variation in VSL becomes more complex once changes in consumption over time are introduced into the analysis. Changes in consumption levels and wealth over the life cycle influence risk-money tradeoffs in a complex manner. Johansson (2002) concluded that the theoretical relationship between the VSL and age is ambiguous and could be positive, negative, or zero. Often theoretical studies, however, have imposed additional structure on the analysis, implying that there is either an inverted U-shaped relationship between the value of statistical life and age or that VSL decreases with age. The simulations by Shepard and Zeckhauser (1984) show a steadily declining value of life if there are perfect annuity and insurance markets, and an inverted-U VSL-age relationship in an economy with no borrowing or insurance, as do Johansson (1996) and Ehrlich and Yin (2004). Rosen (1988), Arthur (1981), and Cropper and Sussman (1988) also present simulation results with VSL decreasing with age.

Empirical evidence based on labor market data may be instructive in resolving the theoretical ambiguity in the VSL-age relationship. Viscusi and Aldy (2003) review eight studies of labor markets in Canada, India, Switzerland, and the United States that included an age-mortality risk interaction term in their hedonic wage analysis. Five studies estimated statistically significant coefficient estimates on the age-risk interaction and all find a negative effect indicating that older workers value risks to their lives less.[3] These results imply implausibly low VSL levels with negative VSL amounts beginning at ages ranging from 42 to 60. The failure of labor market evidence to resolve the age variation issue may stem in part from data limitations. All these labor market studies use fatality risk data that are based on industry averages rather than age-specific values, causing potential biases, where the magnitude of the bias varies with age. If, for example, average industry fatality risks for workers of all ages overstate the risks faced by older workers, the estimated implied VSL amounts for older workers will understate the wage-risk tradeoffs that are actually being made.

---

[3] These studies are reviewed in Section 8 of Viscusi and Aldy (2003). In contrast, a recent study by Smith et al. (2004) has found that the value of statistical life is increasing with age and risk aversion for workers 51–65 years of age.

All previous papers assessing how the compensating differential for job mortality risk varies with age have employed cross-sectional survey data. By using a single cross-section, such approaches confound the cohort-specific influence and age-specific effects on the estimated compensating differential. The cohort influence based on the year of birth should have an unambiguous effect on VSL. Lifetime incomes are rising over time, and the VSL has a positive income elasticity of 0.5 to 0.6.[4] Because older workers belong to an earlier cohort with lower lifetime incomes, they will tend to be willing to pay less for a given risk reduction, implying a lower VSL. The pure age effect is less clear-cut. As a worker ages, there are fewer years of remaining life expectancy, implying lower benefits for a given risk reduction, which should reduce the worker's willingness to pay to reduce risk. This effect is unambiguous if capital markets are perfect. In a world with imperfect capital markets, however, lower income younger workers will not be able to borrow against higher future expected earnings. This will depress their VSLs at young ages until borrowing constraints become less stringent, resulting in an age-related VSL trajectory similar to the inverted-U shape of life-cycle consumption patterns. Extending the traditional analysis to a pooled series of cross-sections will enable us to distinguish age effects from cohort effects. Two separate, but both policy-relevant, questions can then be considered: (1) How does the value of life vary with age across the population? and (2) How do differences in cohorts influence this relationship?

This article extends the previous literature in several respects. Because our focus is on risky labor market decisions, we make job risk decisions a choice variable in a life-cycle consumption model in Section I, deriving an expression for VSL in this context. In Section II, we present empirical estimates how the VSL varies over the life cycle through conventional hedonic wage equations and a minimum distance estimator. These results reflect two innovations to this literature: (1) we employ age-specific job mortality and nonfatal injury risks in our hedonic wage analyses; and, (2) we estimate how the VSL changes over the life cycle by pooling eight years of cross-sectional data and by using a minimum distance estimator that controls for cohort effects based on year of birth. In these empirical approaches, the VSL rises and then falls across the population and over the life cycle. In the cross-sectional analysis, the VSL peaks at age 39 and subsequently declines so that the VSL for workers in their early 60s have values of about $2 million. In the cohort-adjusted analysis, the VSL peaks at age 46, and experiences a more modest

---

[4] See Viscusi and Aldy (2003) for a meta-analysis of the VSL income elasticity value.

decline to about $5 million by age 62.[5] In Section III, we calculate age-specific values of statistical life-years (VSLY) from our age-VSL profiles and find that VSLYs also take an inverted-U shape with a peak at an older age than the VSLs. In the cross-sectional analysis, the VSLY peaks at $375,000 at age 45 and subsequently declines to about $150,000 in workers' early 60s. In the cohort-adjusted analysis, the VSLY peaks at $401,000 at age 54, and experiences a more modest decline to about $350,000 by age 62. Section IV concludes the paper.

I. Wage-Risk Tradeoffs over the Life Cycle

The standard approach in the life-cycle VSL literature employs a time-separable utility function in one consumption good, integrated over the life-cycle subject to a discount function and a survival function, as in Shepard and Zeckhauser (1984), Rosen (1988), Johansson (1996, 2002), and Johannesson et al. (1997). The only choice variable is the level of consumption over time. In these analyses, the value of statistical life is given by a representative agent's expected present value of consumer surplus conditional on having achieved a given age. For example, Shepard and Zeckhauser represent this as the ratio of expected remaining lifetime utility to the marginal utility of consumption.

To motivate our empirical work, we provide a model of wage-risk tradeoffs in a life-cycle setting. We modify and extend the standard life-cycle approach to explicitly account for the choice of job fatality risk on the survival function and the worker's wage. Since a change in job fatality risk affects both the worker's wage and life expectancy, our approach provides an alternative illustration of the VSL varies over the life cycle by characterizing the wage-risk tradeoff given the impacts of both on future consumption. By incorporating a compensating differential framework in our model, we can demonstrate how the wage-risk trade-off varies over the life cycle, which is what we will estimate in our empirical work presented below.

Our simple model indicates variations in VSL, but the linkage is ambiguous. This life-cycle model can illustrate the influences – especially the life-cycle variation in consumption – that can generate an inverted U-shaped relationship between VSL and age. The worker's problem can be characterized by maximizing discounted expected remaining lifetime utility:

(1) $$\max_{p,c} EU(\tau) = \int_{\tau}^{\infty} u[c(t)]\sigma[t;\tau,p(t)]e^{-rt}dt \,,$$

---

[5] All VSL estimates are presented in year 2000 dollars in this paper.

subject to

(2a) $\dot{k}(t) = rk(t) + w[t, p(t)] - c(t) + f(t)$,

(2b) $k(t) \geq 0$,

and

(2c) $\lim\limits_{t \to \infty} k(t)e^{-rt} = 0$,

where

p represents the probability of dying on the job,

$u(c)$ represents the utility of consumption, c, and $u'(c) \geq 0$, $u''(c) \leq 0$,

k represents assets,

w represents labor income,

e^{-rt} represents the discount function,

$\sigma[t; \tau, p(t)]$ represents the survival function, i.e., the probability of surviving to age t, given that the individual has reached age τ,[6]

r represents the return on assets, and

$f(t)$ represents the net amount received through an actuarially fair annuity represented by the condition:

$$\int\limits_{0}^{\infty} e^{-rt} \sigma[t;0, p(t)] f(t) dt = 0 .[7]$$

The worker's expected utility is represented in (1) as the sum of period utilities weighted by a discount factor and the probability that the worker will survive to that period conditional on the worker's current age. The worker maximizes this expected utility expression subject to the constraints: (2a) represents the dynamic budget constraint, and it allows for the worker's assets

---

[6] This expression of the survival function follows Johansson (1996): $\sigma[t; \tau, p(t)] = \sigma[t; p(t)] / \sigma(\tau)$.

[7] To simplify notation, we have followed Shepard and Zeckhauser and assumed that the rate of time preference in the discount function is equal to the rate of return on assets, and that this rate is time-invariant. Allowing for the rate of time preference to differ from the return on assets would not substantively influence the primary conclusion of this analysis that the age-VSL relationship is ambiguous.

to change over time based on capital income $rk(t)$, labor income $w[t, p(t)]$, consumption $c(t)$, and net annuity receipts $f(t)$; (2b) provides a no debt condition; and (2c) is the standard no Ponzi game condition. The actuarially fair annuity envisioned here is similar to that in Shepard and Zeckhauser's (1984) perfect markets case, and the annuity allows for the worker to borrow against human capital during early years of life to provide for consumption smoothing.

The present value Hamiltonian, conditional on having lived to age τ, is given by:

(3) $$H(t) = u[c(t)]\sigma[t; \tau, p(t)]e^{-rt} + \lambda(t)\big[rk(t) + w[t, p(t)] - c(t) + f(t)\big]$$

where $\lambda(t)$ represents the present value costate variable. The first-order conditions for the Hamiltonian are:

(4) $$\frac{\partial H}{\partial c} = u_c \sigma e^{-rt} - \lambda = 0 \, ,^{8}$$

(5) $$\frac{\partial H}{\partial p} = u\sigma_p e^{-rt} + \lambda w_p = 0 \, ,$$

and

(6) $$-\frac{\partial H}{\partial k} = \dot{\lambda} \rightarrow \dot{\lambda} = -r\lambda \, .$$

To see more generally how the value of a statistical life varies with age, we rearrange (5), differentiate with respect to time, where time derivatives are denoted by a dot over the variables in question, and substitute into (6), yielding:

(7) $$\frac{\dot{w}_p}{w_p} = \frac{\dot{u}}{u} + \frac{\dot{\sigma}_p}{\sigma_p}$$

The percentage change over time in the compensating differential for job fatality risk is equal to the percentage change over time in utility and the percentage change over time in the change in the survival function with respect to job fatality risk.[9] This expression holds irrespective of the assumption of actuarially fair annuity markets, although the assumption regarding these markets clearly influences the change in utility over the life cycle. The sign on

---

[8] This is essentially identical to equation 12 of Shepard and Zeckhauser (1984).

[9] Note that the survival function, $\sigma[t; \tau, p(t)]$, and the discount function, e^{-rt}, implicitly enter equation (7) through their influence on the optimal consumption and job fatality risk paths.

equation (7) is ambiguous without imposing restrictions on the survival function and specifying the assumptions regarding annuity markets. This ambiguity is consistent with the life-cycle model provided by Johansson (2002) and the simulation results based on the life-cycle model in Shepard and Zeckhauser (1984). This theoretical ambiguity motivates our interest in resolving empirically how the value of a statistical life varies over the life cycle.

II. Hedonic Wage Methods and Results

To assess empirically the age-VSL relationship, we have expanded the standard hedonic wage framework in two ways. First, using our new and more refined age-specific job-related mortality and injury data, we estimated hedonic wage regressions that allow for the compensating differential for these risks to vary among five age groups. These results indicate how the VSL varies with age across the population. Second, we develop a minimum distance estimator that incorporates age-specific hedonic wage regressions in the first stage and controls for cohort effects in the second stage. This analysis, based on eight years of pooled cross-sections, indicates how the value of life varies with an individual's age.

A. Data

To characterize the fatality risks faced by workers of different ages more precisely than is possible using average risk values by industry, we constructed a novel risk measure conditional upon age and the worker's industry rather than using an industry basis alone, which is the norm for all previous studies of age variations in workers' VSL. The source of the fatality measures is the Bureau of Labor Statistics (BLS) Census of Fatal Occupational Injuries (CFOI), for the 1992-2000 period. We structured the mortality risk cells by 2-digit SIC industries and these six age groups specified in the CFOI data: 16–19, 20–24, 25–34, 35–44, 45–54, and 55–64. To construct the denominator for the mortality risk variable, we used the 1992–2000 Current Population Survey Merged Outgoing Rotation Group files to estimate worker populations for each cell in the mortality data. The annual mortality risk measures are averaged to minimize any potential distortions associated with catastrophic mortality incidents in any one year and to have a better measure of the underlying risks for industry-age groups with infrequent deaths. Our injury risk measure, the probability of a lost-workday injury, also varies by age, and we constructed it in an identical manner for each 2-digit industry and for each of the age groups listed above. While

injury risk decreases with age across most industries, mortality risk increases monotonically with age in all industries, except for in mining.[10]

We have matched these constructed mortality risk and injury risk measures by age and industry with data on adult workers in the Current Population Survey Merged Outgoing Rotation Group data files for 1993–2000. We employed a number of screens in constructing our sample for analysis. The sample excludes agricultural workers and members of the armed forces. We have excluded workers younger than 18 and older than 62, those with less than a 9th grade education, workers with an effective hourly labor income less than the minimum wage, and less than full-time workers, which we defined as those working at least 35 hours per week.

B. Hedonic Wage Regression Framework

The standard hedonic wage model estimates the locus of tangencies between the market offer curve and workers' highest constant expected utility loci. The age variation in the wage-mortality risk tradeoff simultaneously reflects age-related differences in preferences as well as age-related differences in the market offer curve. If older workers are more likely to be seriously injured than are younger workers because of age-related differences in safety-related productivity, then the market offer curve will reflect that, given that age is a readily monitorable attribute. Because workers' constant expected utility loci and firms' offer curves each may vary with age, there is no single hedonic market equilibrium. Rather, workers of different ages will settle into distinct market equilibria as workers of different ages select points along the market opportunities locus that is pertinent to their age group.[11]

Conventional hedonic wage analyses of job risks specify the natural logarithm of the hourly wage or some comparable income measure as a function of worker and job characteristics, mortality risk, and, in more comprehensive specifications, injury risk and a measure of workers' compensation. Our base specification takes the following form:

$$(8) \qquad \ln(w_i) = \alpha + H_i'\beta + \gamma_1 p_i + \gamma_2 q_i + \gamma_3 q_i WC_i + \varepsilon_i,$$

where

---

[10] Refer to Aldy and Viscusi (2004) for more details about the construction of this age-specific job mortality risk measure.

[11] This analysis generalizes the hedonic model analysis for heterogeneous worker groups using the model developed for an evaluation of smokers and nonsmokers by Viscusi and Hersch (2001). Their worker groups differ in their safety-related productivity and in their attitudes toward risk.

w_i is the worker i's hourly after-tax wage rate,

H is a vector of personal characteristic variables for worker i,

p_i is the fatality risk associated with worker i's job,

q_i is the nonfatal injury risk associated with worker i's job,

WC_i is worker i's compensation replacement rate for a job injury, and

ε_i is the random error reflecting unmeasured factors influencing worker i's wage rate.

We calculated the workers' compensation replacement rate on an individual worker basis taking into account state differences in benefits and the favorable tax status of these benefits. We use the benefit formulas for temporary total disability, which comprise about three-fourths of all claims, and have formulas similar to those for permanent partial disability.[12] The terms α, β, γ_1, γ_2, and γ_3 represent parameters to be estimated.

All wage regression specifications used in this paper include the following controls: demographic indicator variables (race and ethnicity, gender of head of household, marital status, union membership, public sector employment, and resident of urban area); educational attainment; indicator variables for one-digit occupation and region of residence; and job mortality risk, job nonfatal injury risk, and expected workers' compensation replacement rate.[13]

The estimated regression then yields a measure of the average value of a statistical life for the sample:

$$(9) \qquad VSL = \hat{\gamma}_1 * \overline{w} * 2{,}000 * 100{,}000 .$$

This equation normalizes the VSL to an annual basis by the assumption of a 2,000-hour work-year and by accounting for the units of the mortality risk variable. As a preliminary check on our age-industry risk variables, we estimated equation (8) with the 1997 CPS MORG and compared this with the results for industry risk variables merged with the 1997 CPS MORG

---

[12] The procedures for calculating the workers' compensation benefit variable are discussed in more detail in Viscusi (2004), which also provides supporting references.

[13] The workers' compensation expected replacement rate represents the interaction of a worker's injury rate and that worker's estimated workers' compensation wage replacement rate based on the worker's wage, state of residence, state benefit formulas, and estimated state and federal tax rates. Given the endogeneity of the wage, we have also estimated instrumental variables regressions. IV estimation does not qualitatively influence determinations of coefficient magnitudes or statistical significance for the mortality risk variable of interest in this study. Refer to Aldy and Viscusi (2004) for additional details.

dataset presented in Viscusi (2004). We estimated a mean VSL of $4.5 million (1997$), which is virtually indistinguishable from the Viscusi (2004) estimate of $4.7 million, and both studies fall within the range of VSLs from hedonic wage regression studies of the U.S. labor market reported in Viscusi and Aldy (2003).[14]

C. Estimated Age Group VSLs

As an initial assessment of how the value of life varies with age across the population, we modified (8) so that the estimated compensating differentials can vary by age. We interacted five age group indicator variables – for age groups 18–24, 25–34, 35–44, 45–54, and 55–62 – with the various risk measures, and included the first four age group indicator variables in the specification:

$$(8a)\ \ln(w_i) = \alpha + H_i'\beta + \sum_{j=1}^{4}\delta_j age_j + \sum_{j=1}^{5}\gamma_{1j} age_j p_i + \sum_{j=1}^{5}\gamma_{2j} age_j q_i + \sum_{j=1}^{5}\gamma_{3j} age_j q_i WC_i + \varepsilon_i,$$

where age_j are the indicator variables for the five age groups and δ_j are parameters to be estimated.

We estimated this modified specification with eight annual CPS MORG samples from 1993–2000 and our industry by age job mortality risk and nonfatal injury risk data.[15] As distinct cross-section regressions, these specifications cannot discern age effects from cohort effects. They do, however, reveal how much an individual currently in one age group at a point in time is willing to pay for a given risk reduction vis-à-vis how much a different individual currently in another age group is willing to pay for such a risk reduction.

Table 1 presents the age-group specific results for this specification. We report two sets of standard errors: White heteroskedasticity-adjusted standard errors and robust and clustered standard errors that account for within-group correlations due to the assignment of the same job risk level to workers in an age-industry cell in each year.[16] The eight annual cross-section

---

[14] In our analysis with the 1997 CPS MORG, the mortality risk coefficient estimate is 0.0019 with a robust standard error of 0.00021.

[15] Note that we used averages of the lagged risk measures in these analyses. For example, the 1995 regression included risk measures averaged over 1992-1994 while the 2000 regression included risk measures averaged over 1992–1999.

[16] Refer to Hersch (1998) and Viscusi and Hersch (2001) as examples of papers in this literature that account for this type of correlation.

regressions reveal similar patterns of the VSL with respect to age: an inverted-U shape with the VSL peaking for the 35–44 age group in six of the eight years. As an illustration, consider the results for the year 2000 cross-section. The coefficient estimate on the 18–24 age group mortality risk variable is 0.0021, and it increases substantially to 0.0039 for the 25–34 age group. The mortality risk coefficient then declines with age: 0.0036 for the 35–44 age group, 0.0028 for the 45–54 age group, and 0.0014 for the 55–62 age group. The five age-group-specific job mortality risk coefficient estimates are individually statistically significant at the 1 percent or 5 percent level. The estimated VSLs for each age group depend on these coefficient estimates as well as age-group-specific average wages, which follow an inverted-U shape over the life cycle. The 35–44 age group has the largest VSL of $9.85 million, more than triple the 18–24 VSL of $3.16 million and nearly triple that of the 55–62 VSL of $3.77 million.[17]

To show how these differences in magnitudes are often statistically significant, we focus on the results for the year 2000 cross-section, which we report again at the top of Table 2. We conducted a series of pairwise Wald tests on the estimated VSLs, and the table presents the F-statistics associated with these tests. The first row of these tests shows that the 18–24 VSL of $3.16 million is statistically different from the VSL estimates for the next three age groups, but does not differ significantly from the 55–62 VSL of $3.77 million. The last column, corresponding to the 55–62 age group, shows that the estimated 55–62 VSL differs significantly from the VSL estimates for the 25–34 age group, the 35–44 age group, and the 45–54 age group. These results indicate that the VSL takes an inverted-U with respect to age across a population. The VSL pattern is relatively flat in the middle age groups as there is no statistically significant difference among the age 25–34, 35–44, and 45–54 categories for the 2000 cross-section.

D. Minimum Distance Estimator and Cohort Effects

We have extended this age-specific regression analysis in subsection C through a two-stage minimum distance estimator using VSL estimates for each year rather than age bands. This approach allows us to infer information about the VSL with respect to age based on a larger number of regressions based on more narrowly defined age bands for each year. While these individual regressions will provide less precise estimates of the compensating differential for risk

---

[17] Refer to Jones-Lee et al. (1985) for an example of a stated willingness to pay for safety study that also finds an inverted-U shaped VSL-age relationship.

than broader age groups, it will then be possible to estimate VSLs as a function of age if age-specific VSLs follow a systematic pattern over the life cycle.

In the first stage, we estimate age-specific hedonic wage regressions of the form expressed in equation 8 and use the mortality risk coefficient estimates to construct age-specific VSL. We estimated age-specific compensating differentials for 45 age levels from age 18 to 62 and eight cross-sections from 1993–2000, yielding 360 separate regressions. With the exception of the youngest and oldest birth-year cohorts, every cohort has eight observations in our constructed panel.[18] We estimated the VSL using the mean real wage for that respective age and year. Based on these first stage regressions, we construct a panel of cohort-specific and age-specific VSL estimates. Each VSL estimate is assigned to a birth-year cohort. For example, the estimated VSL for a 40-year old in 1993 is assigned to the 1953 birth-year cohort; the estimated VSL for a 41-year old in 1994 is also assigned to the 1953 birth-year cohort, and so on. We followed this procedure for all 360 VSL estimates.

In the second stage, we specify these VSLs by age. To characterize how the VSL estimates from the first stage, $\hat{VSL}$, vary with age across a population, the second stage includes a polynomial in age, $a(\theta)$. To characterize how the VSL varies over the life cycle, we account for the differences across cohorts by including a vector of birth-year indicator variables, c, in addition to the age polynomial. We also employ $\hat{V}$, the inverse of a diagonal matrix of the variance estimates of these VSLs, as a weight matrix based on Chamberlain's (1984) analysis of the minimum distance estimator and the choice of the inverse of the variance-covariance matrix as the optimal weight matrix.[19, 20]

---

[18] Refer to Deaton (1985) and Deaton and Paxson (1994) for the advantages of such a constructed panel based on birth-year cohorts.

[19] Because of the potential small sample bias in the optimal minimum distance estimator, we also evaluated the equally weighted minimum distance estimator (Altonji and Segal 1996). To address concerns about the small sample bias, we have presented the results for the equally weighted minimum distance estimator in Figures 1 and 2. The choice of weight matrix has no qualitative impact on our conclusions.

[20] We have employed a test of overidentifying restrictions to assess the appropriate order of the polynomial in age. If we assume that θ is a Kx1 vector, then a restricted parameter vector, α, which is Rx1 where R<K, can be estimated by some function, $b(\alpha)$. The following test statistic can then be used to evaluate the restrictions on the parameter vector:

$$N[\hat{VSL} - b(\hat{\alpha})]'\hat{V}^{-1}[\hat{VSL} - b(\hat{\alpha})] - N[\hat{VSL} - a(\hat{\theta})]'\hat{V}^{-1}[\hat{VSL} - a(\hat{\theta})] \sim \chi^2_{K-R}.$$

An analogous statistic was employed to evaluate the order of the age function in the cohort-based minimum distance estimator.

For the cross-sectional analysis, the minimum distance estimator solves:

(10) $$\min_{\theta \in \Theta}[\hat{VSL} - a(\theta)]'[\hat{V}]^{-1}[\hat{VSL} - a(\theta)].$$

For the life-cycle (cohort-adjusted) analysis, the minimum distance estimator solves:

(11) $$\min_{\theta \in \Theta, \delta \in \Delta}[\hat{VSL} - a(\theta) - c'\delta]'[\hat{V}]^{-1}[\hat{VSL} - a(\theta) - c'\delta].$$

where θ and δ represent parameters to be estimated. We specified $a(\theta)$ in a variety of analyses as a polynomial in age of order one to order eight.

The solid curve in Figure 1 presents the fitted age-VSL functions based on a third-order polynomial in age specification (cross-section VSL), while the dashed line presents the relationship based on a third-order polynomial in age with birth-year cohort indicator variables (cohort-adjusted VSL).[21] In the pooled cross-sections, the value of statistical life increases with age from age 18 with a VSL of $4.87 million through age 39, at which the VSL peaks at $8.27 million. The value of a statistical life then declines with age to a minimum of $1.67 million at the highest age in the sample, which is 62. The cohort-adjusted function, also yields a VSL that follows an inverted-U shape over the life cycle. It starts at $3.39 million at age 18, peaks at $7.79 million at age 46, and then declines to $5.09 million at age 62. Across the population and along the life cycle, the value of statistical life increases, peaks, and then decreases with age. While not presented, the birth-year indicator variables follow a general trend of increasing values with year of birth, consistent with the proposition that the value of life has increased with temporal increase in lifetime income.

The cohort adjustment affects the age-related pattern of VSLs in several ways. The peak of the age-VSL curve is seven years later when accounting for date of birth. The high VSLs for younger age groups is due in part to their higher lifetime wealth, as their cross-section VSLs lie above those in the cohort-adjusted values. For older age groups the pattern is reversed. While there is a steep drop in VSL levels with age in the cross-section results, this decline is due in part to cohort effects. Accounting for cohort differences attributable to changes in lifetime income more than doubles the estimated VSLs for the older age groups and flattens their VSL trajectory. Finally, the counter-clockwise pivoting of the VSL function from the cross-sectional analysis to

---

[21] Based on the specification test presented in footnote 17, we could not reject the hypothesis that a third-order age polynomial fit the data as well as higher-ordered polynomials. We could, however, reject the hypothesis that lower-ordered polynomials fit the data as well as a third-order polynomial.

the cohort-adjusted analysis also illustrates the importance of accounting for lifetime income, implicitly through the birth-year indicator variables, in estimating the age-VSL relationship over the life cycle.

We also tested two economic propositions that are prominent in current policy applications of the value of life. First, many analyses assume that the VSL remains constant, irrespective of age.[22] To assess this proposition, we employed our cohort-based minimum distance estimator and specified the age polynomial function as a constant. We then tested this restriction versus the more flexible, higher-ordered polynomials and we reject the hypothesis that the VSL is constant over the workers' life cycle at the 1 percent level in comparison with all age polynomials of order two or higher. Second, other analyses have assumed that the value of a statistical life is always decreasing with age.[23] To test this proposition, we specified the age polynomial function as linear, but such an approach yielded a negative coefficient estimate that clearly could not be distinguished from zero. The test of overidentifying restrictions rejected the linear specification in comparison to all higher-ordered polynomials. It should also be noted that all order two through order eight polynomials resulted in similar inverted U-shaped relationships between the value of a statistical life and age.[24]

IV. Implications for the Value of a Statistical Life-Year

The preceding section illustrates the estimated age-VSL profile consistent with the theory model presented in Section I and with previous simulations published in the literature. The implicit assumptions underlying the value of a statistical life-year (VSLY) approach, which requires the value of life to be decreasing with age at all ages, are rejected by our data. In light of the common application of VSLYs in evaluations of medical interventions and government regulations, such as those promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency in their sensitivity analyses, we have estimated age-specific VSLYs based on our age-specific VSLs.

---

[22] For example, most U.S. Environmental Protection Agency benefit-cost analyses, including September 2003 revisions to its assessment of the Clear Skies initiative, make this assumption.

[23] For example, this is consistent with the European Commission's proposed position and the life-year approach used by the U.S. Food and Drug Administration.

[24] We also evaluated whether the higher VSLs for individuals in the 25-44 age range reflect major life-cycle events such as marriage or having children, and not variations in age, but find no evidence to support this notion. Refer to Aldy and Viscusi (2004) for more details.

To construct values of statistical life-years, we have annuitized age-specific VSLs based on age-specific years of life expectancy L and an assumed discount rate r of 3 percent:[25]

(12) $$VSLY = \frac{rVSL}{1-(1+r)^{-L}}\ .$$

Figure 2 presents these calculations for the cross-section and cohort-adjusted VSLs derived from the minimum distance estimator. The average VSLY is $296,000 for the cross-section and $302,000 for the cohort-adjusted estimates. VSLYs follow a similar inverted U-shaped relationship over the life cycle as depicted for VSL. The increase in VSLY is clearly expected for young workers because VSL is increasing and life expectancy is decreasing. The monotonic decrease in VSLY after its peak indicates that age-specific VSLs are decreasing at a faster rate than life expectancy. The peak in the VSLY occurs at a higher value and at a much higher age for the cohort-adjusted measure. It peaks at a value of $401,000 at age 54 for the cohort-adjusted measure, as compared to a peak of $375,000 at age 45 for the cross-section measure. The cohort-adjusted VSLY declines at a much slower rate than the VSLY after the peak for the cross-section measure. The influence of cohort adjustments has an even greater relative effect on the VSLY levels for the older workers in the sample than they did on VSL. Interestingly, the VSLY for those age 62 is higher than for all age 39 or younger.

V. Conclusion

The implications of wage-risk tradeoffs for the dependency of VSL on age is consistent based on both age group-specific estimated VSLs and a minimum distance estimator derived from age-specific VSLs. We find that the VSL rises and then falls with age across the population and over the life cycle, displaying an inverted U-shaped relationship. The minimum distance estimator results are perhaps most instructive, as they can more flexibly represent the age relationship while controlling for cohort effects. Failing to account for the secular increase in incomes with birth-year indicator variables yields much lower VSLs for older individuals and higher VSLs for younger individuals in cross-section analysis. Including cohort effects results in a much flatter age-VSL function over the life cycle, and older individuals have a higher value of a statistical life.

---

[25] We have also calculated VSLYs based on a 7 percent discount rate (the current preferred rate by the U.S. Office of Management and Budget for evaluating government regulations). The higher discount rate yields larger VSLYs and a more pronounced inverted U-shaped age-VSLY relationship.

The result that the VSL rises and falls with age is of both theoretical and policy interest. Theoretical analysis of VSL over the life cycle suggests such a relationship may exist, particularly in situations in which there are insurance and capital market imperfections. The results are supportive of these models rather than those that generate steadily declining VSL with age, such as some models with perfect annuity and insurance markets. VSL is not steadily declining with age even though the amount of expected lifetime at stake steadily declines with age. As the life-cycle models indicate, this result is not surprising since the age-VSL linkage depends on factors such as the life-cycle consumption pattern, which also displays a similar age structure.

These estimates may help inform policymakers as they consider policies that would simultaneously reduce mortality risk for individuals of various ages. In terms of the appropriate "senior discount," in the cross-section analysis workers in their early 60s have a VSL of about $1.7–$2.0 million, which is between one-fifth and one-fourth the size of the VSLs for prime-aged workers. Understanding how the value of statistical life varies over the life cycle can inform policymakers as they consider government interventions that would reduce mortality risks posed to individuals over multiple stages of their life. The cohort-adjusted VSL levels for older workers are much higher than in the cross-section analysis, with a VSL of about $5 million for workers in their early 60s. While below the peak VSL over the life cycle, these older workers' VSLs are above the VSLs for very young workers. This analysis does not provide support for approaches that focus only on the remaining quantity of life as the valued attribute. Both the value per life-year approach and the quality-adjusted life year methodology yield a steadily decreasing VSL with age, whereas the revealed preferences of workers' risk decisions indicate a quite different relationship that rises and then declines with age. Explicit construction of age-specific values of statistical life-years from our age-VSL profiles show that the value of a statistical life-year varies with age. Likewise, there is no support for the standard practice of transferring VSLs from studies based on the average of the labor market to risk contexts specific to the elderly population. Individuals make decisions over risk and income that clearly indicates that the value of their life varies with age, but the relationship is not a simple one.

References

Aldy, J.E., and W.K. Viscusi. 2004. Age Variations in Workers' Value of Statistical Life. Olin Center for Law, Economics, and Business Discussion Paper no. 392. Cambridge, MA: Harvard Law School.

Altonji, J.G., and L.M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business and Economic Statistics* 14(3): 353–366.

Anderson, R.N. 1998. United States Abridged Life Tables, 1996. *National Vital Statistics Reports* 47(13).

Arthur, W.B. 1981. "The Economics of Risks to Life." *American Economic Review* 71(1): 54–64.

Chamberlain, G. 1984. "Panel Data." In: Z. Griliches and M.D. Intriligator, (eds.), *Handbook of Econometrics*, Volume II, pages 1247–1318.

Deaton, A. 1985. Panel Data from Time Series of Cross-Sections. *Journal of Econometrics* 30: 109–126.

Deaton, A., and C. Paxson. 1994. Intertemporal Choice and Inequality. *Journal of Political Economy* 102(3): 437–467.

Cropper, M.L. and F.G. Sussman. 1988. "Families and the Economics of Risks to Life." *American Economic Review* 78(1): 255–260.

Ehrlich, I., and Y. Yin. 2004. "Explaining Diversities in Age-Specific Life Expectancies and Values of Live Saving: A Numerical Analysis." NBER Working Paper 10759.

European Commission. 2001. "Recommended Interim Values for the Value of Preventing a Fatality in DG Environment Cost Benefit Analysis." Internet: http://europa.eu.int/comm/environment/enveco/others/recommended_interim_values.pdf

Hara Associates Inc. 2000. *Benefit/Cost Analysis of Proposed Tobacco Products Information Regulations*. Prepared for Health Canada and Consulting and Audit Canada. Ottawa, Ontario. June 5, 2000.

Hersch, J. 1998. "Compensating Differentials for Gender-Specific Job Injury Risks." *American Economic Review* 88(3): 598–627.

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

Johansson, P.-O. 2002. "On the Definition and Age-Dependency of the Value of a Statistical Life." *Journal of Risk and Uncertainty* 25(3): 251–263.

——— 1996. "On the Value of Changes in Life Expectancy." *Journal of Health Economics* 15: 105–113.

Jones-Lee, M.W. 1989. *The Economics of Safety and Physical Risk*. Oxford: Basil Blackwell.

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

Rosen, S. 1988. "The Value of Changes in Life Expectancy." *Journal of Risk and Uncertainty* 1: 285–304.

Shepard, D.S., and R.J. Zeckhauser. 1984. "Survival Versus Consumption." *Management Science* 30(4): 423–439.

Smith, V.K., M.F. Evans, H. Kim, and D.H. Taylor. 2004. "Do the Near-Elderly Value Mortality Risks Differently?" *Review of Economics and Statistics* 86(1): 423–429.

Viscusi, W.K. 2004. The Value of Life: Estimates with Risks by Occupation and Industry. *Economic Inquiry* 42(1): 29–48.

Viscusi, W.K., and J.E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World." *Journal of Risk and Uncertainty* 27(1): 5–76.

Viscusi, W.K., and J. Hersch. 2001. "Cigarette Smokers as Job Risk Takers." *Review of Economics and Statistics* 83 (2): 269–280.

Tables and Figures

Table 1. Age Group-Specific Values of a Statistical Life, Annual Cross-Sections, 1993-2000[a]

Year		18-24 Age Group	25-34 Age Group	35-44 Age Group	45-54 Age Group	55-62 Age Group
1993	Mortality Risk	0.00040 (0.00045) [0.00046]	0.00434 (0.00040)*** [0.00077]***	0.00308 (0.00041)*** [0.00084]***	0.000728 (0.00041)* [0.00068]	0.00089 (0.00066) [0.00087]
	Mean VSL	$0.64	$9.92	$8.36	$2.04	$2.36
1994	Mortality Risk	0.00238 (0.00047)*** [0.00064]***	0.00329 (0.00038)*** [0.00064]***	0.00277 (0.00038)*** [0.00078]***	0.00132 (0.00043)*** [0.00078]*	0.00176 (0.00065)*** [0.00083]**
	Mean VSL	$3.97	$7.73	$7.75	$3.86	$4.87
1995	Mortality Risk	0.00298 (0.00051)*** [0.00064]***	0.00313 (0.00039)*** [0.00063]***	0.00223 (0.00039)*** [0.00079]***	0.00174 (0.00042)*** [0.00078]**	0.00162 (0.00059)*** [0.00080]**
	Mean VSL	$4.87	$7.31	$6.16	$5.02	$4.46
1996	Mortality Risk	0.00319 (0.00077)*** [0.00089]***	0.00350 (0.00043)*** [0.00069]***	0.00310 (0.00044)*** [0.00084]***	0.00163 (0.00043)*** [0.00070]**	0.00124 (0.00056)** [0.00069]*
	Mean VSL	$5.13	$8.08	$8.45	$4.67	$3.39
1997	Mortality Risk	0.00288 (0.00058)*** [0.00075]***	0.00348 (0.00043)*** [0.00071]***	0.00329 (0.00043)*** [0.00076]***	0.00196 (0.00043)*** [0.00073]***	0.00162 (0.00061)*** [0.00080]**
	Mean VSL	$4.60	$8.08	$8.98	$5.64	$4.47
1998	Mortality Risk	0.00346 (0.00064)*** [0.00086]***	0.00283 (0.00045)*** [0.00068]***	0.00305 (0.00044)*** [0.00076]***	0.00159 (0.00045)*** [0.00072]**	0.00158 (0.00058)*** [0.00078]***
	Mean VSL	$5.65	$6.76	$8.61	$4.69	$4.55
1999	Mortality Risk	0.00154 (0.00052)*** [0.00059]***	0.00359 (0.00050)*** [0.00071]***	0.00355 (0.00050)*** [0.00089]***	0.00337 (0.00048)*** [0.00082]***	0.00162 (0.00063)*** [0.00086]*
	Mean VSL	$2.18	$7.18	$8.41	$8.35	$3.95
2000	Mortality Risk	0.00211 (0.00060)*** [0.00073]***	0.00391 (0.00049)*** [0.00074]***	0.00356 (0.00047)*** [0.00088]***	0.00277 (0.00046)*** [0.00074]***	0.00135 (0.00059)** [0.00086]
	Mean VSL	$3.16	$9.03	$9.85	$7.97	$3.77

[a] VSLs are expressed in millions of year 2000 dollars based on age-specific wages. Dependent Variable: natural logarithm of hourly labor income. Each specification includes 9 1-digit occupation indicator variables, 8 regional indicator variables, demographic variables, nonfatal injury risk, and expected workers' compensation replacement rate. Robust (White) standard errors are presented in parentheses, and standard errors accounting for within-group correlation are presented in brackets. ***, **, * Indicates statistical significance at 1 percent, 5 percent, and 10 percent levels, two-tailed test.

Table 2. Age Group-Specific Values of a Statistical Life, 2000[a]

	18-24 Age Group	25-34 Age Group	35-44 Age Group	45-54 Age Group	55-62 Age Group
Mortality Risk	0.00211 (0.00060)*** [0.00073]***	0.00391 (0.00049)*** [0.00074]***	0.00356 (0.00047)*** [0.00088]***	0.00277 (0.00046)*** [0.00074]***	0.00135 (0.00059)** [0.00086]
Mean Age Group VSL (millions 2000$)	$3.16	$9.03	$9.85	$7.97	$3.77

Age Group	H_0: Pairwise Tests of Equality of VSL Estimates, F-Statistics, $F(1, 118{,}639)$				
18-24	-	16.16	17.52	8.89	0.10
25-34	-	-	0.22	0.36	6.94
35-44	-	-	-	1.01	8.39
45-54	-	-	-	-	3.97

[a] $N = 118{,}762$. $R^2 = 0.56$. Dependent Variable: natural logarithm of hourly labor income.   Specification includes 9 1-digit occupation indicator variables, 8 regional indicator variables, demographic variables, nonfatal injury risk, and workers' compensation expected replacement rate. Robust (White) standard errors are presented in parentheses and standard errors accounting for within-group correlation are presented in brackets. ***, ** Indicates statistical significance at 1 percent, and 5 percent levels, two-tailed test.

Figure 1. Cohort-Adjusted and Cross-Section Value of Statistical Life, 1993–2000

VSL (millions 2000$)
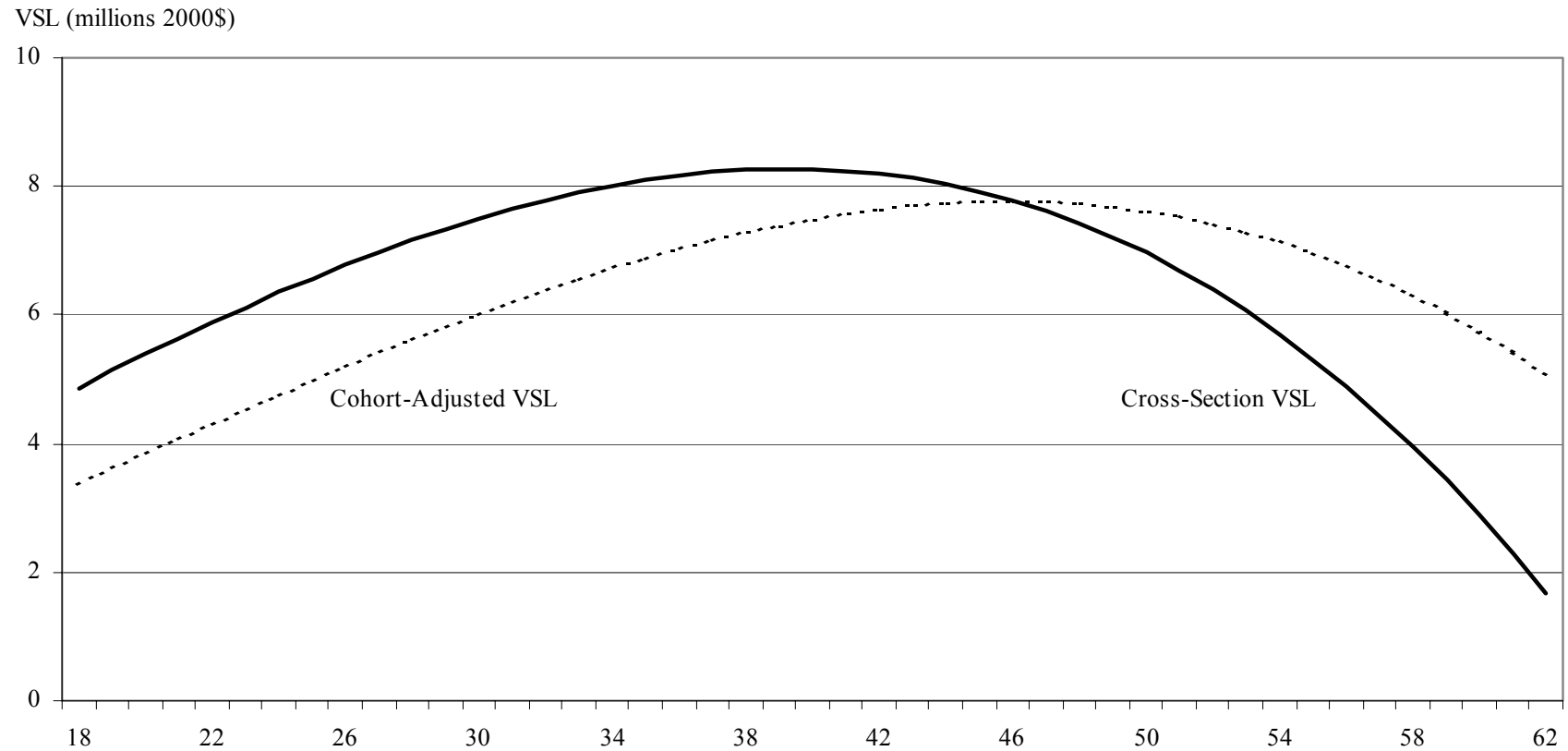
NOTES: Both series are based on equally weighted minimum distance estimator with a third-order polynomial in age. The cohort-adjusted VSL also includes indicator variables for year of birth.

Figure 2. Value of a Statistical Life-Year Based on Cohort-Adjusted and Cross-Section Value of Statistical Life, 1993–2000

VSLY (2000$)
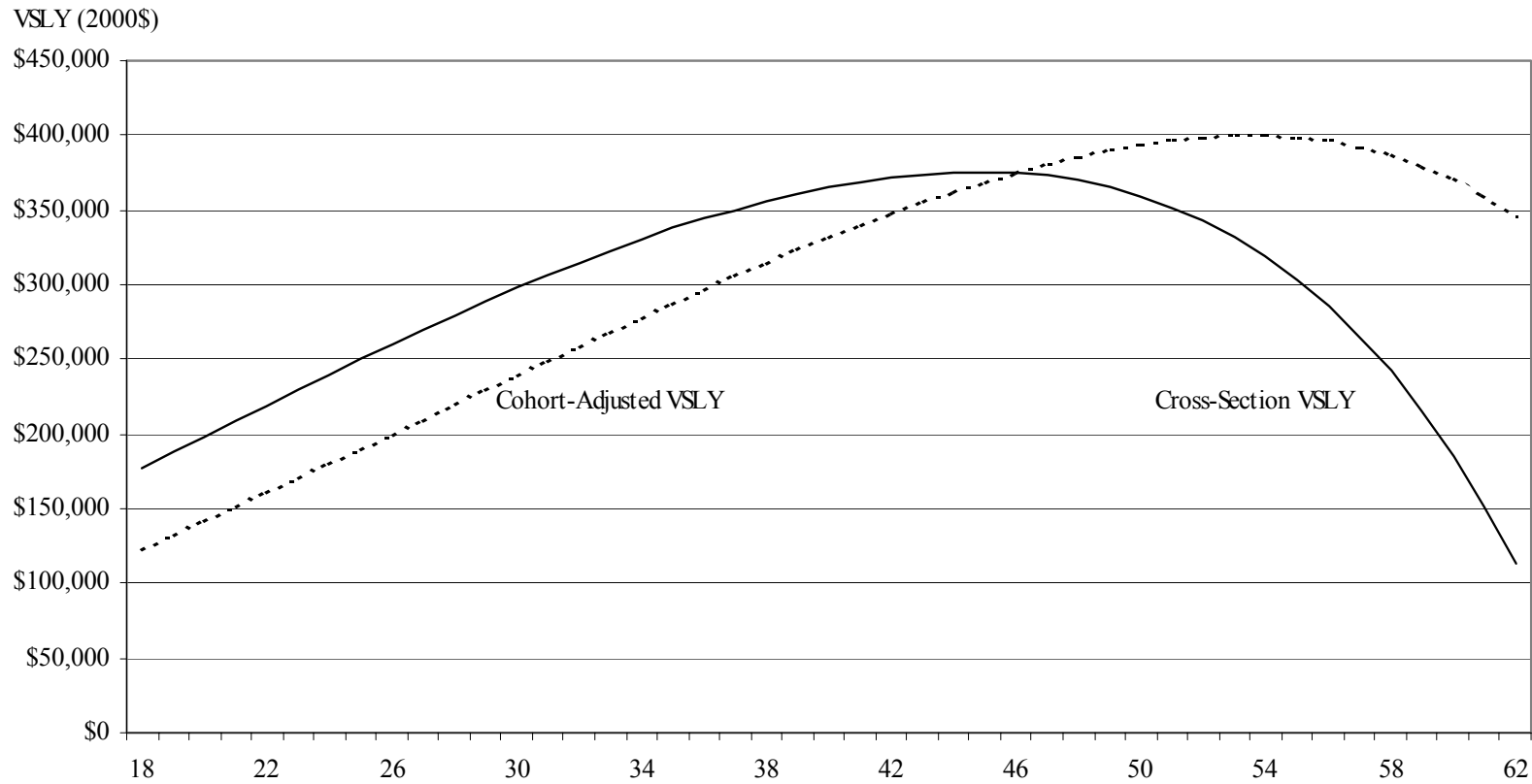
NOTES: Value of statistical life-years based on an assumed 3 percent discount rate and average age-specific life expectancy and derived from the age-specific VSLs presented in Figure 1.

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

Empirical Issues Associated with Mortality Risk Valuation

Joseph Aldy and Kip Viscusi, "Adjusting the Value of a Statistical Life for Age

and Cohort Effects"

George Van Houtven, Melonie Sullivan, and Chris Dockins, "Eliciting Tradeoffs

for Valuing Fatal Cancer Risks"

Comments by Clark Nardinelli

April 11, 2006

     The two papers in this session both explore differences in people's willingness to

pay for a small reduction in the risk of death, or what economists call the value of a

statistical life. Joseph Aldy and Kip Viscusi look at differences in the values of statistical

life across workers by ages, whereas George Van Houtven, Melonie Sullivan, and Chris

Dockins look at the difference between the value of reducing fatal cancer risks and the

value of reducing fatal accident risks. Because of the way I approached the papers, I will

first discuss Aldy and Viscusi.

Aldy and Viscusi's work on the compensating differentials associated with occupational risks forms the basis for the most widely-used estimates of the value of a statistical life. Most regulatory economists use their estimates and will continue to do so. Their new study is especially welcome because it deals with the highly controversial topic of the so-called "senior discount", or more generally the effects of age on the value of a statistical life. The recent dust-up over regulatory analyses that used a lower value of statistical life for older persons, ensures that economists who analyze regulatory policies will pay close attention to the results of this study.

Aldy and Viscusi introduce age-specific fatal and non-fatal occupational risks to estimate the relationship between age and the value of a statistical life. The cross-sectional hedonic regression based on age-specific fatal accident risks generates an inverted U-shaped curve. The curve shows a steep decline from the peak value of $9.9 million per statistical life at age 39 to $3.8 million at age 62. Much of that decline, however, apparently reflects the rise in real incomes over time. Older workers come from generations with lower real earnings and therefore have lower values of statistical life. Aldy and Viscusi show that the steep decline after age 39 comes from the combination of the changes over time within an age cohort and the differences in lifetime earnings across cohorts. Including a cohort adjustment in the regressions substantially flattens the slope of the inverted U, particularly at older ages. In the cohort-adjusted estimates, the value peaks at $7.8 million at age 46, and then declines to $5.1 million at age 62. Indeed, the difference between maximum value and the value at age 62 in the cohort-adjusted estimates is not large enough to be raise serious doubts in someone who believes that the value of statistical life does not vary substantially across age groups. I would like to see if

2

other adjustments would flatten the curve even more. Knowing what those adjustments were might help illuminate the difference between these results and the stated preference experiments that generate a constant value of a statistical life across ages.

As I read this paper and looked at the inverted U-shaped curve, I thought the curve looked familiar. I thought about it for a while and finally realized that the value of statistical life curve generated by their hedonic wage regressions has the same basic shape as a traditional lifetime age-earnings or age-productivity curve. Because they have the data on age and wage rates, I would like to see Aldy and Viscusi compare the age-earnings curve with the age-value of statistical life curve. Do they have similar or different shapes? Do they peak at the same age as earning or at a different age? By doing the comparison, they can show how adding age-specific fatal accident rates to the estimating equations alters the shape of the values of statistical life from that implied by the age-earnings profile alone.

As a regulatory economist I was asked to consider the implications of these results for my work. My first reaction was mild elation because we now have some empirical estimates on how the value of statistical life changes with age. The Aldy-Viscusi estimates cover ages 18 through 62 and most of the illnesses dealt with by my agency affect people younger than 18 and older than 62, but I thought that perhaps we could extend the estimating polynomial equations backward and forward out of sample to generates estimates for our analyses. But then I immediately got discouraged as I read Aldy and Viscusi's warning that "there is no support for the standard practice of transferring VSLs from studies based on the average of the labor market to risk contexts specific to the elderly population."

Feeling a little depressed because Aldy and Viscusi would not let me use their results, I turned to the paper by Van Houtven, Sullivan, and Dockins. I immediately felt better because Van Houtven, Sullivan, and Dockins promised to show me how to use a stated preference survey to derive a relationship between the willingness to pay to reduce the risk of fatal cancer and the willingness to pay to reduce the risk of a fatal automobile accident.

The paper's literature review surprised me. I did not know that some economists have failed to find a premium when estimating willingness to pay to reduce cancer risks compared with other risks. I am more familiar with the risk analysis research on cancer risks; that literature always finds a cancer effect – including more dread, higher aversion, steeper trade-offs, and skewed risk rankings when cancer is one of the risks. In studies of risk perception, cancer always generates a response that cannot be explained by the actuarial risk. Based on risk analysis literature, I did not think that the direction of the cancer effect was in doubt, just the size.

Van Houtven, Sullivan, and Dockins indeed find a cancer premium, which they model as the ratio of the value of a statistical cancer to the value of a statistical life as estimated by the willingness to pay to reduce the risk of a fatal automobile accident. The contrast between the near-instant death through automobile accident and the prolonged illness and other unpleasantness that accompanies death from cancer allows them to identify the full difference in willingness to pay as a cancer premium. They also show how latency (death from cancer occurs years after exposure) and the types of cancer influence the value of a statistical cancer.

Van Houtven, Sullivan, and Dockins extract the value of a statistical cancer and the value of a statistical life from the results of a probit analysis of stated preferences between two hypothetical locations with different risk characteristics. The survey respondents chose between moving to an area with fewer automobile accident deaths per million than their current location and moving to an area with fewer cancer deaths per million than their current location. Using a dichotomous choice to elicit preferences and then deriving the willingness to pay from the results makes fewer demands on participants than many stated preference methods but still gives the full continuous range of results. I like the method and was pleased to hear that the Office of Management and Budget has approved it in principle.

In the description of the participants, however, I noticed something strange in the reported personal experiences of the survey participants. Among the participants, 18 percent reported knowing someone who had died in an automobile accident and 12 percent reported knowing someone who had died of cancer. But the number of annual cancer deaths is 13 times larger than the number of annual automobile accident deaths. Unless people who die of cancer die without friends or relatives, there's something not quite right here. My guess is that the sample, drawn from a web-based national panel, is not in fact made up of persons with unusual life experiences. Instead, the result tells us something about the perceptions and recall biases associated with the risks of fatal cancer and fatal automobile accidents. I suggest that Van Houtven, Sullivan, and Dockins explore this result in later research. It may help them to explain more of the cancer premium itself.

Van Houtven, Sullivan, and Dockins find that, compared with the base case of a fatal automobile accident, the value of a statistical cancer ranges from 2 to 3 times the value of a statistical life, depending on the latency period – a large premium by any yardstick. This finding alone makes the study of value to those of us who must assess the benefits of policies designed to reduce the risks of cancer and other illnesses.

For policy analysis, however, it is important to identify the source of the cancer premium. Van Houtven, Sullivan, and Dockins do not find a strong effect from the period of morbidity that precedes death from cancer. This negative finding may reflect the difficulty of teasing out an estimate of morbidity's contribution to the cancer premium. For policy analysts, the more disturbing possibility is that the cancer premium is not due to morbidity but to something else. Does the cancer premium reflect some fear based not on actual but on imagined outcomes or superstition? The responses of people with some experience of cancer imply that much of the fear associated with cancer may stem from unfamiliarity with the illness.

We need some way to separate the real from the imaginary parts of the cancer premium. Morbidity must account for the real part of the premium. Morbidity can cause real losses above the value of a statistical death directly through the effects of illness on victims and perhaps indirectly through its effects on friends and relatives. Whatever generates the large cancer premium found here has to be related to something that happens before death, during the morbidity phase of the cancer. Once death occurs, the cause ceases to matter.

Many researchers apparently believe that we should treat different causes of death differently based solely on differences in willingness to pay, a practice difficult to justify

6

for regulatory analysis. Suppose that we could survey a representative sample of dead people. If we asked them whether they found it worse to be dead because of cancer or worse to be dead because of an automobile accident, I doubt that we would find a significant difference between the cancer dead and the accident dead. Unless we have some survey evidence from dead persons saying that yes, it is far worse to be dead because of cancer than anything else, we have to find something that occurs before death that makes cancer worse than a fatal accident. The cancer premium derived from stated preference alone does not justify regulatory analysts placing a higher value of preventing statistical cancers.

To make these results of practical use for policy analysis, Van Houtven, Sullivan, and Dockins need to tease the mortality and morbidity effects out of the value of statistical cancers. Doing so would make it possible to de-compose the cancer premium into a realized utility loss and the superstition, stigma, or dread that generates the rest of the premium. I am not suggesting that superstition plays no role in real-world market valuation, only that it should play no role in the valuations used by public health agencies. Public health agencies exist partly because the general public does not always have adequate information on true actuarial probabilities and true severities. In valuing policy alternatives, regulators should ignore superstition, stigma, and irrational fear. As public health economists, we should measure human welfare with actuarially correct risks and real measures of severity, not with dread, superstition, or other imaginary effects. The Food and Drug Administration, for example, came into existence to reduce the consumption of snake oil. To assess the effects of that agency's regulations based on imaginary effects would be the equivalent of introducing snake oil into regulatory

analyses. Public health agencies should stick to concrete, measurable health effects when assessing regulatory policies.

Let me conclude by saying that I am grateful for the opportunity to read and comment on these two fascinating papers.

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

Comments on VSL Papers

Maureen L. Cropper
University of Maryland and World Bank

April 11, 2006

1

Eliciting Risk Tradeoffs for Valuing Fatal Cancer Risks

- **Why This Paper is Important:**

- **SAB's Environmental Economics Advisory Committee determined in 2000 no valid estimates exist of the value of a fatal cancer case**

- **This paper fills this void by estimating the ratio of the value of a statistical fatal cancer (VSC) to the VSL associated with immediate, accidental death in an auto accident (VSL)**

$$\frac{VSC}{VSL} = \frac{-dP_D}{dP_C}\bigg|_{E(U)=k}$$

2

Answers Seem Generally Reasonable

- **78% of respondents pass probability choice quizzes**
- **In choice between City B (fewer Cancer deaths) and City A (fewer Auto deaths), proportion choosing City B FALLS as the relative death ratio (RDR)—the number of auto deaths saved for every additional cancer death—rises**
- **Latency reduces proportion choosing City B, holding the RDR constant**

  BUT:

- Percent choosing City B never falls below 50% even with 25 year latency
- No sensitivity to length of morbidity preceding death

3

Effect of Risk Difference Ratio on P(Choose City B)
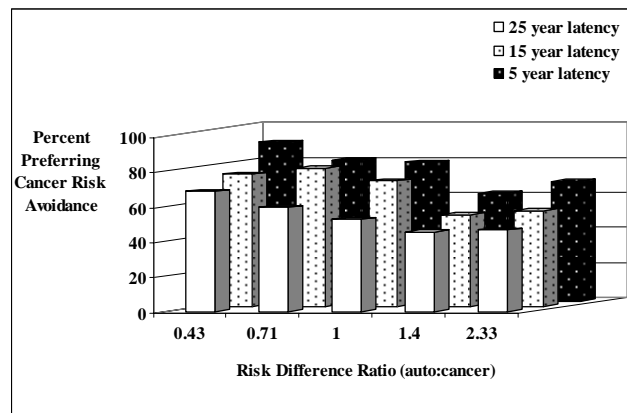
4

How to Value Non-Fatal Cancers?

- Risk-risk tradeoffs for non-fatal cancers effectively produce a QALY weight for cancer:

$$\frac{VSC}{VSL} = \left(1 - \frac{U(C,Y)}{U(H,Y)}\right)$$

- How does this compare with other elicitation methods for obtaining QALY weights?
- Does this effectively monetize QALYs?

5

Adjusting the Value of a Statistical Life for Age and Cohort Effects

- Why this paper is important:
  - Examines how hedonic wage function shifts in wage-risk space with worker age
  - Estimates how MWTP for a change in risk changes with age
  - Uses multiple cross-sections to disentangle age and cohort effects

- However, to use these estimates for policy, one must believe that hedonic wage equations provide unbiased estimates of a change in risk on the wage.

6

Should Hedonic Wage Equations Be Interpreted as Causal Relationships?

- How do we interpret a cross-sectional regression of infant mortality on air pollution levels?

- How do we interpret a regression of property values on air pollution levels using a single cross section of data?

- Due to omitted variable bias problems both results would be suspect: Need to find a natural experiment that causes an exogenous change in air quality (see e.g. Chay and Greenstone, QJE, August 2003; JPE, April 2005).

- Dan Black et al. (2003) raise similar concerns about hedonic wage equations: risk is likely to be correlated with the error term, causing results to be suspect

- Perhaps a natural experiment involving changes in road safety could be used to measure the impact of changes in fatal risk on wages of transport operators.

7

Summary of the Q&A Discussion Following Session V

J.R. DeShazo, (UCLA)
Directing his comments to Dr. Joe Aldy, Dr. DeShazo stated, "We do have some stated
preference estimates that reflect almost exactly the same pattern of age-adjusted VSLs
that you revealed—they're slightly lower, on average." Referring to comments made by
discussant Clark Nardinelli, he added that because he and his colleagues had a large
number of seniors in their sample, these age-adjusted VSLs "actually can be used to
evaluate the senior population." He also stated that "the general pattern is very much the
same, and I think the only difference is that our VSL estimate is about $2 million less, on
average."

Saying that he had two questions, Dr. DeShazo posed the first: "Why do we see this
change with age?" He cited Ehrlich's work as "the best theoretical study to date." He
said that Ehrlich "shows that there are a variety of reasons that we might value reductions
in health risks as we age. First and foremost, we value health in the current period, and
as we age the marginal utility of that is going to increase as our health state declines."
Other factors Dr. DeShazo identified include "changes in the remaining expected
lifespan—changes in the marginal utility of income or consumption—changes in
individuals' discount rates that you would, consistent with theory, expect to increase. Of
course, for any given risk that you're focusing on, the background risk profile is
changing, so that the other risks that you face are going up." Addressing the presenters,
he asked, "Out of those things that vary with age, what do you think explains this
decline?" He added that in their analysis he and his colleagues "were able to identify
changes in the marginal utility of consumption, changes in the discount rate, and changes
in the background risks that people face as they age, all of which might explain this
decline."

The second question from Dr. DeShazo, which he classified as "much more fundamental"
than the first, was "whether or not we should be applying hedonic estimates which give
us current period values for a mortality risk reduction in efforts to value the types of
health risk reductions that FDA and EPA focus on primarily, which follow, typically,
years of chronic or severe morbidity. The basic question is: Is the marginal mortality
risk reduction the same today, if you're perfectly healthy, as it would be if you've
suffered for 10 years from chronic morbidity or maybe 3 years from severe morbidity?"
He added that preliminary results from his studies "suggest that that's not the case—that
the marginal value of a risk reduction is highly context dependent, and your willingness
to trade off morbidity and mortality health states is such that your value of mortality
reductions falls as you experience more prior morbidity." In closing, Dr. DeShazo asked,
"So, what's the best argument for transferring the sort of hedonic wage analysis?"

Joe Aldy, (Resources for the Future)
"To get to the first question about why we think we see this kind of age profile with
respect to the value of a statistical life—I think when we look at the young workers that
there are two things driving that result. One is that for the 18- to 24-year-olds we're

focusing on just the full-time workers. That group is going to have a disproportionally large sample of full-time workers who do not have a college degree relative to older age groups in our sample. To the extent that those are people who never have a college degree, they have lower lifetime income, and we would expect them to have a lower risk-income tradeoff in the labor market. I think the other thing that's driving that is that you can't really borrow against future income. With the exception of college loans, it's really hard to be able to do that. This is why we see very little savings behavior among most households until they get into their late 30's or early 40's. You see a little bit of precautionary savings, but other than that, very little saving occurs early in life, and I think that's one reason we see the lower value for the younger workers." He added that the impact seen with older workers is, he feels, "being dominated by life expectancy." He also said he felt it would be great to get a sense of what's driving "stated preference results that show relatively small declines or no decline in the value of life," and he asked "is it because we see changes in discount rates as they get older, and we see differences in risk attitudes? Are there ways in which we can try to structure future surveys to try to get at those questions explicitly so we can have a better understanding of why we see that impact?" Citing existing literature and specifically naming Zeckhauser, Rosen, Ehrlich (two papers), and Johannson (several papers), Dr. Aldy stated that most of this work shows that the value of life is going to decline as people get older. He added, however, that almost all of these studies "have assumed that attitudes toward risk are constant across the lifecycle and the discount rate is constant over the lifecycle. Also, with the exception of the Ehrlich stuff, health doesn't really enter in at all. As we get more complex, there's the question: Are we able with a richer model and a richer theory to explain the fact that the value of life may not decline much with age? With what we're finding among the workers, though, we don't have any basis for saying it's because of health—their health can't be changing that much because they're still full-time workers. It might be declining some, but I think at the end of the day it's life expectancy that's really going to be driving the results on that."

Dr. Aldy then turned to the other question of why we should be using hedonic wage VSLs when we look at policies that have latent impacts. He commented that "part of the point we made at the end of our paper is that if most of those you are looking at are elderly who are enjoying the benefits—whether it's clear skies, whether it's the tier-2 rule that reduced sulfur in gasoline and had a lot of PM reduction benefits—in that context it's difficult to reconcile a VSL where the average age of the worker is 35-40 years old. You bring up the latency issue; I think the age issue makes it difficult." He went on to explain, "There's a question of whether or not in ours you say: Well, you come up with this value of life for someone who is 62—that still doesn't really help me much if the average age of the beneficiary is 70. I haven't even really gotten to that person yet, and if you think, for example, there may be differences in attitudes after one leaves the workforce. One can come up with plausible stories about how one's attitudes toward risk would vary from what they were before and raise questions about whether or not one should be trying to transfer a hedonic wage estimate over. So, I'm not going to come out and forcefully say that hedonic wage is the way to go. As I already mentioned, I have a personal bias towards revealed preference. My co-author may have a slight bias towards that, but he's done a good number of CV studies, too."

He closed by saying, "I think we're getting closer to doing a better job in the hedonic wage literature of trying to get to interesting questions. For about 15 years, there wasn't much that was very interesting going on. I think if you look at the last 5 years, it's getting more interesting to try to better understand the heterogeneity in the value of life. I think that's something that if you work through theory models, you start seeing that there should be a lot of heterogeneity in the value of life, not just with respect to age but to other issues and attributes as well. We're moving in that direction, but I'm not standing up here and saying you should ignore all the stated preference stuff and go with hedonic wage, because I don't think we're really cracking that nut yet."

_____

Bryan Hubbell (U.S. EPA)
Dr. Hubbell stated, "Just for the record, we're actually not using VSLY either anymore, even in sensitivity analyses," and added "but what I really actually wanted to raise a question about is: One of the things I find most intriguing about your results were the graphs that showed that fatal risks increase with age." Saying that this issue has bugged him for a number of years, Dr. Hubbell asked, "How well do we really understand the wage trajectory over the lifespan of an individual? I think about a person entering an occupation, and with that occupational choice they're making a decision at that point about the level of risk they want to accept in the wage tradeoff. From that point forward, however, how much are they actually able to renegotiate based on their own individual age-level risk with their wage trajectory? . . . Say, for example, that risk didn't change over your lifespan for the particular occupation that you're in. You would then actually see an increase in VSL over time, simply because your skill level and wages are going up, and so forth, but your risk level is going to stay the same. So, you're getting an implied VSL that perhaps would seem large if you didn't adjust for individual specific factors well enough. So, one question that arises is: Without a panel study—if you're not controlling random effects in panel fashion—are we getting confounding with individual-level effects in terms of the wage-risk relationship? The other question is: How would individuals actually understand how those risks change over time? Certainly, it was unexpected to see that change, and the question is how much is that information actually out there so there's difference between perceived risk and actual risk in those particular cases."

He closed by mentioning, "One of the things I think is interesting is that there are not a lot of studies out there looking at things other than hedonics as a way to identify these marginal values using the labor market decisions," and he questioned whether hedonics were the only way to use this information or "are there other revealed-preference methods, such as discrete choice type models which look at occupational or job choices and job switching, that could help capture some of this information." He stated that "hedonics tend to assume that there is no bundling of attributes, that you can have any kind of combination of experience with risk and everything else in a very continuous fashion so that you can get these derivatives—and if that *doesn't* occur, you can actually end up with some biased results. Hence, the question is: How much have we explored the bundling of attributes and jobs and whether we can disentangle that."

Joe Aldy

Dr. Aldy answered, "There at the end, Bryan, you started addressing how I would respond to the very first question that you raised, which is whether or not over their lifecycle workers can really adjust their wage or salary in response to changes in the risk. Clearly when one applies the model, we're making the assumption that the labor market has enough freedom and mobility so that one *can* move—so that you can get these kind of equilibria. In the case of what we've done, these equilibria that are age specific, we looked at the relationship between what the firms offer in terms of a combination of safety and labor compensation and what the workers are demanding in terms of that combination of risk and income. We so see that labor income does increase over much of the lifecycle, but then it does start to decrease for some older workers in their 50's. We actually see in our sample a slight decline in labor income for our oldest group." He explained that this could be partly explained by "the standard story that for those who stay within one firm for a long time there's sort of an agreement that they will be paid less than their marginal product when they are young and then will tend to be paid more than their marginal product when they're older." He added that "there's some concern that because of this we're not really getting the right measures. Having said that, we still should be seeing among workers that if they really don't like what's being offered to them they should be going to a different job. That does raise the question of whether workers are really that mobile when they're at older stages of their lives." This led back to the Dr. Hubbell's comment regarding "if there isn't sort of a continuous set of job market characteristic bundles, then you could have some potential problems with this."

Dr. Aldy summarized that "unfortunately there's not much one can do when looking at these cross sections. The benefit of using the CPS is that it's a *massive* cross section; the downside is that it's *just* a cross section. You could construct a quasi-two-year panel, but that's actually very problematic with how they've designed the CPS. We're actually thinking about trying to explore the PSID, where we could have a pretty long panel." He added that he has not "seen any evidence of what perceived risks are and how they vary with age" and he said that he wasn't sure "if in academia professors know how their risk profile changes with their age, but if you're in a blue-collar job where there is a good number of injuries, there's probably a decent sense of that." He said he's sure that there's a sense of that within the firms, "because they're the ones who have to pay workmans' compensation premiums to the state governments, so they should have a sense of how the more serious incidents that can lead to either long-term hospitalization or fatality can vary with age."

_____

Mary Evans, (University of Tennessee)

Dr. Evans said that she first "wanted to applaud your efforts in this paper and other papers in refining the occupational job risk measure that we're able to use—I think that's an important contribution to the literature." She then picked up on something that was mentioned in the response to J.R. DeShazo's earlier question, which was "the issue of possible selection effects within the sample." She said, "You focused on the lower age

range, the 18- to 24-year-olds, but my concerns are more about the upper end of the age distribution, the 55- to 62-year-olds. My question is: Are you able to estimate some sort of selection model?—Have you thought about doing that?—and how do you think doing so might impact your results?"

Joe Aldy
Dr. Aldy responded, "Again, the problem here is in using the CPS—it doesn't really give you anything to identify the decision to work or not." He went on to say that although they had not looked at non-labor income in the household, they did "try to look at non-head-of-household labor income" and added that "it yielded virtually no impact on our estimates, but it wasn't a good instrument—it clearly was not exogenous." Dr. Aldy closed by saying, "As I mentioned, we're thinking about trying to go with the PSID, where we can use a panel. There we can definitely use asset measures to try to identify that . . . it's a much, much richer data set that might enable us to identify selection. It's one reason why we decided to cut the age at 62, the age of early social security retirement. We recognize that there's still a potential problem there, but there are these tradeoffs with the sample that we wanted to use, because at the end of the day when you want to cut this thing by a specific age, it helps to have a 100,000 observation sample. You know, if you want to get the VSL for people who are just age 60, you're not going to get much if you're using the PSID. In the end, that's the tradeoffs one has to make when using a really large data set."

Greg Poe, (Cornell University)
Addressing his comment "to George [Van Houtven] and to the audience," Mr. Poe stated, "It really looks desirable that we're doing risk-risk tradeoffs, and that pulls out the money, and we immediately think that this is going to be a much simpler type decision framework and it might be better. However, a long body of literature from marketing to psychologists to political choice to even birds and bees has shown that these two choices, these tradeoffs, are not that stable. You can add a third choice, which would pass some sort of dominance test such as Maureen [Cropper] referred to, which nobody would choose, but it greatly changes the proportions of people who choose both of those. So, just because we're getting rid of money doesn't mean that we've solved everything and it immediately makes it a preferred technique—it's just *another* technique, and we need to investigate that."

George Van Houtven
"I would agree that it certainly doesn't solve everything, and I wouldn't want to make that claim. I do think, though, that the framework helps in terms of cognitive burden the way we've set it up—in terms of not having to spend as much time explaining the *absolute* value of the risks but rather the relative risks. It's easier for people to trade off—you know, when the denominators are the same, they sort of get canceled out of the equation as long as we're willing to assume and expect the utility framework or something close to that. *That's* the sense in which I think it maybe offers some advantage, but otherwise I agree—it's not a silver bullet."

Reed Johnson, (RTI)

Dr. Johnson commented that "both Joe [Aldy] and Maureen [Cropper] made passing reference to possible connections between the value of a statistical life year and willingness to pay for QALY. I just want to emphasize Alan's [Krupnick] point yesterday that it's important to be vigilant about keeping separate welfare theoretic measures from quality-adjusted health indices like QALY. We get hit over the head all the time by non-economists about this term "value of a statistical life." It's a bad choice of terms, and I think Trudy [Cameron] has advocated something that really describes what we're getting at, which is the willingness to pay for a small change in risk. I think we deserve to be ridiculed if we make the same mistake they make in thinking of the value of a statistical life year as the value of a life year, which of course is what willingness to pay for QALY is supposed to get at."

He went on to say that he's "not sure, though, about George's [Van Houtven] manipulation of ratios of values of a statistical life, whether there's some way of backing out some QALY-like measure out of that. I think it's worth thinking about that, but let's not make the same mistake non-economists make about dropping the statistical part of these value measures."

David Risley, (U.S. EPA)
Starting with the disclaimer that he just started with the Clean Air Markets division about a month ago, so this is all very new to him, Mr. Risley addressed this comment/question to Joe Aldy: "Maybe I'm just offended that my VSL seems to be lower than most peoples' but . . . I'm three years out of undergraduate studies; I have no savings; I moved to the most expensive part of D.C.; and I'm about to start grad school. I know that my debt will be growing, but I *hope* that in the future I'll have earnings potential. I was just wondering if there's any thought of perhaps adding my current VSL to some fractional VSL that represents the likelihood that I'll get to be 40 make more money and have a family."

Joe Aldy
Dr. Aldy replied, "The good news is that your value is going to go up for a while. When Kip [Viscusi] and I were working on this, he didn't like the fact that the peak was always at an age younger than his current age. . . When I think about whether or not the younger population should have a lower value of life, I sometimes wrestle with the question of— you know, this is reflecting imperfections in the labor market, in the capital markets, and that's why they have this lower value. Having said that, this is still what people are using to make actual decisions. Your compensating differential for mortality risk in your job at EPA I presume is probably pretty low—I actually haven't looked at the data closely enough to know how risky it is to work at EPA—I hope it's low. What one is able to infer from that suggests that you're making what is, for you, a rational decision. The idea here is that you're valuing your current consumption enough that you're willing to take less compensation for that probability of dying on the job right now—that's what's implicit in the modeling framework that we have."

Wrapping up, Dr. Aldy commented, "You can take some offense. I can tell you that my father, whom I've talked to about this, takes probably greater offense, for the same reason

Kip does. At the end of the day, what we're trying to say here is that these are the values through which people are trying to reveal their preferences about labor income and risk in their labor market decisions. There are a host of issues in terms of how we try to estimate this and interpret it, as we've already discussed, but at the end of the day that's what we're trying to achieve empirically. It would be nice if you could go to a bank right now and say, "Hey, I'm about to go to graduate school and I'm going to make a lot of money—why don't you give me a lot of money right now?" They're probably not going to do that, but if they did, then you would probably be demanding a larger compensation for the probability of dying at EPA sometime over the next couple of years."

_____

END OF SESSION V Q&A

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

PROCEEDINGS OF SESSION VI: VALUING MORBIDITY AND MORTALITY: DRINKING WATER

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

Table of Contents

Combining Psychological and Economic Methods to Improve Understanding of Factors Determining Adults' Valuation of Children's Health: The Case of Nitrates and Infants

Cheryl Asmus, John Loomis, Helen Cooney, Paul Bell and Bryon Allen (Colorado State University)
May 11, 2006

Abstract

The objective of this research is to evaluate the gain in explanatory power from adding independent variables from the a psychology model of predicting behavior, the Theory of Planned Behavior (TPB) to an economic model, conjoint analysis for determining adults' willingness to pay (WTP) to protect children's health, with the method to be adapted for policy-making. For the development of this method, nitrate in drinking water will serve as the risk factor because it only affects children's health. A questionnaire is used to assess knowledge, attitudes, beliefs, norms, and perceived control with respect to the risk factor, as well as the components of TPB. Respondents also complete a choice task for a conjoint analysis to assess their preferred choices of behavior for averting this risk. One half of the groups are told the choice is hypothetical. The other group is told that one of their four choices will be binding and they will actually buy the amount of bottled water using the money given to them at the beginning of the experiment. We test whether the behavioral responses of these two groups are equivalent or not. The majority of the data collected to date have been in the English-speaking (88%) and hypothetical (76%) treatments.

There was a statistically significant difference in the real/cash cost coefficient and when the costs were hypothetical. The real/cash cost coefficient was far more negative (price sensitive) than the hypothetical cost coefficient, although the hypothetical cost coefficient was still negative.

A household would pay $2.64 in the real cash treatment and $18 in the hypothetical treatment for bottled water that would result in a .0001 (1 in a thousand) reduction in the chances of an infant going into shock from nitrate in water. A household would pay $5.25 in the real cash treatment and $36 in the hypothetical treatment for bottled water that would result in a .0001 (1 in a thousand) reduction in the chances of an infant experiencing permanent brain damage from nitrate in water.

Dividing the coefficient on infants present in the household by the cost coefficients allows us to calculate to investigate the extent of altruism of households without children in terms of their willingness to pay to buy bottled water for households with infants. While willingness to pay (WTP) rises by $49 with real money and $332 for hypothetical payment for households with infants at risk, WTP is still positive for households without an infant. This suggests there is some measure of altruism reflected in our WTP results.

The Theory of Planned Behavior (TPB) variables were only significant at between the .14 and .19 levels and added about 2.5% to the explanatory power of the logistic regression model. Of the variables in the TPB, *attitudes* about infant health issues were not significant p = .48, *health perceived control* (one can protect an infant from environmental contaminants) was significant at p = .19, *water perceived control* (one can control the quality of one's drinking water) was

significant at p = .14, and *water norms* (subjective norms for being concerned about drinking water quality) was significant at p = .15. The results indicate that perhaps perceived control and community norms would be most useful for a policy maker.

Legal Background
Increasingly federal agencies are being called upon to explicitly factor children's health into their regulatory decisions and benefit cost analyses. For example Executive Order 13045 issued by President Clinton on April 21, 1997 required making children's health a high priority in federal agency decision making. In that same year, EPA established the Office of Child Health Protection to give increased emphasis on children health in the agency's many programs. See U.S.E.P.A. (2003) for more details on the Executive Order.

Study Objective
There are two basic issues when valuing children's health. One is selecting the appropriate risk-reducing policies and actions and the other is the value of reducing these risks. Although it is important to economically put a value on the reduction of an environmental health risk to a child, doing so does not necessarily give public and private stakeholders the information they really need to decide upon the appropriate policies or actions.

The kind of information that is most useful to these stakeholders would not necessarily be a dollar figure. It may be an understanding of how and if knowledge, education, belief systems, cultural or societal norms and general attitudes actually lead to the decisions each individual makes when they put a value on a child's health.

To that end, the objective of the proposed research is to test a combining of the explanatory variables from the Theory of Planned Behavior (TPB) with conjoint analysis for determining adults' willingness to pay (WTP) to protect children's health, with the method to be adapted for policy-making.

Study Design
The overall study design focuses on:
(a) Deriving adults' willingness to pay to reduce their infants' risk of shock, brain damage and death from nitrate in drinking water during their first year of life;
(b) Deriving these values using a choice experiment, which involves a hypothetical WTP for bottled water.
(c) Using a consequential treatment in which adults will be asked to pay real money for the bottled water, with a pre-paid coupon for the bottled water provided to those agreeing to pay.
(d) Using the Theory of Planned Behavior to see if attitudes, beliefs, knowledge, norms, and perceived control increase the predictability of adults' WTP choices.
(e) Testing for whether there is altruism toward children's health by testing whether people without infants at risk would pay for bottled water for other households with infants at risk.

2

Hypotheses:
1. In tests of the internal validity of the choice experiment, adults' demand for children's health will be reduced at higher prices (i.e., negative own price), and positive with respective to the amount of risk reduction.
2. In tests of the external validity of the experiment, marginal value for risk reduction i from the traditional hypothetical choice experiment ($MV_i(h)$) will equal the marginal value for risk reduction i from the consequential (real money) choice experiment ($MV_i(c)$).
3. In tests of the predictive power of the Theory of Planned Behavior, regression analyses of WTP for bottled water as a function of risk reduction and cost will show increased predictability by adding beliefs, attitudes, knowledge, subjective norms, and perceived control as predictors and which of those predictors may be more relevant to stakeholders and policy-makers as they make decisions around education or potential mitigation.

Literature Review
Agee and Crocker provide an evaluation of the available methods for valuing children's health. They suggest that stated preference methods such contingent valuation are one of two methods that are most theoretically tenable and analytically tractable. Stated preference methods are not only able to measure parents' willingness to pay for their children, but may also allow elicitation of community public good values toward children's health as well.

While there is a rising demand for children's health information, there have been very few primary valuation studies of children's health issues using stated preference methods. One of the first was Viscusi, et al. (1987) where adults are asked their WTP to reduce adverse health effects to children (in this case pesticide poisonings). Dickie and Messman (2004) perform a very thorough stated preference study of parents' WTP to reduce their own acute illnesses versus those of their children. They used WTP for a medicine that would treat the acute respiratory symptoms such as cough, chest pain, shortness of breath, fever and the untreated duration of these symptoms. For severe acute illness parents are WTP about $217 to reduce one symptom day (Dickie and Messman, 2004: 1167). The values for younger children (age three) is nearly double that of children ages 12 to 17.

WTP of parents to reduce latent skin cancer chances were studied by Dickie and Gerking based on parents WTP for a sunscreen product. Liu, et al. (2000) studied mother's WTP to reduce their own and their child's multiple day, multiple symptom episodes of colds in Taiwan. Converting WTP into U.S. dollars average WTP was $71, and upwards of $121 if adjustments made for differences in income levels and a mid-range income elasticity of WTP.

Valuation Methodology
The methodological approach used in this study is based on the conjoint or choice experiment approach (Holmes & Adamowicz, 2003). This is a stated preference method, in which a respondent makes a series of contingent choices. These choices are contingent upon the characteristics in the choice set. Our choice set has cost as one attribute, and risk of the child going into shock, risk of the child suffering brain damage and risk of death as the key variables

we wish to value. By dividing the attribute coefficient by the cost coefficient the marginal value of a one unit change is monetized.

Following theoretical foundation of Hanemann (1984) on utility difference from random utility models and Roe et al. (1996)'s application to conjoint, we make the first choice a "no action" or baseline risk level associated with no cost. Then the action alternative that reduces the three health risks to the child is offered at a one time cost of X, that varies across the sample. We do this in pairwise fashion, whereby each choice task or choice set is a no action and a single action alternative. As Carson et al. suggest, having just two choices increases the likelihood that the choice will be incentive compatible (even in the hypothetical treatment).

The probability a respondent will choose the action alternative should be related to the expected gain in the parents' well being obtained from their infant receiving the health risk reduction, over and above the satisfaction lost due to paying higher cost. To be more specific, a state-dependent utility function is posited focusing just on the risk of death, to keep the notation simple. Thus UL and UD is the utility to the parent when the child is alive and dead, respectively. Further let PD be the baseline probability of the child dying with and without the risk reduction intervention (e.g.., bottled water). Baseline expected utility (EU) to the parent can be defined as:

$$EU = PD[UD(I)] + (1-PD)[UL(I)] ,$$

where I is income.

The parents' purchase of bottled water reduces the probability of premature death from PD to P'D, but at a proposed cost to the respondent of $X each year. If the reduction in the probability of premature death from PD to P'D yields more expected utility than the loss of $X in income, the parent will select the action alternative in the choice question. Specifically, the expected utility difference (EUD) is given by:

$$EUD= \{P'D[UD(I-\$X)]+ (1-P'D)[UL(I-\$X)]\} - \{PD[UD(I)]+ (1-PD)[UL(I)]\}$$

If this expected utility difference is linear in its arguments, and if the associated additive random error term is distributed logistically, then the probability a respondent will select the action alternative to a question asking him or her to pay $X for the bottled water that would reduce the risk of the child's death from PD to P'D is:

$$\text{Probability of buying bottled water} = P(Y) = 1 - [1 + e^{Bo-B1(\$X)}]^{-1}$$

Maximum likelihood statistical routines such as logistic regression can be used to estimate a transformation of this equation in the form of:

$$\text{Log } \{P(Y)/[1-P(Y)]\} = Bo - B1(\$X) + B2 \text{ (Reduction in Risk of Death)}$$

The marginal value to the parent of reducing a child's risk of death (or parental WTP) is: B2/B1.

4

Theory of Planned Behavior
Besides deriving the adults' value of each type of risk reduction, we wish to explore whether the Theory of Planned Behavior (TPB) adds explanatory power to this model.
According to TPB, there are certain factors that influence behavior. Attitudes toward behavior, knowledge, subjective norms (beliefs about whether the behavior is appropriate), and perceived control have a combined influence on behavioral intentions (whether the individual intends to engage in the behavior or not). In this study, the choices made in the contingent valuation task served as a measure of behavioral intentions. Attitudes, beliefs, knowledge, and perceived control were measured via a questionnaire.

Theory of Planned Behavior
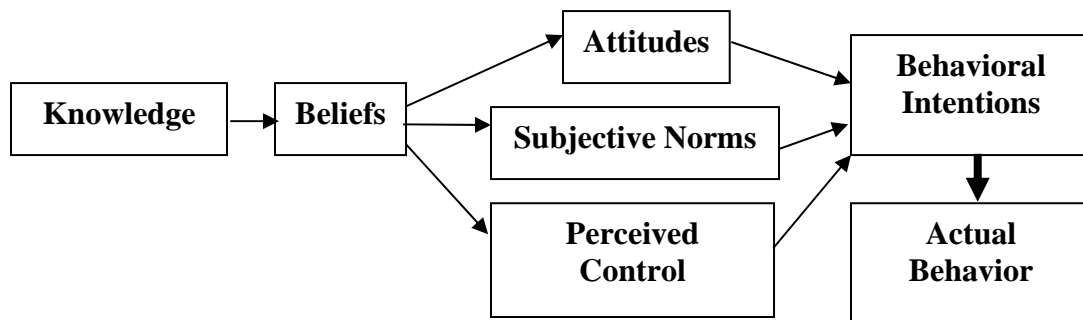
Some sample items are: "I am not aware of any potential negative health effects for children caused by drinking water contaminated with nitrate" (knowledge); "Overall, the children in my community are healthy" (beliefs); "Children's health is an important issue" (attitudes); "Most of the people I know would take steps to ensure that their drinking water is safe" (subjective norms); and "I can ensure that my children are healthy" (perceived control). Behavioral intentions will be assessed via a contingent valuation task. Actual Behavior will be assessed in the experimental conditions via a consequential choice treatment in which the participants will be instructed that the decision they make on one of the choice tasks will be binding.

Choice Experiment Design
The choice experiment involves four attributes (cost, risk of shock, risk of brain damage and risk of death). There were four levels of the risk attributes and seven levels of the cost attribute. We utilized a main effects design to develop an orthogonal choice set with ten different survey versions.

Peer Review of Study Design
The overall study design evolved with numerous discussions with water quality specialists and economists. Several versions of the survey were reviewed by economists that were experts in the area of contingent valuation and choice experiments.

Key Elements of the Survey Design
The key elements of the choice task involves the information provided the respondent and the nature of the alternatives before them.

Figure 1. Key Elements of the Choice Task Given in the Survey

Section 5 → **This section contains a choice task for you to complete. We have listed below some important information, which you may or may not be aware of, about nitrate in water. Please read this information before you continue.**
- ✓ Your community is one of many in Colorado that is at risk for nitrate contamination of its drinking water.
- ✓ Both public water supplies and private wells can be affected.
- ✓ Because infants do not have fully developed digestive systems, drinking nitrate contaminated water can have negative effects on infants' health, but it will not affect adults.
- ✓ Consuming nitrate contaminated drinking water places infants at risk for a condition called "blue baby syndrome" that is caused by depleting the oxygen in the blood.
- ✓ Symptoms of "blue baby syndrome" include a bluish tint to the infant's skin, shortness of breath, shock, brain damage, coma, and death.
- ✓ Using bottled water or water that has had the nitrate removed to prepare formula will eliminate negative health effects caused by nitrate contaminated drinking water for infants, but will not reduce risks from other sources.

What follows is some information concerning different choices you have to reduce health risks to infants associated with exposure to nitrate contamination of drinking water. Please read through the following information and for each pair of options, choose the option that you feel is best.

Options for Preparing Infant Formula

Option A   **Option B**
Use tap water   Use bottled water

Option B may have other potential benefits in addition to reducing exposure to nitrate.

Effects of Over-exposure to Nitrate Contaminated Drinking Water

Cost	Risk of Temporary Shock	Risk of Permanent Brain Damage	Risk of Death
Total, one-time cost of the option in dollars	Risk of infant experiencing decrease in blood pressure and a weak, rapid pulse	Risk of infant experiencing damage to the brain	Risk of infant dying

6

Adults with infants were told the following in the Non Consequential Treatment:

In the next part of the survey you will be asked whether you would purchase or not purchase various amounts of bottled water. This water would help to reduce your infant's exposure to water with excessive levels of nitrate.

If you purchased the water, the health risks to your child from nitrate contaminated drinking water (as well as other potential drinking water contaminants) would be reduced. The amount by which these risks would go down for a given amount of water is presented on the sheet for each choice. Purchasing the bottled water would not reduce risks to your child to zero because she would still face all of the normal risks that do not come from drinking contaminated water.

If you would not purchase the water, your child would continue to face the risks associated with drinking contaminated water (either by drinking the water by itself or by drinking formula that was prepared with contaminated water). The total risk that your child would face if you chose not to purchase the water is also presented on the sheet for each choice. You will be asked to make 4 choices in total.
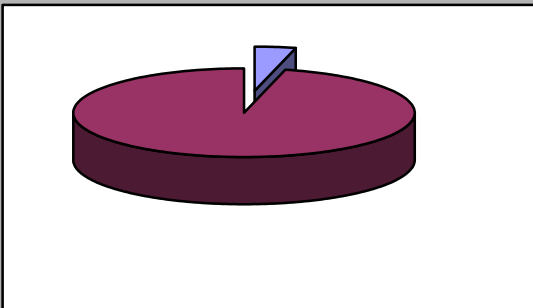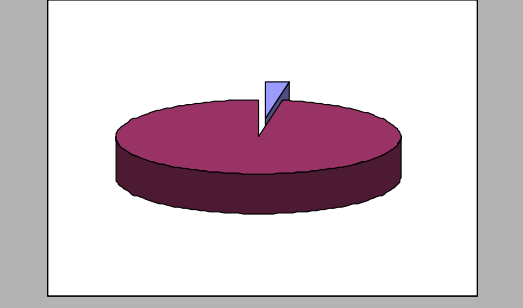
Households without children were told the following in order to allow for investigation into altruism:

In the next part of the survey you will be asked to imagine (pretend) that you have to choose between purchasing or not purchasing various amounts of bottled water for a needy family in your community to help reduce their infant's exposure to water that may contain excessive levels of nitrate.

If you purchased the water, the health risks to the infant from nitrate contaminated drinking water (as well as other potential drinking water contaminants) would be reduced. The amount by which these risks would go down for a given amount of water is presented on the sheet for each choice.

If you chose not to purchase the water, the infant would continue to face the risks associated with drinking contaminated water (either by drinking the water by itself or by drinking formula that was prepared with contaminated water). The total risk that the infant would face if you chose not to purchase the water is also presented on the sheet for each choice. You will be asked to make 4 choices in total.

CONSEQUENTIAL SURVEY TREATMENT

Adults with infants were told the following in the consequential survey treatment.

In the packet containing this survey, you were also given a voucher for $_____. In the next part of the survey you will be asked whether you would purchase or not purchase various amounts of bottled water. This water would help to reduce your infant's exposure to water with excessive levels of nitrate.

If you purchased the water, the health risks to your child from nitrate contaminated drinking water (as well as other potential drinking water contaminants) would be reduced. The amount by which these risks would go down for a given amount of water is presented on the sheet for each choice. Purchasing the bottled water would not reduce risks to your child to zero because she would still face all of the normal risks that do not come from drinking contaminated water.

If you would not purchase the water, your child would continue to face the risks associated with drinking contaminated water (either by drinking the water by itself or by drinking formula that was prepared with contaminated water). The total risk that your child would face if you chose not to purchase the water is also presented on the sheet for each choice.

You will be asked to make 4 choices in total. Choosing between Option A and Option B will allow you to either: actually purchase bottled water for your infant using money provided by Colorado State University or keep the money that it would take to purchase the water.

*At this time, look over the voucher that was attached to your survey. You will see that it is good for a dollar amount that matches the highest cost given for bottled water on the four choice tasks. Once you have completed the survey, send the completed survey along with the signed voucher back to us in the self-addressed postage-paid envelope that we have provided. Once we have received the surveys and vouchers back, we will randomly select one of your four choices between A and B in Section 5. If on that particular task you chose "Do Nothing," you will receive a check for the full amount listed on the voucher. If, on the other hand, you chose "Purchase Bottled Water," you will receive a pre-paid punch-card to obtain the bottled water from a local grocery store. If the value of the punch-card is **less** than the dollar amount given on the voucher, you will be sent a check for the difference.*

Adults without infants were told the following in the consequential survey treatment.

In the packet containing this survey, you were also given a voucher for $_____. In the next part of the survey you will be asked whether you would purchase or not purchase various amounts of bottled water. This water would go to a needy family to help to reduce their infant's exposure to water with excessive levels of nitrate.

If you purchased the water, the health risks to the child from nitrate contaminated drinking water (as well as other potential drinking water contaminants) would be reduced. The amount by which these risks would go down for a given amount of water is presented on the sheet for each choice. Purchasing the bottled water would not reduce risks to the child to zero because she would still face all of the normal risks that do not come from drinking contaminated water.

If you would not purchase the water, the child would continue to face the risks associated with drinking contaminated water (either by drinking the water by itself or by drinking formula that was prepared with contaminated water). The total risk that the child would face if you chose not to purchase the water is also presented on the sheet for each choice.
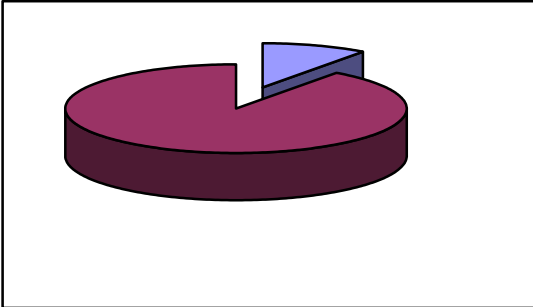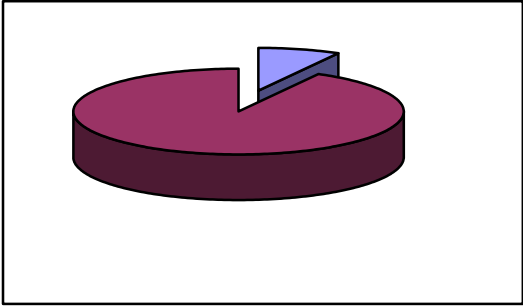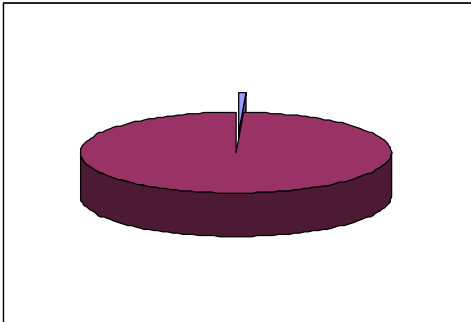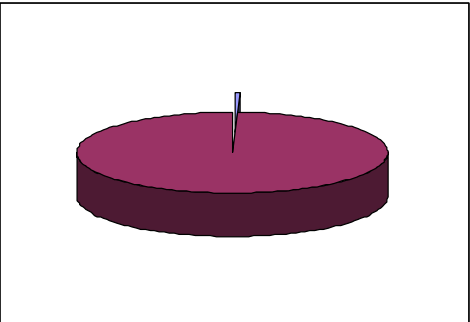
You will be asked to make 4 choices in total. Choosing between Option A and Option B will allow you to either: actually purchase bottled water for an infant in a needy family using money provided by Colorado State University or keep the money that it would take to purchase the water.

*At this time, look over the voucher that was attached to your survey. You will see that it is good for a dollar amount that matches the highest cost given for bottled water on the four choice tasks. Once you have completed the survey, send the completed survey along with the signed voucher back to us in the self-addressed postage-paid envelope that we have provided. Once we have received the surveys and vouchers back, we will randomly select one of your four choices between A and B in Section 5. If on that particular task you chose "Do Nothing," you will receive a check for the full amount listed on the voucher. If, on the other hand, you chose "Purchase Bottled Water," a needy family with an infant will receive a pre-paid punch-card to obtain the bottled water from a local grocery store. If the value of the punch-card is **less** than the dollar amount given on the voucher, you will be sent a check for the difference.*

ACTUAL CHOICE TASK

*For this task, we want you to compare Option A to Option B and choose the option you would actually pick if you had to pay the cost shown. *Risk information is presented in the number of infants in your community out of 1,000 who will be affected.*

Effects	Option A Do Nothing	Option B Buy Bottled Water for an Infant in Your Household
Cost	$0	$300
Risk of Temporary Shock*	100/1000	80/1000
Risk of Permanent Brain Damage*	40/1000	30/1000
Risk of Death*	9/1000	6/1000

Which option do you choose? _____

Data Collection
The survey was pilot tested with two groups, one English-speaking and one Spanish-speaking, in the San Luis Valley area of Colorado. Due to pilot results, the survey was revised to decrease its length and to improve clarity. Data collection was to take place through in-person sessions with participants conducted at various recruitment sites (day care, childbirth classes, etc.). However both participants and sites proved reluctant to participate in this manner. As a result, the data collection methods were altered to include a mail survey mode and "hosted sessions," as well as recruiting from a broader range of areas in Colorado.

For the mail surveys, the survey packets were sent to five early childhood sites, such as Head Start, family centers, or preschools. The packets include a self-addressed stamped envelope for the participants to return the survey. From the time the surveys were mailed to the sites to the time the first participants picked up surveys was approximately three weeks. Participants complete a contact sheet when they pick a packet up at the site and the contact sheets are sent back to the experimenters. Participants are asked to date the slips so that the experimenters know when to begin the reminder phone calls. Using this survey tracking method, the experimenters call participants who have not returned the survey within two weeks and remind them to mail back the survey or send them a new one if necessary. If respondents have simply forgotten to return the survey, they are reminded to do so. If they have lost the survey and are still interested in participating, they are mailed another. In another two weeks they are contacted by phone again and if they don't return the survey, they are counted as a non-respondent and dropped from the study.

To date, information on hosting a session has been disseminated via word of mouth. Starting May 15th, fliers for hosted sessions will be given to individuals who attend in person sessions. For the "hosted" sessions, individuals who are interested in being a host set up a time when they can meet with any friends, family, or acquaintances who are in the demographic groups of interest. An experimenter attends and conducts an in-person session with the guests. Individuals participating in an in-person session received $25 for their participation and those completing the survey via mail receive $15. In the case of "hosted" sessions, participants receive $25 and the host receives $5 for each completed survey.

The target number of participants is 280 (see Table 1). To date, data have been collected from 92 individuals. About one third of them in the adults with infants or expecting category and most have been non-consequential (Non).

Table 1

	Expecting		Child(ren) under 1		Child(ren) 1 to 3		No Children	
English-Speaking	Non	Con	Non	Con	Non	Con	Non	Con
	2	0	20	4	27	6	21	1
Spanish-Speaking	Non	Con	Non	Con	Non	Con	Non	Con
	0	0	0	5	0	4	0	2

RESPONSE RATE

Over the last two months 216 survey packets have been sent to five sites and at least one survey has been returned from each site for a 100% recruiting site response rate. Of the 216 surveys sent to the five sites, 55 participants (25%) have completed a contact card. Of those 55, 23 or 42%, have returned a survey. In addition to the mail recruitment, there have been 2 hosted sessions and 19 surveys have been completed there. There have also been two in-person sessions, both at family centers in southern Colorado.

Response rates to health surveys tend be lower than other types of valuation surveys. For example, Dickie and Messman (2004) who did a parental health survey regarding themselves and their children obtained response of 7.5% of eligible households (those with children). This is on a par with other health valuation surveys such as Johnson, et al. (1997) obtained about 8.8%. So our response rate to date is on a par with these other surveys.

ECONOMIC MODEL RESULTS

Table 2 provides the basic economic model that focuses primarily on the cost and risk reduction variables.

Table 2 Logistic Regression of the Binary Choice Model

Dependent Variable: YESPAY
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/08/06   Time: 21:03
Sample(adjusted): 3 370
Included observations: 363

Variable	Coefficient	Std. Error	z-Statistic	Prob.
CONSTANT	-0.610367	0.503933	-1.211206	0.2258
COST	-0.010853	0.002351	-4.616077	0.0000
HYPCOSTDUM	0.009256	0.001889	4.899394	0.0000
SHOCK RISK REDUC	0.028697	0.010585	2.711134	0.0067
BRAIN DAM RR	0.056937	0.024426	2.331006	0.0198
DEATH RISK REDUC	0.026172	0.081679	0.320430	0.7486
INFANT	0.530914	0.271195	1.957684	0.0503

Mean dependent var	0.707989	S.D. dependent var	0.455315
S.E. of regression	0.431550	McFadden R-squared	0.094395
Sum squared resid	66.29976		
Log likelihood	-198.5367	LR statistic (6 df)	41.38877
Restr. log likelihood	-219.2311	Probability(LR stat)	2.43E-07

Obs with Dep=0	106	Total obs	363
Obs with Dep=1	257		

Where:
Cost is the one time cost to you.

HypCostDum is whether the survey is hypothetical-consequential dummy variable (Hypothetical equals 1) times the one time Cost.

Shock Risk Reduc is the reduction in risk of shock to your child

BrainDamRR is the reduction in risk of brain damage

Death Risk Reduc is the reduction in risk of death to your child

Infant is whether the respondent has an infant (ages 0-1) that would be at risk from drinking water with nitrates in it.

Note that the one time cost is negative and statistically significant at the 1% level. However, the HypCostDum is positive and significant. Thus, when the cost is hypothetical (not actual or consequential), then the net or overall price coefficient becomes much less price sensitive, although still negative suggesting that the higher the price the less likely households are to purchase the risk reduction through bottled water. The difference in the real cash cost coefficient and the hypothetical cost coefficient, provides results of our hypothesis test regarding whether there is a statistical difference in responses of people facing a hypothetical cost and an actual cost. There is quite a difference, with households facing the hypothetical cost being much less sensitive to the cost than households that face an actual monetary opportunity cost. For purposes of comparing marginal values calculated using the actual monetary cost versus the hypothetical cost treatment, we set the HypCostDum to one for hypothetical and adding its coefficient to the Cost coefficient results in a net Cost coefficient of -.001597. Thus to calculate marginal values for the real cost, we divide the attribute coefficient by Cost variable of -.010853, while for the hypothetical cost we use the -.001597.

The positive signs on Brain Damage Risk Reduction, Shock Risk Reduction and Death Risk Reduction make sense. People are willing to pay more the greater the reduction in risk of shock and brain damage is provided by using bottled drinking water. However, the Death Risk Reduction coefficient is not statistically significant and therefore we will not calculate marginal values for this coefficient.

The coefficient on Infant is positive and statistically significant, indicating individuals with an infant in their household are more likely to pay, than those without.

Calculating Marginal Values of Risk Reduction

Marginal Value is Shock or Brain damage risk reduction coefficient divided by the absolute value of the cost coefficient. It is the willingness to pay to reduce shock or brain damage by 1 per 1000 infants. Performing such calculations with our data yields the following results.

A household would pay $2.64 in the real cash treatment and $18 in the hypothetical treatment for bottled water that would result in a .0001 (1 in a thousand) reduction in the chances of an infant going into **shock** from nitrate in water. A household would pay $5.25 in the real cash treatment and $36 in the hypothetical treatment for bottled water that would result in a .0001 (1 in a thousand) reduction in the chances of an infant experiencing **permanent brain damage** from nitrate in water.

Dividing the coefficient on Infant by the cost coefficients allows us to calculate to investigate the extent of altruism of households without children in terms of their willingness to pay to buy

bottled water for households with infants. While WTP rises by $49 with real money and $332 for hypothetical payment for households with infants at risk, WTP is still positive for households without an infant. This suggests there is some measure of altruism reflected in our WTP results.

Comparison of Economic Results to Results with the Theory of Planned Behavior

Table 3. Logistic Regression of the Binary Choice Model with Theory of Planned Behavior Variables.

Dependent Variable: YESPAY
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/08/06   Time: 21:05
Sample(adjusted): 3 370
Included observations: 343

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.204187	1.452501	-1.517511	0.1291
COST	-0.012265	0.002641	-4.644250	0.0000
HYPCOSTDUM	0.010640	0.002184	4.870890	0.0000
SHOCKRISKREDUC	0.022966	0.011058	2.076906	0.0378
BRAINDAMRR	0.038686	0.026311	1.470324	0.1415
DEATHRISKREDUC	0.015765	0.085622	0.184118	0.8539
INFANT	0.749895	0.301004	2.491310	0.0127
HEALTH ATTITUDES	0.149765	0.212629	0.704350	0.4812
HEALTH PERCEIVED CTRL	0.263524	0.199599	1.320263	0.1867
WATER PERCEIVED CTRL	0.461788	0.312784	1.476383	0.1398
WATER NORMS	-0.240546	0.167906	-1.432628	0.1520

Mean dependent var	0.720117	S.D. dependent var		0.449598
S.E. of regression	0.420073	McFadden R-squared		0.121784
Sum squared resid	58.58504			
Log likelihood	-178.5810	LR statistic (10 df)		49.52848
Restr. log likelihood	-203.3452	Probability(LR stat)		3.26E-07

Obs with Dep=0	96	Total obs		343
Obs with Dep=1	247			

Health Attitudes is positive or negative evaluation of health-related behaviors.
Health Perceived Control is perceived control over means of reducing risks to infant health.
Water Perceived Control is perceived control over drinking water safety.
Water Norms is subjective norms for being concerned about drinking water quality.

Evaluation of the Contribution of the Theory of Planned Behavior Variables
Comparison of the McFadden R square of the standard economic model at .094 and the .12 McFadden R square of the full model with the inclusion of the Theory of Planned Behavior suggests these Planned Behavior variables add a small amount of explanatory power (roughly 2.5%). Health Attitudes variable was not significant (p=.48) and Health Perceived Control is significant at the 19% level. The Water Perceived Control and Water Norms were significant at the 14% and 15% levels, respectively.

The Health Attitudes items were scored so that a high score indicates an orientation toward viewing infant health issues as a community problem. Health Perceived Control items were scored so that a high score indicates a strong feeling that one has control over keeping infants free from harm caused by environmental contaminants. Water Perceived Control items were scored such that a high score indicates a strong feeling of personal control over drinking water quality. Water Norms items were scored such that high score indicates a strong subjective norms for being concerned about drinking water quality.

Conclusions
The results support the first hypothesis, indicating that respondents' WTP was negatively correlated with one time cost for bottled water and positively correlated with risk reduction. The second hypothesis was not supported, with respondents in the consequential treatment being more cost sensitive than respondents in the hypothetical treatment. The third hypothesis was partially supported with TPB components accounting for a very small amount of the variance.

The fact that respondents were willing to pay more in the hypothetical treatment than in the consequential treatment makes sense and indicates that for such choices there is a hypothetical bias. The data indicate that individuals who believe infant health is a community issue and have a high degree of perceived control over both infant health issues and water quality issues are more likely to choose to purchase the bottled water, which makes intuitive sense. On the other hand, individuals who perceive strong subjective norms for being concerned about water quality were less likely to choose the bottled water option. It is possible that such individuals feel that the norms are extreme and to a certain extent are reacting against them.

It is hoped that with a complete data set the TPB components will have better explanatory power. A complete data set will also allow more in-depth analyses, including testing for differences between English-speaking and Spanish-speaking participants and a more detailed test of the differences between the different demographic groups (expecting parents, parents of infants, parents of children 1-3 years old, and adults with grown children or no children). The difficulties initially encountered with participant recruitment are informative. Despite the offer of compensation, both sites and participants were generally unwilling to participate in in-person data collection sessions. Sites were much more receptive to distributing mail surveys and the response rate for this targeted quasi-mail was much higher than that obtained in previous research on this topic.

References

Agee, M. and T. Crocker. 2002. On Techniques to Value the Impacts of Environmental Hazards on Children's Health.

Carson, Richard, Theodore Groves and Mark Machina. 2000. Incentive and Information Properties of Preference Questions. Dept. of Economics, University of California-San Diego.

Dickie, M. and V.L. Messman. 2004. Parental Altruism and the Value of Avoiding Acute Illness: Are Kids Worth More than Parents? Journal of Environmental Economics and Management 48: 1146-1174.

Dickie, M. and S. Gerking. 2003. Parent Valuation of Latent Health Risks to their Children. In J. Wessler, H. Weikard, R. Weaver, Eds. Risk and Uncertainty in Environmental Economics. Elgar Publishing. Pages 251-278.

Hanemann, M. 1984. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. American Journal of Agricultural Economics 66(3): 332-341.

Holmes, T. and V. Adamowicz. 2003. Attribute Based Methods. In Champ, P., K. Boyle and T. Brown. A Primer on Non Market Valuation. Kluwer Academic Publishers.

Johnson, R., M.R. Banzhaf and W. Desvousges. 2000. Willingness to pay for Improved Respiratory and Cardiovascular Health. Health Economics 9: 295-317.

Liu, J-T, J.K. Hammitt, J-D. Wang, J-L Liu. 2000. Mother's Willingness to Pay for Own and Her Child's Health: A Contingent Valuation Study in Taiwan. Health Economics 9: 319-326.

Roe, B., K. Boyle, and M. Teisl. 1996. Using Conjoint Analysis to Derive Estimates of Compensating Variation. Journal of Environmental Economics and Management 31: 145-159.

U.S. Environmental Protection Agency. 2003. Children's Health Valuation Handbook. EPA 100-R-03-003. Washington DC.

Viscusi, K., W. Magat, and J. Huber. (1987). An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks. Rand Journal of Economics 18: 465-479.

APPENDIX A – Sample Mail Survey

Valuation of Infant Health Survey Directions

The survey you are going to be completing contains questions concerning water quality, infant health, nitrate, environmental attitudes, and some demographic questions such as age and gender. Part of the survey will also ask you to make a series of choices between two different options for averting risks to infant health that are associated with unsafe levels of nitrate in drinking water.

Please answer all the questions honestly. There are no right or wrong answers to any of the questions. We are only interested in your opinion and attitudes. Your responses will be completely confidential. Even though we have your names, they will not be associated with your responses in any way.

Please feel free to contact Helen Cooney at (970) 491-2119 if you have any questions. Your participation is voluntary and you may quit at any time without any negative consequences.

Please remember to return a signed copy of our informed consent form along with your survey.

Thank you for your participation!

Section 1 → This section asks some general questions about you and your drinking water.
Note: "Tap water" means water that comes out of the faucet in you kitchen.

1) How long have you lived in _____ County, Colorado? _____

2) a) Overall, how would you rate the taste of your tap water?
     ○ Poor   ○ Below Average   ○ Average   ○ Above Average   ○ Excellent

  b) Overall, how would you rate the smell of your tap water?
     ○ Strong unpleasant smell   ○ Somewhat unpleasant smell ○ Noticeable smell   ○ No smell

  c) Overall, how would you rate the appearance of your tap water?
     ○ Colored (brown, red, yellow)   ○ Very Cloudy   ○ Cloudy   ○ Slightly cloudy ○ Clear

  d) Overall, how would you rate the safety of your tap water?
     ○ Poor   ○ Below Average   ○ Average ○ Above Average   ○ Excellent   ○ Don't Know

3) List any problems that you think your tap water has.

_____

_____

_____

4) Do you use a water filter system at home to purify your tap water?
○ Always   ○ Often   ○ Sometimes   ○ Never (Go to question 5)
    If you use a filter system in your home, what type is it?
     ○ Filter Pitcher   ○ Faucet Mounted ○ Under-sink ○ Refrigerator

5) How much money do you spend on each of the following over the course of a typical month?
    Bottled Water (for use at home only)
    ○ None   ○ $1-$10   ○ $11-24   ○ $25-$49 ○ $50 or more
    Filter System at home (maintenance or replacement filters)
    ○ No System   ○ Less than $25   ○ More than $25

6) Does the water in your home come from a well on your property?
    ○ Yes   ○ No (if "No" skip to question 7)
    6a) Do you have your well water tested?
      ○ Yes   ○ No (if "No" skip to question 7)
    6b) How often do you have your well-water tested?
    ○ Once a year   ○ Once every two years   ○ Every five years
    6c) Does your well water meet standards when tested?
    ○ Yes   ○ No

7) Check any of the items below that you think can be a source of nitrate contamination in drinking water.

○ Fertilizer Runoff   ○ Natural Deposits   ○ Decaying Plant Matter

○ Fossil Fuels   ○ Sewage   ○ Landfill Runoff

○ Steel Factories   ○ Discharge from Coal-burning Factories

○ Leaching from Ore-processing Sites   ○ Leaching from Septic Tanks

8) Check any of the items that you think can help you avoid drinking water with high levels of nitrate.

○ Under-sink Filter   ○ Faucet-mounted Filter   ○ Filter Pitcher (e.g., Brita™ filters)

○ Bottled Water   ○ Boiling Tap Water

9) Have you heard about the quality of your community's drinking water?

       ○ Yes   ○ No

10) Do you read the water quality information included in your water bill?

○ Always   ○ Sometimes   ○ Never   ○ Don't receive a water bill

11) Do you prepare formula for an infant (a child under one year old)?

       ○ Yes   ○ No (if "No" skip to question 12)

       11a) How old is the infant?_____

       11b) Do you use bottled water to prepare infant formula?

       ○ Always   ○ Often   ○ Sometimes   ○ Never

12) Have you or a woman in your household been pregnant in the last three years?

         ○ Yes   ○ No (if "No" skip to question 13)

       12a) While pregnant, how often did you or a woman in your household buy bottled water to drink at home?

         ○ Always   ○ Often   ○ Sometimes   ○ Never

       12b) While nursing, how often did you or a woman in your household buy bottled water to drink at home?

         ○ Always   ○ Often   ○ Sometimes   ○ Never   ○ Didn't Nurse

13) Do you have health insurance?

       ○ Yes   ○ No (If no, skip to question 14)

       13a) Does your insurance cover emergency room care?

         ○ Yes   ○ No

       13b) Is your family (spouse and/or children) covered?

         ○ Yes   ○ No

14) If you have children, how much does a visit to the doctor for your child usually cost you?

   ○ $0   ○ $5 - $20   ○ $21 - $30   ○ $31 - $50   ○ $51 - $70

   ○ $71 - $90   ○ $91 - $100   ○ $100 +

  14a) Does an adult in your household have to miss work in order to take a child to the doctor or hospital?

○ Yes   ○ No

Section 2 → This section asks about your beliefs regarding infants' health (consider infants to be children under 1 year of age).

Please check the box corresponding to your responses for questions 1 through 17.

	Strongly Agree	Agree	Don't Know	Disagree	Strongly Disagree
1) If drinking water is safe for adults, it is also safe for infants.	○	○	○	○	○
2) If infants consume water contaminated with nitrate, it can be harmful to their health.	○	○	○	○	○
3) If adults consume water contaminated with nitrate, it can be harmful to their health.	○	○	○	○	○
4) It is natural for infants to become ill more often than adults.	○	○	○	○	○
5) The infants in my community are never ill due to pollution.	○	○	○	○	○
6) My friends and family are concerned with infants' health issues.	○	○	○	○	○
7) The parents I know are worried about the health of their infants.	○	○	○	○	○
8) It is possible to reduce the exposure infants have to pollution.	○	○	○	○	○
9) It is possible to prevent infants from becoming seriously ill due to environmentally caused illnesses.	○	○	○	○	○
10) Only people with infants living in their home need to be concerned about pollution.	○	○	○	○	○
11) Parents, not the public, have the sole responsibility for protecting their infants from harm.	○	○	○	○	○
12) More state and community resources need to be devoted to infant health issues.	○	○	○	○	○
13) There is too much emphasis placed on issues regarding infants' health.	○	○	○	○	○
If you are NOT currently caring for an infant, skip to question 1 of Section 3.					
14) My infant(s) are not exposed to dangerous environmental contaminants.	○	○	○	○	○
15) I can ensure that my infant(s) do not become ill due to environmental contaminants.	○	○	○	○	○
16) I can afford to take my infant(s) to the doctor when they are ill.	○	○	○	○	○

	Strongly Agree	Agree	Don't Know	Disagree	Strongly Disagree
17) I can prevent my infant(s) from becoming seriously ill.	○	○	○	○	○

Section 3 → This section asks what you think about the quality of your drinking water. Please fill in the bubble corresponding to your responses for questions 1 through 7.

	Strongly Agree	Agree	Don't Know	Disagree	Strongly Disagree
1) My community has safe drinking water.	○	○	○	○	○
2) My home's drinking water (straight from the faucet) does not have unsafe levels of nitrate.	○	○	○	○	○
3) My friends and family are worried about our drinking water.	○	○	○	○	○
4) Most of the people I know would take steps to ensure that their drinking water is safe.	○	○	○	○	○
5) Nitrate in drinking water is an unavoidable occurrence.	○	○	○	○	○
6) It is important to me to test the quality of my home's drinking water.	○	○	○	○	○
7) It is the government's responsibility to ensure that my drinking water is safe.	○	○	○	○	○

Section 4 → **We are now going illustrate some risk information for you to help you get used to the way in which risk information is presented as pie charts. Please read the information and then choose which chart represents the greatest risk.**

In the first example, the gray pie wedge represents the fraction or proportion of 1000 accidents which involve Car A and Car B. The larger the gray slice, the greater the risk. As long as the bottom numbers in the fractions (as in this case, 1000) are the same, the larger the top number, the larger the risk.

1) The following charts represent the risk (in number of accidents out of 1000) of being involved in a fatal car crash in two different types of car.

Car A	Car B
150/1000	60/1000
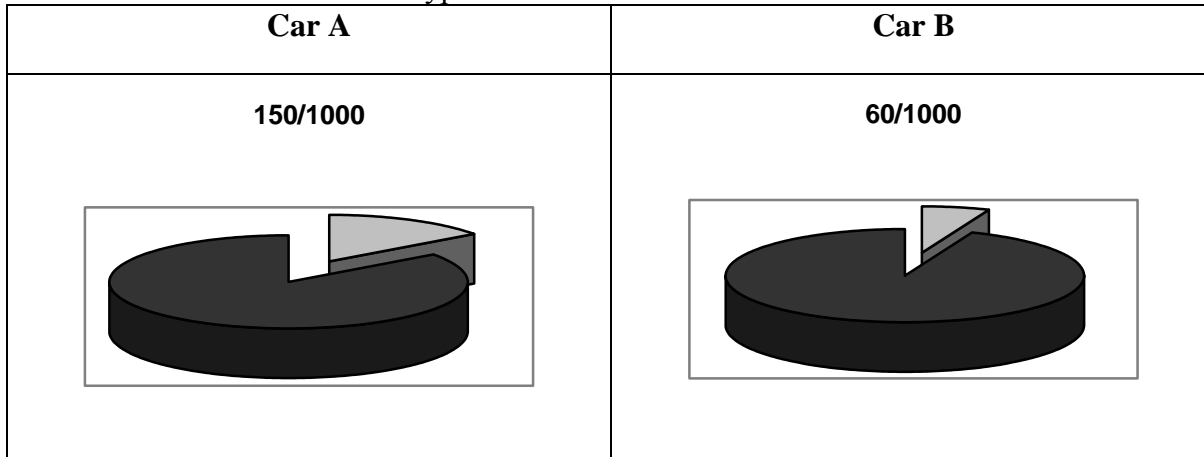
Which car poses the greatest risk? _____

2) The following charts represent the risk (in number of park visitors out of 1000) of being attacked by a mountain lion in two different national parks.

Park A	Park B
15/1000	6/1000
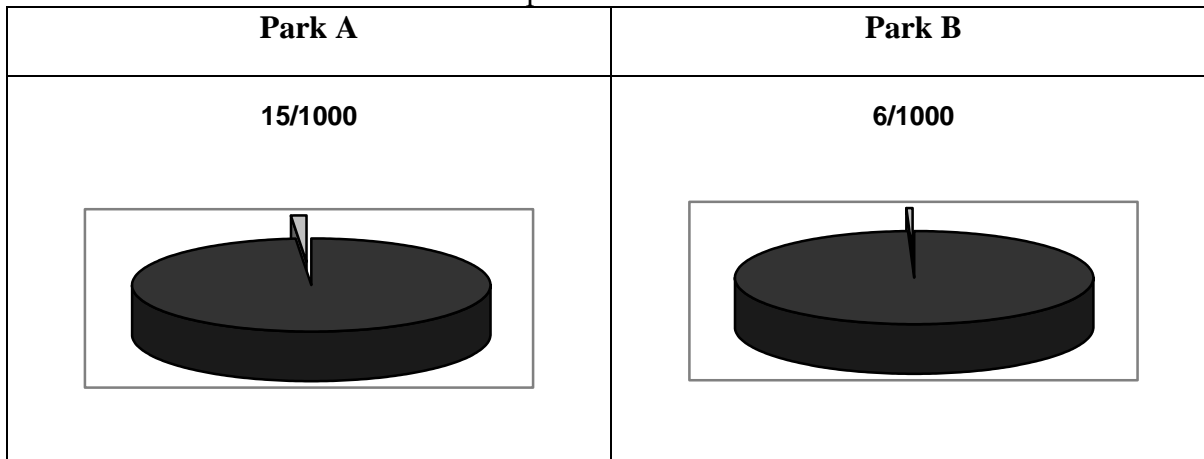
Which park poses the greater risk? _____

1) **The correct answer is A. The top number for A (150) is greater than the top number for B (60).**
2) **The correct answer is A. The top number for A (15) is greater than the top number for B (6).**

Section 5 → **This section contains a choice task for you to complete. We have listed below some important information, which you may or may not be aware of, about nitrate in water. Please read this information before you continue.**

- ✓ Your community is one of many in Colorado that is at risk for nitrate contamination of its drinking water.
- ✓ Both public water supplies and private wells can be affected.
- ✓ Because infants do not have fully developed digestive systems, drinking nitrate contaminated water can have negative effects on infants' health, but it will not affect adults.
- ✓ Consuming nitrate contaminated drinking water places infants at risk for a condition called "blue baby syndrome" that is caused by depleting the oxygen in the blood.
- ✓ Symptoms of "blue baby syndrome" include a bluish tint to the infant's skin, shortness of breath, shock, brain damage, coma, and death.
- ✓ Using bottled water or water that has had the nitrate removed to prepare formula will eliminate negative health effects caused by nitrate contaminated drinking water for infants, but will not reduce risks from other sources.

What follows is some information concerning different choices you have to reduce health risks to infants associated with exposure to nitrate contamination of drinking water. Please read through the following information and for each pair of options, choose the option that you feel is best.

Options for Preparing Infant Formula

Option A	**Option B**
Use tap water	Use bottled water

*Option B may have other potential benefits in addition to reducing exposure to nitrate.

Effects of Over-exposure to Nitrate Contaminated Drinking Water

Cost	Risk of Temporary Shock	Risk of Permanent Brain Damage	Risk of Death
Total, one-time cost of the option in dollars	Risk of infant experiencing decrease in blood pressure and a weak, rapid pulse	Risk of infant experiencing damage to the brain	Risk of infant dying

24

CR

In the packet containing this survey, you were also given a voucher for $_____. In the next part of the survey you will be asked whether you would purchase or not purchase various amounts of bottled water. This water would help to reduce your infant's exposure to water with excessive levels of nitrate.

If you purchased the water, the health risks to your child from nitrate contaminated drinking water (as well as other potential drinking water contaminants) would be reduced. The amount by which these risks would go down for a given amount of water is presented on the sheet for each choice. Purchasing the bottled water would not reduce risks to your child to zero because she would still face all of the normal risks that do not come from drinking contaminated water.

If you would not purchase the water, your child would continue to face the risks associated with drinking contaminated water (either by drinking the water by itself or by drinking formula that was prepared with contaminated water). The total risk that your child would face if you chose not to purchase the water is also presented on the sheet for each choice.

You will be asked to make 4 choices in total. Choosing between Option A and Option B will allow you to either: actually purchase bottled water for your infant using money provided by Colorado State University or keep the money that it would take to purchase the water.

At this time, look over the voucher that was attached to your survey. You will see that it is good for a dollar amount that matches the highest cost given for bottled water on the four choice tasks. Once you have completed the survey, send the completed survey along with the signed voucher back to us in the self-addressed postage-paid envelope that we have provided. Once we have received the surveys and vouchers back, we will randomly select one of your four choices between A and B in Section 5. If on that particular task you chose "Do Nothing," you will receive a check for the full amount listed on the voucher. If, on the other hand, you chose "Purchase Bottled Water," you will receive a pre-paid punch-card to obtain the bottled water from a local grocery store. If the value of the punch-card is **less** than the dollar amount given on the voucher, you will be sent a check for the difference.
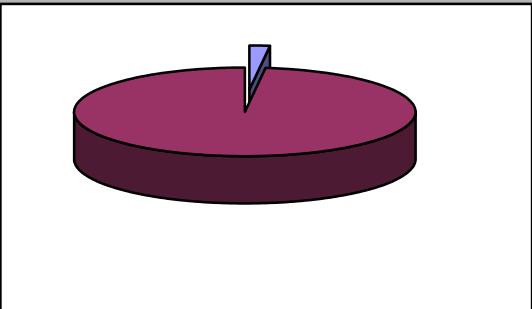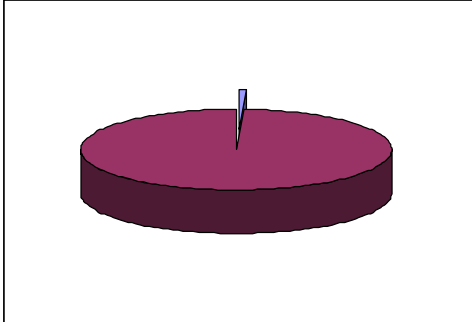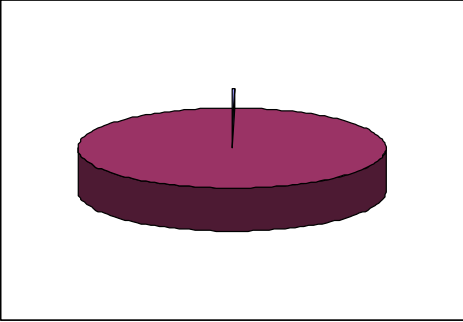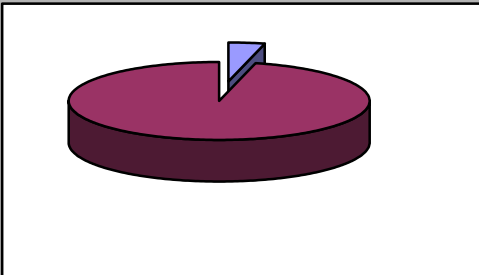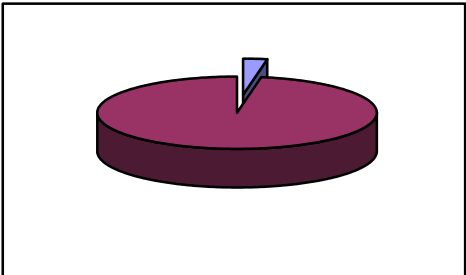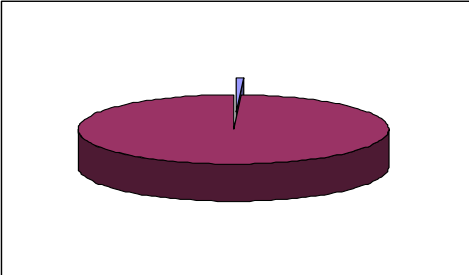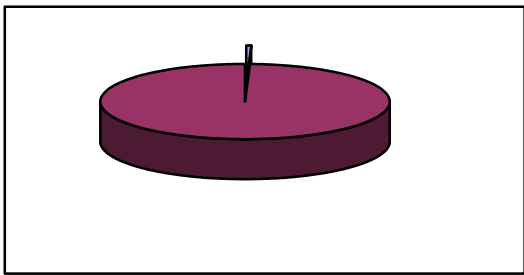
For this task, we want you to compare Option A to Option B and choose the option you would actually pick if you had to pay the cost shown.
*Risk information is presented in the number of infants in your community out of 1,000 who will be affected.

Effects	Option A Do Nothing	Option B Buy Bottled Water for an Infant in Your Household
Cost	$0	$300
Risk of Temporary Shock*	100/1000 	80/1000
Risk of Permanent Brain Damage*	40/1000 	30/1000
Risk of Death*	9/1000 	6/1000

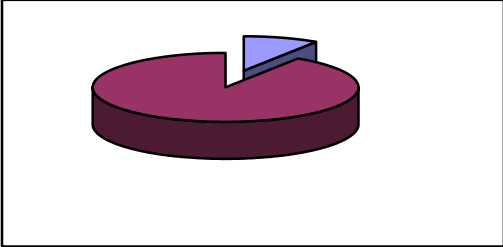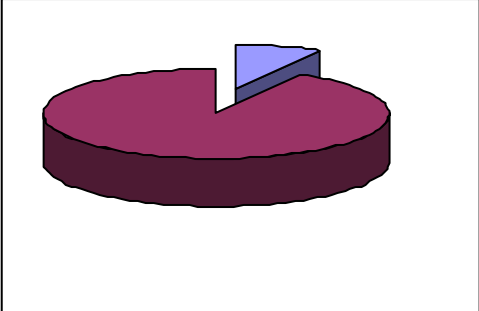Which option do you choose? _____
Why did you choose that option?

_____

_____

For this task, we want you to compare Option A to Option B and choose the option you would actually pick if you had to pay the cost shown.
*Risk information is presented in the number of children in your community out of 1,000 who will be affected.

Effects	Option A Do Nothing	Option B Buy Bottled Water for an in Your Household
Cost	$0	$450
Risk of Temporary Shock*	100/1000	60/1000
Risk of Permanent Brain Damage*	40/1000	20/1000
Risk of Death*	9/1000	3/1000

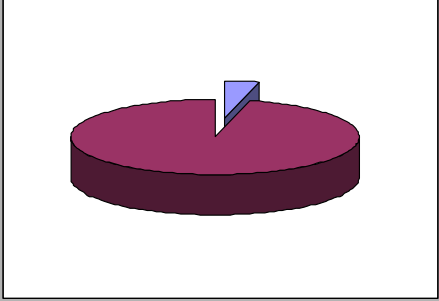Which option do you choose? _____
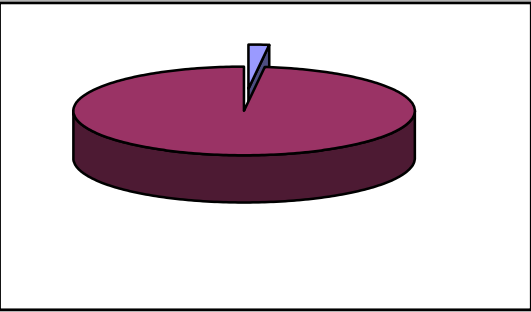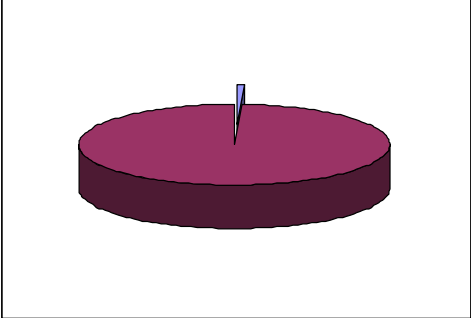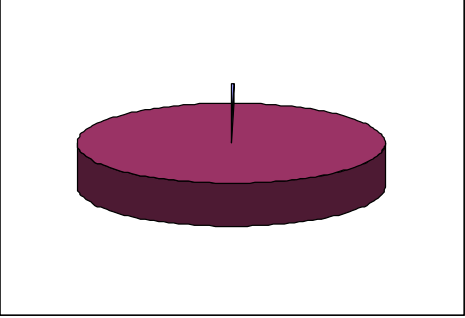Why did you choose that option?

_____

_____

_____
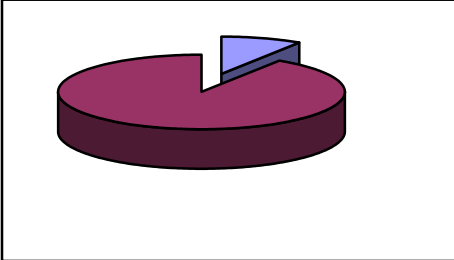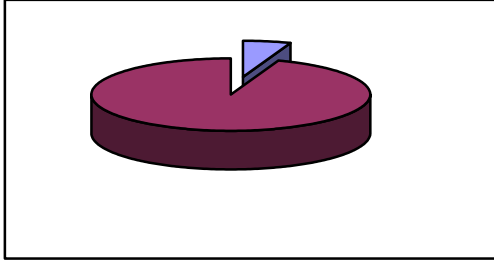
For this task, we want you to compare Option A to Option B and choose the option you would actually pick if you had to pay the cost shown.
*Risk information is presented in the number of children in your community out of 1,000 who will be affected.

Effects	Option A Do Nothing	Option B Buy Bottled Water for an Infant in Your Household
Cost	$0	$400
Risk of Temporary Shock*	100/1000	60/1000
Risk of Permanent Brain Damage*	40/1000	30/1000
Risk of Death*	9/1000	6/1000

Which option do you choose? _____
Why did you choose that option?

For this task, we want you to compare Option A to Option B and choose the option you would actually pick if you had to pay the cost shown.
*Risk information is presented in the number of children in your community out of 1,000 who will be affected.

Effects	Option A Do Nothing	Option B Buy Bottled Water for an Infant in Your Household
Cost	$0	$500
Risk of Temporary Shock*	100/1000	80/1000
Risk of Permanent Brain Damage*	40/1000	20/1000
Risk of Death*	9/1000	3/1000

Which option do you choose? _____
Why did you choose that option?

_____

_____

_____

Section 6 → This section asks for some general demographic information. Your responses will be confidential. No information about your identity (name, SSN, etc.) will be associated with your data. Only researchers on this project will have access to your data.

1) Age _____

2) What is your gender?   ○ Male   ○ Female

3) Occupation _____

4) Number of Years of Schooling:_____

5) Ethnicity (Check all that apply)
○ African American
○ American Indian
○ Asian American
○ European American
○ Hispanic/Latino
○ Native Hawaiian/Pacific Islander
○ Other (_____)

6) Do any of your children (under the age of 18) live in your community?
        ○ Yes   ○ No   ○ I have no children.

7) Do any of your grandchildren (under the age of 18) live in your community?
        ○ Yes   ○ No   ○ I have no grandchildren.

8) Do any of your nieces or nephews (under the age of 18) live in your community?
        ○ Yes   ○ No   ○ I have no nieces or nephews.

9) Yearly Household Income from all Sources
○ $0 - $10,000   ○ $10,001 - $20,000   ○ $20,001 - $30,000   ○ $30,001 - $40,000
○ $40,001 - $50,000   ○ $50,001 +

Colorado State University
Family and Youth Institute Study on Valuation of Infant Health
Voucher

$250

Sign this voucher where indicated and return with your completed survey. Once the choice has been randomly selected, **you will be sent one of three things**:

--A check for the full amount of this voucher (you chose "Do Nothing" on the selected choice)

--A pre-paid punch-card for bottled water worth the dollar amount listed as the cost for the choice (you chose "Purchase Bottled Water" and the randomly selected choice was the one with the highest dollar amount)

--A pre-paid punch-card for bottled water worth the dollar amount listed as the cost for that choice **and** a check to make up the difference between the worth of the punch card and the amount listed on this voucher (you chose "Purchase Bottled Water" and the randomly selected choice was not the one with the highest dollar amount)

_____   _____
Staff Signature   Participant Signature

Economic Valuation of Avoiding Exposure to Arsenic in Drinking Water

Kathleen P. Bell1
University of Maine
Kevin J. Boyle
Virginia Polytechnic Institute

U.S. EPA Morbidity and Mortality Workshop
April 10-12, 2006

Research Team

- **Kelly Maguire (US EPA)**
- **Andrew E. Smith (Maine Bureau of Health)**
- **Laura Taylor (Georgia State University); Tom Crocker (University of Wyoming); Anna Alberini (University of Maryland)**

2

Research Objectives

- **Economic Valuation of Avoiding Exposure**
  - Scrutinize behavioral response of households to information regarding levels of arsenic in private wells
    - private actions at home
    - transactions of residential properties
  - Examine public support for government programs aimed at reducing arsenic levels in drinking water
    - coverage (public and private water supplies)
    - level of reduction

3

Central Research Questions

- **What will be the relationships among valuation estimates derived using different valuation methods?**
  - averting behavior
  - hedonic property value
  - hybrid conjoint / contingent valuation
- **Do household composition and location factors influence behavioral responses?**
  - children, age, gender, health status
  - household location – proximity to arsenic "cluster" areas
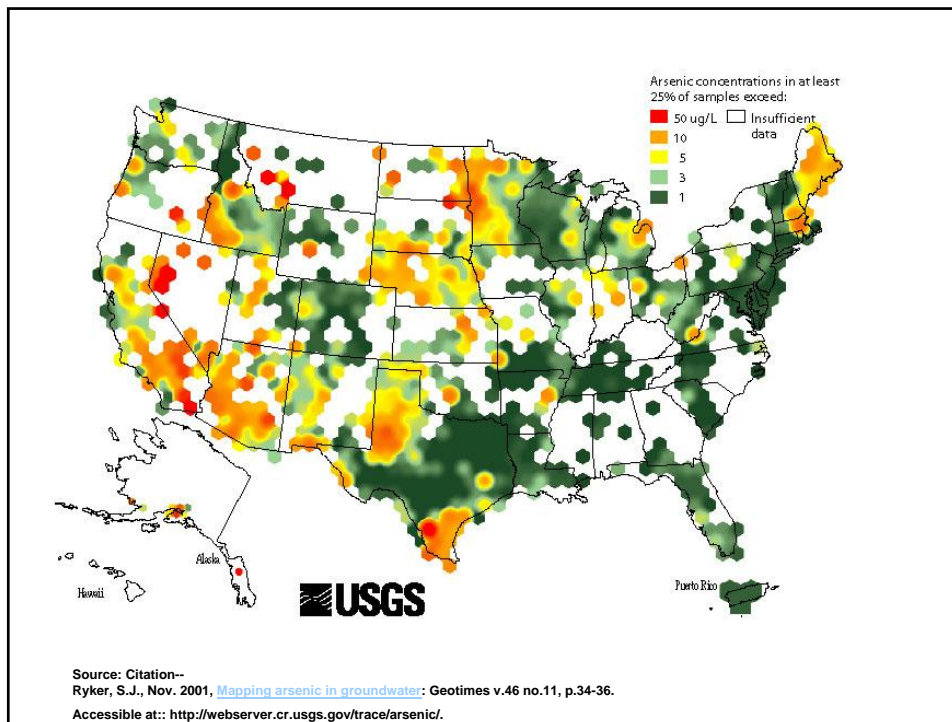
4

Multiple Valuation Methods

- **Revealed Preference**
  - Hedonic Property Value
  - Averting Behavior
- **Stated Preference**
  - Hybrid Conjoint / Contingent Valuation

5

Study Area: Maine

- **Upwards of 50 % of Maine Households Rely on Private Wells for Drinking Water**
- **Assessment of Risks (Loiselle, Marvinney, and Smith 2001)**
  - 10% exceed 10 micrograms per liter
  - 6% exceed 20 micrograms per liter
  - 2% exceed 50 micrograms per liter

6

Arsenic concentrations in at least 25% of samples exceed:
- 50 ug/L
- 10
- 5
- 3
- 1
- Insufficient data

USGS

Source: Citation--
Ryker, S.J., Nov. 2001, Mapping arsenic in groundwater: Geotimes v.46 no.11, p.34-36.
Accessible at:: http://webserver.cr.usgs.gov/trace/arsenic/.

Sample Selection

- **Town Sample**
  - **1,000 randomly selected households from arsenic "cluster" towns**
    - **Buxton, Hollis, Northport, Standish**
- **State Sample**
  - **1,000 randomly selected households**
    - **split – general population (500) versus private well / prior arsenic test (500)**
- **Property Sample**
  - **Sales data from arsenic "cluster" towns**

8

Comparative Approach

- **Samples**
  - **Town Sample**
    - averting behavior
    - hybrid conjoint / contingent valuation
  - **State Sample**
    - hybrid conjoint / contingent valuation
  - **Property Sample**
    - hedonic property value
- **Permits joint estimation**
- **Facilitates comparison and contrast of valuation estimates**

9

Relevant Literature

- **Hedonic Property Value Studies**
  - Contamination of Private Wells (McCormick 1997; Malone and Barrows 1990)
  - Health Risks/Stigmas (Gayer et al., 2000; Gayer et al. 2002; McCluskey and Rausser 2001; Kiel 1995; Kiel and McClain 1995)
- **Conjoint and CV Studies of WTP for State Programs**
  - Safe Drinking Water Supplies (Boyle et al. 1994; Edwards 1986; Bergstrom et al. 2001; Poe et al. 2001)
- **Averting Behavior**
  - Gerking and Stanely 1976; Dickie and Girkie 1991; Shogren and Crocker 1999; Abdalla 1994; Abdalla et al. 1992; Bartik 1998

10

Relevant Literature (continued)

- **Health and Risk Communication**
  - Lead in Tap Water (Griffin and Dunwoody 2000)
  - Risk Communication (Fischhoff 1995; Slovic 1987; NRC 1989; Covello et al. 1989)
- **Environmental and Health Economics**
  - Hazardous Waste Sites (Gayer, Hamilton, and Viscusi 2002; Gayer, Hamilton, and Viscusi 2000)
  - Smokers (Smith et al. 2001)
  - Radon (Smith and Johnson 1988; Smith and Desvousges 2001)
  - Chemical Industry Workplace (Viscusi and O'Connor 1984)

11

Hedonic Property Value Study

- **Objective**
  - Examine evidence of impacts on property values of arsenic levels
    - "elevated"
    - spatial spillovers
- **Valuation estimates**
  - Marginal WTP to avoid exposure

12

Averting Behavior Study

- **Objective**
  - Examine evidence of relationships between averting expenditures/decisions and potential causal factors
    - household composition
    - household location
    - arsenic level in drinking water
- **Valuation estimates**
  - WTP to avoid exposure
  - Value of a statistical life
  - Value of a statistical cancer

13

Hybrid Conjoint / Contingent Valuation Study

- **Objective**
  - Examine evidence of relationships between support for State Programs and potential causal factors
    - household composition
    - household location
    - arsenic level in drinking water
    - household drinking water source
    - program coverage
      - private wells, public supplies, both private and public
    - program scope (level of protection)
- **Valuation estimates**
  - WTP for State Programs

14

Progress

- **Hedonic Property Value Study √**
- **Averting Behavior Study**
  - **Focus Group √**
  - **Survey Design/Approval √**
  - **Survey Implementation**
  - **Analysis**
- **Hybrid Conjoint / Contingent Valuation Study**
  - **Survey Design/Approval √**
  - **Survey Implementation**
  - **Analysis**
- *** Risk Communication Study √**

15

Results

- **Hedonic Property Value Study**
  - **Devanney (2005)**
- **Risk Communication - Aggregate Analysis of Household Testing Decisions**
  - **Huang (2005)**
- **Focus Group Research on Averting Behavior**

16

Hedonic Results (Devanney 2005)

- **Sample**
  - **Buxton and Hollis**
  - **1991 to 2003**
  - **2,212 transactions**
- **Arsenic level**
- **Other explanatory variables**
  - **acreage, structures (age, sqft), time**

17

Measurement of Arsenic

- **continuous or discrete**
  - **elevated levels (> 50 ppb)**
- **property and test**
  - **1 to 1 correspondence**
  - **"closest" test**
  - **average test result within a radius of ¼ mile, ½ mile, or 1 mile**

18

Estimated Parameters (Arsenic Variables)

- **1 to 1 correspondence**
  - insignificant (Buxton)
  - significant (0.1) and negative (Hollis)
- **Closest test result > 50 ppb**
  - significant and negative (Buxton)
  - insignificant (Hollis)
- **Average test result in buffer**
  - ¼ mile
    - significant and negative (Buxton)
    - insignificant (Hollis)
  - ½ and 1 mile
    - insignificant

19

Marginal WTP (Devanney 2005)

¶
Table·4.6.··Estimated·implicit·prices·of·arsenic·concentrations·in·groundwater.¶

Variable	Buxton		Hollis	
	Marginal·implicit·price·($)			
	ML	OLS	ML	OLS
ASHETL	-254.15	-334.03	**-1,497.37***	-1,134.41
AS50	**-400.73****	**-337.67***	-62.92	-26.60
AS50AVGQT	**-216.76***	**-195.68****	-200.91	-125.83
AS50AVGHALF	-31.08	-36.25	-91.63	29.61
AS50AVGONE	-16.07	-19.19	-158.39	-278.37

Notes·–·***,·**,·* ·denotes·significance·at·the·0.01,·0.05,·and·0.1·level,·respectively.··Arsenic·concentrations·are·measured·in·parts·per·billion·(ppb).¶

20

Sales Price Effects (Devanney 2005)

Table 4.7. Sales price effect of arsenic contamination of groundwater.

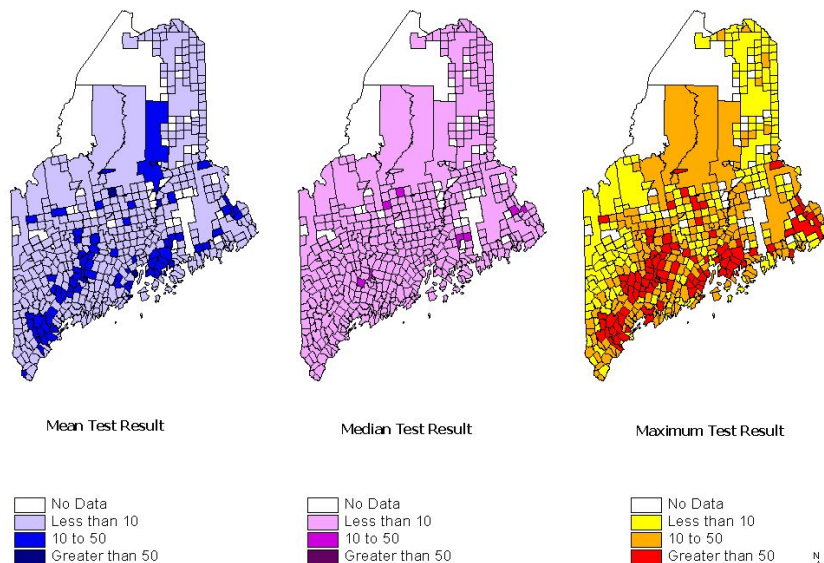| Variable	Δ in Arsenic concentration (ppb)	Buxton		Hollis	
		Sales price effect ($)			
		ML	OLS	ML	OLS
ASHETL	90 to 50	6,318.21	7,715.17	15,764.57	15,422.50
	50 to 10	7,064.88	8,935.10	29,131.23	24,557.76
AS50	90 to 50	9,049.29	7,430.47	1,664.40	732.56
	50 to 10	10,791.36	8,619.16	1,707.89	740.58
AS50AVGQT	90 to 50	5,508.58	5,044.03	5,403.47	3,479.40
	50 to 10	6,059.81	5,497.01	5,867.52	3,663.66
AS50AVGHALF	90 to 50	887.21	1,029.54	2,695.11	-936.61
	50 to 10	900.06	1,046.06	2,797.57	-925.32
AS50AVGONE	90 to 50	465.32	550.42	4,264.64	3,315.02
	50 to 10	469.29	555.75	4,550.82	3,482.20

Risk Communication Study (Huang 2005)

- **Panel Count Regression Model of Annual Arsenic Test Requests by "town"**
  - **Household-level Data on Tests**
    - **Maine State Testing Lab (HETL)**
    - **1990 – 2003**
    - **Final sample size of 16,854 tests(private residential) over 520 towns**

Descriptive Statistics -Private Residential Tests (1990-2003)

Mean Test Result   Median Test Result   Maximum Test Result

No Data	
Less than 10	
10 to 50	
Greater than 50	

Explanatory Variables

- **Demographic Characteristics**
  - Census of Population and Housing (1990 and 2000)
    - WELLS, GENDER, CHILDREN, EDUCATION, INCOME
- **Newspaper Coverage of Arsenic in Drinking Water**
  - Number of total articles
  - "town" referenced in articles

24

Results (Huang 2005)

- **Newspaper coverage (general and town-specific)**
  - positive and significant
- **Household Composition (at "town" level)**
  - education, gender, and income
    - positive and significant
  - proportion of town under age 3 and age 17
    - negative (?) and significant

25

Focus Group (Averting Behavior)

- **Joint Production**

- **Uncertainty / Misinformation**
  - treatment methods in place

- **Share Information with Neighbors**

26

Reflections on Current Results

- **Hedonic Property Value Analysis**
  - further exploration of measurement of arsenic concentration
  - past mitigation
  - timing of sample
- **Household test decisions**
  - role of test
  - sample selection
  - perverse incentive - disclosure

27

EXTRA / BACKGROUND SLIDES

28

Background - Arsenic in Drinking Water: Federal Policy

- **1976 SDWA**
  - MCL of 50 micrograms per liter (1942)
- **1999 NRC Report**
  - Proposed MCL of 5 micrograms per liter (5 ppb)
    - Evaluated 3, 5, 10, and 20
- **2001 SDWA**
  - MCL of 10 micrograms per liter (10 ppb)
  - Public Water Supply Systems Must Comply by 2006

29

Health Effects (NRC 2001)

- **Cancer effects**
  - skin, bladder, lung
- **Non-cancer effects**
  - diabetes, high blood pressure
  - * adverse reproductive outcomes, respiratory effects

30

Variation in Exposure to Arsenic in Drinking Water

- **Ground water sources of drinking water**
  - **Public Water Supply Systems**
    - **100 million persons (2000)**
  - **Private Wells***
    - **40 million persons (2000)**

31

Session VI: Valuing Morbidity and Mortality: Drinking Water

"Perceived Mortality Risks and Arsenic in Drinking Water: Preliminary Research"
Presenter: Douglass Shaw, Texas A&M University

Dr. Shaw opened by stating that he had "*way* too much to say and was going to just launch right in." He went on to put a disclaimer on the upcoming discussant comments, admitting that what he was about to present had evolved since the submission of the paper and, therefore, there might be little correlation between the two. In anticipation of "running out of time" and perhaps not getting to his scripted wrap-up to the presentation, he set out the following summary in advance.

"Here's what we're trying to do—I think it's way different than anything you've heard at this conference. There's been a little talk about what to do with people who look like they're irrational, or what do you do with people who don't get the probabilities, and that kind of thing. There's a sense that you either ignore that and you don't know about it at all or that you throw those people out [of your study]. We're not going to do either one of those things. What we're trying to do is to really bridge the gap between what the decision theorists and the psychologists have been saying about risk and uncertainty for the last 25 or 30 years, but which economists, to some extent, are ignoring—and I don't mean the theoretical economists. The theoretical economists are *loaded* with that stuff— they all know it. If you ask theorists in risk and uncertainty what they think about the expected utility model, they'll say, "It doesn't work, and we have 6 billion experiments to show that it doesn't work.""

Now, our task then is to determine how we adopt a more-general framework in an empirical setting. Can we do something to bring some of what they're telling us is true into an empirical model? So, our agenda is to try and develop an empirical model of one of the non-expected utility models, make it work with survey data, *and* derive a formal derivation of an [unintelligible word] welfare measure that's consistent with that generalized or non-expected utility model.

So, if this sounds like mumbo jumbo, take a look at the paper that we have coming out in the *Journal of Risk and Uncertainty*, technically coming out this month although it probably won't be. That one's on nuclear waste. We're going to try to do it better when we do this with arsenic in drinking water. So, I won't talk a lot about arsenic in drinking water, but I'm going to talk more about the *theory*, and I'll hope you're awake enough to catch it, because I think a lot of this is important.

So, let me jump to the important things. First of all, what we're interested in is called "ambiguity," and ambiguity goes back a long way. Daniel Ellsberg thought of this in 1961 and talked about how when he did experiments, people were averse to ambiguity. So, what is ambiguity? Ambiguity is *uncertainty about the risks.* So, if you think that you know the risks but people say, "I heard what you said, and I saw all your visuals and everything you communicated to me, and I still don't get it—I still am not certain about these risks"—that's ambiguity. In a lot of conventional settings, you'd say that we can't do anything with ambiguity. Well, we *can*, so that's what we're trying to do.

When you introduce ambiguity, what happens is that the conventional expected utility model is not going to be that useful. It's going to provide a benchmark, so we're not going to throw it out—we're going to try and use it—but we're going to try to expand on that and let it provide a benchmark for us. So, Kathleen [Bell] told you all about arsenic risks, but let me get a little bit of the nitty-gritty of where the problem is [he refers to slides.] If you go to the reports that were done for the arsenic rule, it's as Kathleen said: there were some thresholds. The old threshold was, of course, 50 parts per billion (ppb); the new one is 10. What do we know about 50? Well, that in itself the experts don't agree on. Even one of the members of the Science Advisory Board for arsenic said, when I asked him if 10 was safe, "Well . . . we didn't all agree about that." When I asked then why they set it at 10, he responded, "Cost—the cost of compliance. If you go below 10, it's going to be a real problem, particularly for rural areas."

So, is 10 safe? Well, we're going to *say* that 10 is safe, but the reality is if we tell people that 10 ppb is safe (which may or may not be true, but let's assume for now that it is), the *real* problem for people, as Kathleen suggested, is between the 10 and wherever they are. So, if someone is at 22 ppb, and they're wondering if they're safe, what *exactly* are the health risks? We know some things from the Science Advisory Board. For instance, we're looking at 30-to-60 times higher than baseline risk for the incidence of lung and bladder cancers. That's good that we know that, but we have no exact credible relationship that's been mapped out between 0 and 10, or 0 and 22, or 0 and 50. Now, there was extrapolation done in the EPA report that looks pretty good, but a lot of the physical scientists are arguing and disputing with that—there are some papers out in some of the science journals that say it was just an extrapolation and if we apply a different approach, we get different results.

So, this is a perfect setting to allow for ambiguity, because if the experts don't even know what the risks are, then how can we expect the respondents to know what the risks are? Then we add to that that you've got all of these complicating factors. It matters *hugely* in a drinking water setting that risk is completely endogenous. It can be completely controlled by averting behavior, either through your drinking water behavior itself or adopting a treatment—you can solve the arsenic problem pretty quickly if people are willing to install a reverse osmosis or distillation treatment method in their homes or if they're willing to support a program that brings the public drinking water system into compliance. But, that means that the risk is endogenous and there is going to be a lot of averting behavior that we have to watch out for.

Okay, I'll tell you what I know so far about arsenic, and then I'll try to do a little more of the theory to give you an idea about where we're going with this. We did a pilot study for USDA. If you want the papers, there are two of them out that are already published—you can ask me about those—I have a couple of copies of them. Here's the bottom line, and I think Kathleen suggested that they're finding the same thing. Drinking water behavior is very complicated. When we started the pilot study for USDA, we thought it was really simple. We thought that all you have to do is ask people, "Do you drink tap water?" Wrong! What we found out in the pilot study is when you ask people this question, a whole bunch of them say, "No." Then when you follow that up and ask what they do, they might say they drink bottled water. Then when you ask them if they use bottled water in all their cooking and to make tea and to fill up the ice trays and all those kinds of things, they say, "Oh no, we don't do that. We drink bottled water if we're at work, and then when we're at home we use the tap water . . ." They said, "No," but the reality, and what the published studies report, is that they *do* use tap water. So, you have to ask them very detailed questions about what their drinking water behavior is. People who live in two-story houses will ask whether you mean in the kitchen or whether you're asking about the glass of water they drink when they get up in the middle of the night. If you're really going to get a detailed report on drinking water, you have to ask all of that.

[Again, referring to slides, he continued] People don't treat because of cost. We have a relationship in the data that we have, where we were looking at a rural area of Nevada. They were completely aware of the arsenic in the drinking water of the rural area community that we studied. They had been studied to death—the CDC had come in there—Hillary Clinton had gone there—everybody had gone there—they had a very well publicized cancer cluster—they were all very aware, and they're still drinking it. We have people in this rural area who are drinking water with arsenic at 100 ppb—we have some with 500 ppb—the risks are very, very high. We were astonished to see that. So, when you ask them why they're doing that, they say, "Well, you know, the government doesn't know what they're talking about. This stuff is safe—I've lived here all my life, and I don't have cancer." That's a common response that we received. So, we learned a lot from that. We learned that we have to expand the surveys that we are doing to substantially rethink and rework the kinds of questions about tap water that we're going to ask.

Trudy [Cameron] has a paper that I don't think many people know about. It's in the *Journal of Risk and Uncertainty* the year before ours. It's very different from our paper, but she does have a measure of ambiguity, and she finds that it is important in explaining behavior.

Here's another way to think about ambiguity that might be useful. The decision theorists say that you can view a lot of complex problems as a two-stage lottery. In normal expected-utility frameworks, we think that everybody can take a compound lottery and they can do the multiplication and they can reduce that to an easy single-lottery problem. That's called the "reduction of compound lottery axiom," and that's something you would adhere to and believe in if you believe in expected-utility theory. With ambiguity, that's not so. What we think, in fact, is that under ambiguity, people cannot do that. In

experiment after experiment after experiment—many of these get reported on in the *Journal of Management Science*, which is a big decision theory journal, and some other ones—people can't do it. *And*, in conjunction with Reed's comments this morning, [F. Reed Johnson] they are much better able to do it in the context of financial gambles because when you talk about flipping a quarter and getting a head or a tail, they can figure that out. So, if you say there's going to be an expected outcome, they'll say, "Oh yeah, I get that. I know what an expected outcome is"—because it's simple. What we've found in the work that we're doing is that they can't do the same thing when you talk about mortality. It's emotional, and it's very difficult for them to do.

So, what does the utility function look like with ambiguity in it? [again, referring to a slide] There's one right there. They're complicated. You can put in two different probabilities, and you can put in what's called an "absolute risk aversion coefficient." In empirical work, we don't want to *assume* that people are averse to ambiguity—we want to test for it in an empirical model and see if it turns out to be true. In that utility function there [referring to a slide], there are two states, and in state 1, the person doesn't know— the probability could be equally likely to be very small or very large. We see that in behavioral experiments all the time.

So why do people do this? They do this for a lot of reasons. One, psychologists think they do this because they get confused. You give them too many sources of information, and those sources conflict with one another. For instance, in climate change, gee, it might get hotter or it might get cooler—it might get wetter or it might get drier. So their perception is that the questioners don't even know, so they get totally confused and that creates ambiguity. It's more pronounced when they're not confident and, if it can be overcome, if they're quite confident. Chip Heath and Amos Tversky have a paper on that.

[referring to a slide] Here's the mess that you get into: If you allow for ambiguity, then the probabilities that you're dealing with may not do the things that you're hoping they'll do. We all think, for example, that probabilities are supposed to sum to 1 if we've given people a complete space of probabilities. With ambiguity they may not, and the decision theory people are well aware of that. David Schmeidler has this whole new mathematical technique called [unintelligible word] integration which allows for that. You can also have violations of stochastic dominance—and then something I'm working on as a side paper to all of this: What is the welfare measure in all of this? The decision theory people don't care about welfare measures. If we went to them right now and we told them that we're running around estimating willingness to pay for this and that kind of risk reduction, they'd wonder what we're talking about. They'd say, "Do you mean like Pratt's Risk Premium?"—because that's the only thing they know about. And if we say "option price"—some of us in this room know that if you send out a paper to a journal and you say "option price" they send it to a finance person because they think you're talking about pricing financial options. So, there's sort of a gap between them and us on all of that.

Okay, here's the key thing: In all of the alternatives to the expected utility model, the heart, foundation, and soul of these approaches is a probability-weighting function. What we know from observed behavior is that people can over-weight very low probabilities and they can under-weight high ones, and each person can do something different. So, that weighting function there has an inverse-S shape—that's the one that Kahneman and Tversky thought that we would most often observe when we observe behavior. So, for everybody in the sample that we're going to try to go after, we're going to try to get their probability weighting function. The question is "How do you get it?"—and that turns out to be very hard to answer. So, in the focus groups we have been trying to uncover this probability weighting function. Again, the risk and decision theory people have done this in experimental settings, but they've primarily done it with incentive-compatible financial gambles. They'll tell people, "We'll give you $100 to participate in this experiment—we're going to try to do these things on your probability weighting function—if you get it right, you can win a whole bunch of money"—and people try very, very hard to get it right. This is where a lot of this stuff about the trade-offs comes from, so when you talk to people about doing kinds of risk/risk trade-offs or life trade-offs, this is really coming from Peter Fokker's work on whether we want to ask people about the difference between a certainty equivalent and risk measures or whether we want to ask them about something different than that. That led people to wonder, "Gee, could we do this with mortality risks?"

So, that's what we've been up to for the past six months—running experiments and doing focus groups to try to see if we could get at this probability weighting function. We're not having much luck, unfortunately. It's turning out to be quite hard. There is one paper in the entire literature that I know of (and I'm probably wrong—somebody's probably got another one somewhere) where they did it for mortality risks and they were convinced that it was right.

The last little bit that I wanted to talk about is this willingness to pay issue. It has come up time and time again over the past two days, and I wanted to put this slide up about what is this welfare measure that we're trying to get? I want to remind people: We're trying to get the option price. The reason we're trying to get the option price is because Daniel Graham and all that work that he did was trying to help all those great people (Rich Bishop among them) who had struggled with trying to figure out "What do we want?—Do we want an option value?—Do we want an option price?—Can we use the expected surplus?—and all of those kinds of things.

The bottom line is that it's a state-dependent concept. We have two states. We've got the balance of the left-hand side with the right-hand side. What I wanted to make a plug for is that if you look at Trudy's [Cameron] paper in the *Journal of Public Economics* last year and you look at ours that's coming out, you'll see that we're very careful about the derivation of that option price. In ours it's even more complicated because we have ambiguity in the model and we derive the "quasi-" option price—because I don't know for sure what it is yet. But, we take the expected utility difference, and when you solve an equation like this one [referring to a slide], and you set it equal to zero in the top, the expected utility difference, and you solve for some sort of *payment* that balances these

two things, that's how you're going to get the formal expression for what the willingness to pay is. What I'm interested in finding out is whether a lot of the really good people who have been here the last couple of days are doing that.

Okay, that's kind of the theoretical stuff that I wanted to talk about, and if I lost you, I'll be happy to talk about it afterwards. Here's what we're thinking, and it sounds as if we're thinking a lot of what Kathleen [Bell] and them are thinking: With public systems, we're trying to do both public and private wells—the private wells are not regulated by the federal government, whereas the public systems are. What we have figured out from our focus group work so far is that with public systems people can certainly choose to treat within their own homes, and many people often do because they don't like the taste of the drinking water that comes out of the public system. For example, in the town that I live in the water is considered by EPA to be perfectly safe, yet thousands of households have water treatment systems because they can't stand the taste. If you've ever been to Texas, you know that the tap water has a lot of salt in it and tastes terrible. Another interesting health issue is whether people's blood pressure is going up because of drinking the tap water down there, so people are working on that down there too.

Now, their rates may increase if they're on the public system, or they may have already, to pay for getting into compliance, and if it was an [unintelligible word] framework before it happened, of course we would want to find out whether they'd support that rate increase. We're going to also be trying to tackle the child vs. adult health issue, so thank you for all those great papers—I've learned a lot on that, because we haven't been sure how we're going to deal with that. We're going to try to allow for ambiguity for the public system. For the private, we have to ask them if they treat. And remember, when you ask this, it is a very complicated question, because in our focus groups and in the pilot study we did for USDA people say they don't know what is meant by that. So, you have to explain to them very carefully what treatment you mean and which types of treatment that they can adopt actually get rid of arsenic. I never thought in a million years to ask this or to explain this to people, but we've been asked in every focus group so far: What about our refrigerator—is the water that comes out of the door dispenser treated? They do actually put a little filter on the back of many refrigerators to filter the drinking water, but as it turns out, it's a charcoal filter and it's not the kind that actually gets rid of arsenic. But when you tell people that they need a reverse osmosis filter, they don't know what that is. Or, they may respond that they treat, but it turns out to be a Brita water pitcher, which doesn't do anything to get rid of arsenic.

Now, we were thinking of doing complete private welfare measures for those on private wells and then we thought: These people could have a public-goods-related welfare measure. So, again, similarly to what Kathleen and them are doing, we also are going to try to take advantage of different valuations approaches. So, we can have a revealed preference value that comes out of adopting a treatment method, which is going to be averting behavior and can reveal their value for protecting themselves, but there's no reason to think that those people in private homes don't also have some sort of value for a public good. If we can get both out of the same household, we'll be able to cross-validate and look at the preference functions in both cases, which would be really nice. That's

one of the things that—those of you who have never cared about recreation demand modeling, you missed a good thing. The good thing is that those guys figured out that you can use stated preference and a revealed preference in the same model and then do preference restriction tests, which are quite nice, actually.

So, we're going to tackle trying to get all of that—both a public and a private value—out of the people on private wells, and we'll try to get the averting behavior if they've done it.

I'll only share one of the focus groups so far and what came of that. [again, referring to slides] For this group, which was on a public system—this was Eagle Mountain, Utah— we knew that their drinking water was 26 ppb arsenic, so they're not in compliance with the new arsenic rule. We did two focus groups there, and we thought that in the first one we would not tell them. Well, in the first five minutes, when they started to figure out that it was all about arsenic in drinking water, they began to demand that we tell them what their level was. We realized that once you open this door, you have to tell them pretty quickly what their arsenic level is. That's going to be a challenge for our survey team. We're using a risk ladder, and J.R. [DeShazo] and I had a really good discussion on the phone one day about risk ladders vs. grids. We have tested both risk ladders and grids. Surprisingly, our folks are doing a lot better with the risk ladder than they are with the grid, even though some people are saying that that grid is the way to go now. It *may* be that you want to do baseline risks with the ladder and changes with the grid. But then that raises the issue of "Are you overloading them with two different kinds of risk communication devices?" In all of the risk experiments that we've done so far, when we communicate to them that we think the risks are different for children and that they're probably higher, they get that. So, when they come back with subjective risks, we let them mark on the ladder what they think the risks are after we tell them what the experts think the risks are. When they come back and mark them on the ladder, they mark very different points than the experts did, but they get that relative difference between an adult's and a child's risk.

[referring to slides] Here are some summary results from this particular focus group: There were eleven subjects in the focus group and all but one said that the child's risks were much higher. On the risk ladder, they all say very different orders of magnitude from what we told them, which is interesting. That's borne out in the paper on nuclear waste that's coming out in the JRU. Jim Hammitt was talking about having very, very low risks of 1 in 100,000. *Ours for nuclear waste were 2 in 10,000,000!* There is no way that people can understand what 2 in 10,000,000 means—they just cannot do it. So, when we get their subjective risks in that study, and we were looking at people that live along the proposed transportation corridor for shipments to Yucca Mountain, they come back with risks *thousands* of times higher than the DOE says that those risks will be. I told that to Paul Slovik a couple of years ago and he laughed and said, "I told them that. I've been telling DOE that for 20 years and they won't listen to me." But, again, if you have very low risks, in the scheme of things all of this becomes much more important.

So, that's good enough—I'll stop there. Thanks.

Willingness to Pay to Reduce Community Health Risks
from Municipal Drinking Water: A Stated Preference Study

Vic Adamowicz, Department of Rural Economy, University of Alberta
Edmonton, Alberta, T6G 2H1
Phone: 780.492.4603; Fax: 780.492.0268; Email: vic.adamowicz@ualberta.ca

Diane Dupont, Department of Economics, Brock University
St. Catharines, Ontario, L2S 3A1
Phone: 905.688.5550 ext 3129; Fax: 905.688.6388; Email: diane.dupont@brocku.ca

Alan Krupnick, Senior Fellow and Director,
Quality of the Environment Division, Resources for the Future
1616 P Street NW, Washington, DC 20036-1400
Phone : 202.328.5107; Fax : 202.939.3460; Email : krupnick@rff.org

October 2005
Please do not quote without permission.

ABSTRACT

This paper examines the value of health risk reductions to Canadians in the context of clean and
safe drinking water. The health risks we examine pertain both to microbial illnesses and/or
deaths and bladder cancer illnesses and/or deaths. The cancer risks arise because chlorine, the
most common disinfectant used to remove microbial contaminants, has been implicated in the
production of Trihalomethanes (a disinfection by-product) that are linked to increases in bladder
cancer cases. We evaluate results from an panel-based Internet survey of 1,600 Canadians
conducted in the summer of 2004. The survey included text and graphical information regarding
risk changes and employed contingent valuation and attribute-based stated choice benefit
valuation techniques. The valuation questions were designed to elicit consumer preferences for
public programs to reduce health risks associated with improved tap water. Our analysis of the
stated preferences of consumers reveals several types of values that are of interest to policy
makers. These include: the value of mortality risk reduction and the value of morbidity risk
reductions for both microbial contaminants and cancer. In addition, the value of reducing cancer
risks versus microbial risks in a public context is revealed. Our results suggest that reducing
mortality risks from microbial illness has greater value than reducing mortality risks from cancer.
Similarly, overall microbial risk reductions programs (mortality and morbidity) have higher
value than cancer risk reduction programs in this context. In addition, we provide separate
estimates of the value of statistical life associated with cancer and microbial risks, and the value
of statistical illness cases associated with these two risks. The results also include a host of

comparisons between contingent valuation and attribute-based methods, as well as different formats within each of these classes of methods. The values estimated in this study can be used to evaluate investment decisions associated with water treatment, or as estimates of mortality and morbidity value in benefit transfer cases.

Background

Ninety percent of Canadians receive their tap water from public water systems (Environment Canada, 2004). With assistance and scientific input from Health Canada, the Federal-Provincial Subcommittee on Drinking Water (DWS) has developed a set of national drinking water guidelines. The publication *Guidelines for Canadian Drinking Water Quality* lists substances found in drinking water that are known or suspected to be harmful. The most recent summary was published in 2004 (Federal-Provincial-Territorial Committee on Drinking Water, 2004). Substances include both pathogens (microbes such as E. coli, cryptosporidium, giardia, etc.) and potentially carcinogenic chemical by-products (such as Trihalomethanes or THMs). These are formed when chlorine – used for disinfecting water to destroy bacterial and viral contaminants – reacts with other chemicals present in the water.

Provincial regulations require municipal water utilities to provide tap water that is as free as possible from pathogenic micro-organisms called microbes. While many people are familiar with the harm caused by the bacteria, E.coli O157:H7 in Walkerton, it is not the only microbe of concern. Over the last 10 years communities all across Canada have experienced problems with other microbes including: cryptosporidium and giardia. Microbes are generally transported into surface water through agricultural runoff. While most municipalities employ both primary and secondary disinfection technologies - typically chlorine-based - to remove microbes, recent work shows that some microbes are present, even in disinfected tap water (Payment, Berte, Prévost, Ménard, and Barbeau, 2000).

Concern has been expressed about the predominant use of chlorine for disinfection (Carson and Mitchell, 2000). It is implicated in the production of a number of disinfection by-products commonly called Trihalomethanes.[1] These are considered to be potentially carcinogenic. Health Canada convened an expert workshop in 2000 to look into the health risks of drinking water chlorination by-products (Mills, Bull, Cantor, Reif, Hrudey, and Huston, 2000). After reviewing the available evidence, the experts noted that, five epidemiological studies show a statistically significant positive association of chlorinated by-product exposure with risk of bladder cancer. The expert panel concluded "… that it was possible (60% of the group) to probable (40%) that chlorination by-products pose a significant risk to the development of cancer, particularly bladder cancer." Furthermore, they stated that "… this is a moderately important public health problem."

For each substance, the Guidelines establish the maximum acceptable concentration (MAC) permitted in tap water used. A change in any MAC level generally means that water suppliers must improve disinfection techniques in order to meet more stringent requirements. In general, these new methods are more expensive than the traditional chlorine-based methods. While they may or may not be as effective at removing microbial contaminants, they are generally considered to produce fewer THMs. Thus, it is possible that there is a tradeoff between reducing THMs and reducing microbial contaminants.

To inform such tradeoffs, public preferences towards reducing bladder cancer from THMs and microbial disease from pathogens in water must be gauged. Thus, the main research question examined in this paper is how much Canadians are willing to pay on their municipal water bills

in order to reduce these types of health risks from drinking tap water in their community. To our knowledge, ours is the first effort to elicit tradeoffs from individuals between reducing microbial risks and reducing cancer risks within the context of publicly supplied water quality. From a methodological perspective, ours is also one of the few attempts to ask for mortality risk and morbidity risk preferences in the same survey (see Cameron and DeShazo for another example), albeit in a public goods, rather than private goods context.   In addition, our study examines the performance of two stated preference techniques within the same basic survey, i.e., contingent valuation and choice experiments. While there have been several comparisons between CVM with ABSCM (see e.g. Adamowicz et al 1998 or Hanley et al, 2001 for a survey) our comparison includes controls for various context factors including information provision, number of alternatives presented, and a referendum approach.

Using data from an Internet-based survey conducted across Canada during the summer of 2004, the paper presents estimates of the value of reducing one more death or one more illness in the overall population. Values such as these can be used to actually inform choices of technologies for treating drinking water at the plant level and may also be used to help evaluate policy options at the Provincial or Federal level.   For instance, on the one hand, the status quo disinfection technology implies a set of baseline risks for microbial illnesses and deaths and cancer illnesses and deaths. On the other hand, alternative disinfection programs using ozone or ultra-violet light are expected to reduce the health risks associated with cancer illnesses and deaths and with microbial illnesses and deaths. However, these programs are more costly to the household (US EPA, 1999). From the point of view of the public, the decision problem is whether it is worth the

5

additional cost to have reduced risks of both morbidity and mortality effects and whether effort should be focused more on microbial illness reduction versus reductions in cancer cases.

The next section discusses a number of methodological issues addressed in the surveys. This is followed by a description of the survey versions employed in this study. Survey administration and a brief description of the data are presented next. After this, the models and empirical results, along with a number of statistical tests, are described in detail. A discussion of how these results can be useful in a policy context follows. Conclusions and suggestions for future research directions complete the paper

Methodology

Our goal is to obtain information about consumer preferences and tradeoffs relating to household water bill increases and the morbidity and mortality health risks associated with the consumption of tap water. Given the inefficient pricing structure adopted by water utilities and the absence of competitive markets for the sale of tap water, virtually no information exists that yields the value of potable water to Canadians, or indicates which aspects of water are subject to potential tradeoffs according to preferences. In order to obtain this information we constructed a hypothetical market, which allows respondents to express their preferences. We discuss below in detail some important aspects of what we did: preference elicitation methods, presentation of health risks, and public versus private risks.

Preference Elicitation Methods

We employ two non-market valuation methods for eliciting information about consumer preferences for the public good "tap water" -- contingent valuation methods (CVM) and the

Attribute Based Stated Choice Method (ABSCM) (Adamowicz, Louviere, and Swait, 1998). CVM requires the researcher to describe in detail the characteristics of the good to be valued (scenario). Respondents then answer choice questions (we used a double-bounded dichotomous choice format) about whether they would be willing to pay for the described good in its entirety at a stated price. The researcher constructs the willingness-to-pay for the good, where the expressed WTP is for the good in its entirety as described in the scenario, from the pattern of responses. In the ABSCM framework a good is described expressly as a <u>bundle</u> of characteristics or attributes. Each attribute provides valuable services to the consumer. While the individual attributes have value, they cannot be purchased separately but are acquired by the consumer at some stated price for the entire good. With this approach, then, the price paid for a particular bundle of characteristics becomes itself an attribute. In contrast to the CVM method, which provides an overall willingness-to-pay for the bundle of attributes, the ABSCM approach permits us to determine separate willingness-to-pay values for each identified attribute, as well as to examine tradeoffs between individual attributes. For the purposes of this project, the relevant tap water attributes are household water costs and morbidity and mortality health risks from microbial and bladder cancer.

Describing the Health Risks

In presenting the program choices to survey participants we need information about the health effects (a description of each health risk in terms of the symptoms) and baseline risk levels (the likelihood of contracting the disease and or dying from it), as well as changes in risk levels and costs of different programs. A range of reasonable program cost increases was estimated from

7

information on alternative disinfection technologies (US EPA, 1999). These were presented as dollar increases per year in one's household water bill effective January 2005. Information describing symptoms of microbial and bladder cancer illnesses is readily available from a number of sources including Health Canada and the United States Centers for Disease Control. (See Appendix 1 for descriptions used in the survey.)

Baseline information for the number of microbial illnesses and deaths attributable to waterborne microbes is needed for our survey but difficult to ascertain. While outbreak data are collected by regional health officials, they are generally considered to be lower bound estimates of endemic health risks (Mead et al. 1999). This is for three reasons. First, some people become ill prior to general knowledge of an outbreak and are not tested by the doctor for the presence of the microbe, so these cases are not counted. Second, symptoms are often attributed to another cause such as food poisoning or flu. Third, some microbial illnesses are not considered "notifiable" diseases, so doctors are not required to report cases. A second source of data for water-based microbial illnesses is from medical practice cases, which are generally considered to better present the endemic risks (Wheeler et al. 1999; De Wit et al. 2001). A third source of data, which presents the highest estimates of health risks from microbial illnesses, is from microbiological studies examining water supplies for presence of pathogenic micro-organisms.[2] These represent the high end because they assume a dose-response model that links the number of organisms to the number of affected persons (Payment and Riley, 2002). Data on the number of deaths attribute to waterborne microbes are even scarcer. However, Ronchi and Wald (1999), writing in the OECD Observer, claim that "in the United States about 900,000 cases of illnesses and 900 deaths occur every year as a result of microbial contamination of drinking water".

Determination of the baseline risks of becoming ill and/or dying from bladder cancer from consuming water that contains elevated levels of THMs also poses problems. While there is some disagreement in the scientific/medical literature about the relationship between chlorine in water and the incidence of bladder cancer, there are a number of studies that show an association. Recent work under the auspices of Health Canada reports on a study of individuals living around the Great Lakes. The research shows a link between the presence of THMs in drinking water and increased cases of bladder cancer. These results suggest that long-term exposure (on the order of 20-35 years or more) to THMs in water may cause between 14-16 % of all bladder cancer cases in Canada (King and Marrett, 1996). Similar numbers from the United States EPA are between 2-17 % (Mills, Bull, Cantor, Reif, Hrudey, and Huston, 2000). Cancer statistics are available from Health Canada (Cancer Surveillance on-line) by site. Status quo bladder cancer cases attributed to water consumption can be estimated by applying the attribution rates to all bladder cancer cases. Mortality rates are also presented on the Health Canada Cancer Surveillance web site.

With our baseline numbers established (See Appendix 1), we review the engineering and microbiological literature for estimates of anticipated reductions in microbes and/or THMs associated with improvements to water disinfection systems. Numbers from US EPA (1999), Havelaar et al. (2000) and Barbeau et al. (2000) form the basis for our estimates of changes in baseline risks presented to survey respondents (See Appendix 1).

Appendix 1 shows three pages of health risk information presented to our survey respondents. They review this information prior to answering the preference elicitation questions. The first page describes potential health effects associated with using chlorine for water treatment. In particular, it describes symptoms of bladder cancer and clearly identifies the potential tradeoff between the beneficial aspects of reducing microbial contaminants and the potential adverse effects in terms of enhanced risks of contracting bladder cancer. The second page places the baseline microbial and cancer risks together and shows typical linkages between illnesses and deaths for each health condition. In addition, it puts health risks from tap water consumption into a more general perspective. It is important to present the contextual setting to respondents, so that tap water health problems are not viewed in isolation from other health risks. The third page summarizes the baseline health risks from the four health outcomes: microbial illness, death from microbial illness, bladder cancer illness, and death from bladder cancer. Again, the magnitude of health risks from tap water consumption are contrasted with all health risks for each of these health outcomes.

Since we are asking our respondents to assess these health risks, we need to ensure that they are able to evaluate changes in health risks in a meaningful way. Some respondents find numerical representations difficult to interpret. There is a large literature on how best to communicate risk and researchers have used visual aids such as graphs, pie charts, risk ladders, and tables (Jones-Lee et al., 1985; Hammitt, 1990; Corso, Hammitt, and Graham, 2001). We adapt probability communication techniques from Krupnick et al. (2002). After experimenting with a number of options we use what we call our "snake in the sand" design. This begins with a blue rectangle representing a population of 100,000. To this rectangle we add yellow squares representing

individuals who get microbial illnesses from drinking tap water and red squares representing individuals who get bladder cancer from drinking tap water. We superimpose black squares onto either the red or yellow squares in order to illustrate the deaths arising from either microbial illnesses or cancer illnesses. An example of this graphic is shown in Appendix 2 for a CVM format question and in Appendix 3 for an ABSCM format question.

After reviewing the background information (in Appendix 1), the survey respondent is presented with a discussion about changes to water disinfection methods that can alter health risks. The respondent is told that he/she will be faced with a series of choices regarding alternative municipal water treatment programs for his/her community. Each choice includes a status quo (do nothing) option. Alternative programs presented generally lower the health risks and involve an annual increase in the existing water bill for the household. A given respondent answers questions either in the CVM format (example in Appendix 2) or the ABSCM format (example in Appendix 3).

Private versus Public Risks

An issue arising from the approach adopted in this research is whether this particular problem should be treated as an individual (private) decision or a social (public) decision. The private decision context would readily yield individual specific measures of value (e.g. values of statistical life or VSLs) that could be compared to other private good estimates (e.g. Krupnick et al, 2002). A public context, however, is more realistic in this setting since drinking water is

11

consumed by an individual at home as well as at other places (office, school, etc.) and most people view drinking water treatment as a municipal or public responsibility. Therefore, the decision context chosen for this case is a public or social decision. Carson and Mitchell (2000) make the same choice in their open ended CVM survey to obtain willingness to pay for carbon filtration to reduce the risks associated with trihalomethanes,

Thus, with our approach respondents are asked to indicate their preference for one program for drinking water improvement over another (or the status quo). A potential drawback of this approach, however, is that the resulting estimates of the willingness to pay for water quality improvement and for the specific attributes of reduced microbial and cancer risks may contain elements of altruism. That is, when individuals make their choices they may be thinking about their family members, friends, and others in the affected community who will benefit from this program in addition to themselves. Thus, we elicit the individual's preferences including, at once, that for their own health and for the rest of their community. While, in principle, we would like to have these "total social values" to make policy decisions, summing altruistic values from all individuals can introduce an unknown, possibly large degree of double-counting, as opposed to the summing of individuals' values for their own risk reductions, where the latter provides, perhaps, a reasonable lower bound to social value. While this is a challenge, it may also provide us with important and interesting information. Since so many individuals in certain provinces and areas of Canada rely on tap water substitutes, and thus may believe the benefits of such programs will be enjoyed wholly by others, ultimately we hope to be able to sort out altruistic and individual values. This is a topic for future papers.

Outline of Survey Versions

We developed two versions employing the CVM format and 6 versions employing the ABSCM format. In this paper, only four versions of the ABSCM format are referenced. Details of each version follow. Table 1 describes some of the key features of these different versions.

Versions 1 and 2: Contingent Valuation Methodology Format (CVM)

The CVM format (example shown in Appendix 2) presents the respondent with the option to choose status quo (no increase in water bill, no reduction in health risks) or a new municipal water treatment program (increase in water bill, reduction in some or all health risks). Regardless of the versions, each respondent was presented with three separate double-bounded dichotomous choice questions. For Version 1, when compared with the status quo, the first question presented a reduction in bladder cancer illness (from 100 to 50) and a proportional reduction in the risk of death (from 20 to 10), holding constant microbial illness and death risks at their status quo levels. For the second question, respondents were asked to consider a reduction in microbial illnesses from 23,000 to 7,500 and a proportional reduction in the risk of death from 15 to 5, holding constant cancer illness and death risks at their status quo levels. For the third question, the reductions in health risks pertained to all four risks and were the same as those in questions one and two. The payment vehicle was additional costs to the household water bill. Payment levels ranged between $25 per year to $350 per year.[3]

For Version 2 the ordering of the first and second WTP questions was reversed; however, the risk reduction and payment levels are the same as those in Version 1. The third question was identical to that in Version 1.

Versions 3 to 6: Attribute Based Stated Choice Format (ABSCM)

The ABSCM approach begins with the determination of a number of attributes that characterize the good to be valued, along with the setting of the number of separate levels of those attributes. For each choice task the respondent compares a status quo option of no change (risks or household water bills) with either one (Versions 3 and 6) or two (Versions 4 and 5) alternative municipal water treatment programs, where attribute levels for these programs are varied systematically according to the experimental design. Each combination of attributes/levels represents a unique bundle of the good to be valued (Cochran and Cox, 1957). A fractional factorial experimental design procedure is needed to identify those combinations that best reveal the underlying consumer preferences (Louviere, J., D. Hensher and J. Swait, 2000). We identify 32 combinations and divide these into 8 blocks of four questions each. In order to avoid respondent fatigue, each respondent is randomly chosen to face a particular block of 4 choice tasks only. Appendix 3 presents an example of one of the choice tasks faced by survey respondents who received the ABSCM format of the questionnaire.

In order to facilitate a direct comparison of the results from Versions 1 and 2 with the ABSCM format, Versions 3 and 6 present respondents with a status quo option, along with a single alternative program choice. Programs describe three attributes: cancer cases, microbial cases and

household water bill. Cancer and microbial cases are each defined to have four levels of attributes, while household water bill has five levels (including a status quo level of zero increase in a water bill). These values are the same as those used in the CVM questions. We maintain the fixed proportions ratio between morbidity and mortality effects.

One problem with assuming a fixed proportions relationship between illnesses and deaths is that it does not permit us to disentangle the willingness to pay for cancer or microbial morbidity risk reduction from that for cancer or microbial mortality risk reduction. While we could have created a large number of sub-samples of CVM questions using varied proportions, this approach would have been costly since it would have required at least 100 respondents per sample in order to have confidence in the statistical properties of the estimates. A solution is to employ a desirable feature of the alternative ABSCM format to obtain separate WTP values for each of the health risks of interest.

Versions 5 and 6 relax the assumption of proportionality between morbidity and mortality health effects. This requires us to specify five attributes: cancer illnesses, cancer deaths, microbial illnesses, microbial deaths and household water bill. Version 5 presents the respondent with a status quo option, along with two alternative programs. Version 6 is similar to versions 3 in that the respondent may choose only between status quo and one alternative program.

Regardless of the version the ABSCM formats share a common framework. Each respondent provides us with four separate choices across a number of different levels of each attribute.

These choices are pooled and added to choices made by other respondents in order to obtain WTP estimates for the various attributes, along with the tradeoffs between these attributes.

We identify results from the 6 versions as follows. We call results from Versions 1 and 2 our CVM results. We call results from Versions 3 and 4 our Proportional (ABSCM) results and we call results from Versions 5 and 6 our Non-Proportional (ABSCM) results. In all cases we first estimate separate models using data from each version without covariates and follow this with estimates that include covariates. Furthermore, we estimate all models using firstly the full sample of data and secondly a reduced sample of data that removes individuals whose responses identified them as "yea-sayers." (See discussion in next section of how these individuals were identified.) Finally, we also estimate pooled versions of the models: CVM (using data from versions 1 and 2), Proportional ABSCM (using data from versions 3 and 4) and Non-Proportional ABSCM (using data from versions 5 and 6). We also perform series of statistical tests to determine whether these data can be pooled.

- The CVM results are used to obtain a willingness to pay estimate for reductions in cancer risks, a willingness to pay for reductions in microbial risks, and a willingness to pay for reductions in both types of risks together. In the discussion of the empirical results we examine the role played by question order upon these willingness to pay values. We also examine whether the results support both a weak adding-up test (the willingness to pay from question 3 is greater than either the willingness to pay from question 1 or question 2) and a stronger form of the same test (the sum of the individual willingness to pay values from questions 1 and 2 is equal to the willingness to pay value obtained for both

16

items in question 3). We also examine the effects of screening out for yea-saying responses and, finding such effects, do most of our analyses with these respondents removed.

- The Proportional ABSCM results are used to calculate both an overall willingness to pay for the same health risk reductions as described in the CVM scenario for the three items (microbial risk reduction alone, cancer risk reduction alone, and combined cancer and microbial risk reduction). In addition, we present results on the marginal willingness to pay for a one unit reduction in either of these items. As for the CV analysis, we remove respondents who answered questions leading to categorizing their answers as "yea-saying."

- We compare the results from the CVM approach with those using the Proportional ASBCM format in order to determine whether question format has an impact upon the estimated willingness to pay values.

- The Non-Proportional ABSCM results are used to calculate overall willingness to pay values for the same health risk reductions described in the CVM scenario. However, we can now separate out the WTP for the cancer deaths from that associated with cancer illnesses. Similarly, we calculate the WTP for reductions in microbial deaths, as separate from the WTP for microbial illness risk reductions.

- We compare results from the Proportional ABSCM and Non-Proportions ABSCM versions to determine whether a relaxation of the fixed proportions assumption of deaths to illnesses has an impact upon the estimated WTP.

Survey Administration and Data Description

Survey Administration

We employed Ipsos-Reid, a marketing and public research agency to administer and put the survey onto a secure on-line website. Respondents were solicited from amongst a panel of Internet users maintained by Ipsos-Reid. The panel consists of over 100,000 members and reflects an accurate, balanced representation of Internet-enabled Canadians, recognizing that this does not necessarily mean that the panel is representative of all Canadians. These households have been recruited primarily to the panel over the telephone using random digit dialing. After focus groups and pilot testing to refine the survey, we implemented the final version in two waves during the summer of 2004. The waves are only important because they gave us an opportunity to make "mid-course corrections" to the survey, of which there were virtually none. As, after analyzing the data, we have found no reason to distinguish responses by wave, we drop this distinction from here on out. On our behalf Ipsos sent out 4,563 email invitations to its panel of Internet users, of which 2,520 respondents began the survey. Of these 1,633 completed the survey and 419 individuals quit the survey before completion. Additionally, 466 were dropped because they did not obtain any of their tap water from a local municipal water supplier. Finally, 2 responses were deleted after errors arose when the Ipsos server went down in the middle of completing a survey. Assuming that ineligibles are found in the same proportions to those contacted as to those responding (466/2520 = 18.5%), the overall response rate is 46% (1,633/3,536). As we utilize only six of the 8 versions of the survey in this paper, the sample size is 1219.

Table 2 presents summary statistics from the data for all the variables used in this paper, both for the full samples (by stated preference approach -- CVM or ABSCM) and samples that remove "yea-saying" observations. (See below for discussion of how this was done). The Table also reports 2001 Census average statistics for the Canadian population. For most characteristics, average values for survey respondents are virtually the same as these average values. The only socio-demographic characteristic that differs in any appreciable way from that of the general Canadian population is the percentage of individuals educated beyond high school. The 2001 Census estimate is 55 per cent, while the corresponding value for our sample, collected in 2004, is 79.1 per cent. In the previous five years, the percentage of people educated beyond high school increased 5 points. So, the 2004 percentage is likely to exceed 55 percent.

The most important implication of our overeducated sample is in the implication for non-response bias because of the Internet nature of the panel. Statistics Canada (2004) notes that two thirds of Canada's 12.3 million households have at least one family member who regularly used the Internet in 2003. Thus the degree of bias suggested by the education level in our sample may not be very large.

Beyond the issue of sample representativeness is the issue of what is called in the stated preference literature "warm glow" (see, for instance, Kahneman and Knetsch, 1992, "Valuing Public Goods: The Purchase of Moral Satisfaction," JEEM **22** 57-70; and Andreoni, J. 1989. "Giving with Impure Altruism: Applications to charity and Ricardian Equivalence," *Journal of Political Economy* **97**(6), 1447-1458) and "yea-saying" (see, for instance, R. K. Blamey, J. W.

Bennett, M. D. Morrison. (BBM) 1999. "Yea-Saying in Contingent Valuation Surveys," *Land Economics*, Vol. 75, No. 1 (Feb.) , pp. 126-141), the former being a more narrowly defined phenomenon than the latter. The warm glow issue is that, when asked to value a public good, people may derive satisfaction, and be willing to pay something, just from the act of giving. The latter is defined by BBM as "the tendency to subordinate outcome-based or 'true' economic preferences in favor of expressive motivations…." (pg. 126). There are two implications for our results. First, we might expect that people's responses to either CV or CE questions will be insensitive to the commodities or attributes being put up for purchase, and that, therefore, if present in large numbers in the sample, such responses will make it less likely for various tests of sensitivity to scope to be passed. Second, such people are likely to be insensitive to the money – health tradeoffs being posed and are therefore likely to drive WTP estimates inappropriately upwards. We will use the term "yea-saying" to describe this phenomenon.

To address these issues, we used responses to one or two questions to remove "yea-saying" respondents. For the CV analysis, we removed people who said that they would pay anything for health risk reductions and who answered Yes-Yes (YY) to all three dichotomous choice questions with follow-up posed to them. This amounted to 44 respondents (11% of the sample of 407). Interestingly, 10 people who said they would pay anything actually did not answer YY to all three WTP questions. For the ABSCM analysis, we could not use the "YY to all three WTP questions" condition because in a choice experiment set-up there is no equivalent to the YY condition. Even if we had identified those respondents choosing the alternative with the largest health improvement all six times, we would still not necessarily be removing yea-saying effects since attributes are able to differ across programs. Thus,, in very few cases are there

20

clear and unambiguous "yea-saying" answers. Therefore, we used only the first criterion (a statement that the respondent was willing to pay anything for health risk reductions) to screen out respondents. This amounted to dropping 86 respondents out of 812 (10.6 %). The remaining respondents are distributed among 4 ABSCM versions, 361 in the two versions that are directly comparable to the CV approach and 366 in the two approaches that varied morbidity and mortality attributes, permitting separate valuation of these endpoints for both cancer and microbial cases.

As shown in Table 2, the various samples have similar demographic and other characteristics and responses (variables are defined in table 3). The exceptions include URBAN, where a higher percentage of the respondents in Versions 5 and 6 live in urban areas, and ASSETS, where wealthier people were randomly slotted into ABSCM versions (We generally did not use this variable in further analyses because it is missing for too many respondents).

Models

The econometric model used to analyze the CVM survey data is the one appropriate for interval data. We use this model to obtain estimates of WTP

$$\log WTP_i^* = X_i \beta + \varepsilon_i \quad (1)$$

In this equation, *WTP** is the underlying willingness to pay for a selected risk reduction; *X* denotes a vector of age, health, and other attributes; β is a vector of coefficients; and ε is an extreme value Type I error term. Effectively, equation (1) describes a survival time model based on the Weibull distribution. The log-likelihood function for this model is

$$\log L = \sum_{i=1}^{n} \log\{F[(\log WTP_i^H - X_i\beta)/\sigma] - F[(\log WTP_i^L - X_i\beta)/\sigma]\} \quad (2)$$

where F is the type I extreme value distribution with scale σ, $WTPi^H$ and $WTPi^L$ are upper and lower bounds for the payments as presented to respondents in the CVM questions, and X is a vector of age, health, and other attributes with β as the corresponding coefficients. σ is the scale parameter of ε, as well as the reciprocal of the shape parameter of the Weibull distribution describing WTP. The scale parameter for the Weibull distribution is $\exp(X\beta)$. A similar model is also estimated assuming preferences can be described by a lognormal distribution.

A random utility model is used to analyze the responses from the ABSCM format. Random utility theory begins with the assumption that individual consumers choose alternatives that provide them with the greatest utility. It is assumed that an individual's utility is composed of a deterministic component (V) and an unobservable or stochastic component (ε), where V is an indirect utility function. Respondents may choose amongst a number of alternatives. If the stochastic component or error term is distributed extreme value, McFadden (1981) shows that the conditional choice probability of selecting alternative i is:

$$\Pr ob(i) = \frac{\exp(\mu\beta_k Z_i)}{\sum_{j \in C} \exp(\mu\beta_k Z_j)} \quad (3)$$

where Z is a vector of attributes of each program, μ is a scale parameter and C is the choice set. Note, however, that μ is confounded with the parameter vector β and cannot be identified. Normally, μ is set equal to 1.0 and the parameters are estimated using maximum likelihood methods.

An individual application of the method involves the generation of a number of bundles of attributes, and these are presented to respondents in series of choice tasks. Thus, the attributes of each alternative offered in a task comprise the Z vector and the sets of alternatives in each task comprise C, the choice set.[4] In our case, respondents are required to answer four choice tasks. In Versions 3 and 6 the choice tasks consist of comparing the status quo and one alternative program and in Versions 4 and 5 the choice tasks consist of comparing the status quo and two alternative programs. The resulting information is viewed as four individual choices from either a binary or a trinary universe. The econometric analysis (maximization of the likelihood employing the probabilities derived from the equation above) provides the estimates of the marginal utilities associated with the attributes and allows for their use in welfare measures.

Empirical Results: Contingent Valuation Method

There are three sets of results in this section. The first presents the most assumption-free results behind the estimates of willingness to pay – the percent of sample voting Yes to the first bid they are given. The second presents estimates of WTP and subjects these estimates to a variety of validity tests. The third presents regression results to examine construct validity and estimate marginal effects of covariates explaining WTP.

% Yes results

Figure 1 shows "percent Yes" responses by bid separately for each of three public goods being valued (cancer risk reductions, microbial risk reductions, and both types of risk reductions together). Note that each bar in the chart refers to a separate subsample. In general, there is

concern for ordering effects, in that the answers to the microbial reduction program that follow answers to the cancer reduction program may be different than answers to the microbial reduction program when it appears first in the order. We use a likelihood ratio test to show that the two samples may be combined, therefore, we present only the results of pooling the % Yes responses across versions. Our expectation is that % Yes should fall with the size of the bid, and that % Yes should be greater for the third WTP question (the one that combines cancer and microbial reductions offered in the first two WTP questions) than either the responses to cancer or microbial reductions alone. Figure 1 generally and visually supports these expectations, which are confirmed statistically using a Wald Test in Table 5. The figure also reveals that the % Yes for a microbial risk reduction program are generally larger than for a cancer risk reduction program.

CVM WTP Results

Mean and median WTP results appear in Table 4. They are presented for a variety of combinations of cancer and microbial endpoints and for two assumptions about the underlying error distribution (lognormal and Weibull). Estimates are shown both for the full sample and for the sample that removes yea-saying observations. In addition, we present separately results from the two CVM versions (1 and 2) in order to examine issues related to question ordering. Thus, the mean household willingness to pay from the full sample for a reduction in 50 cancer cases of which 10 would have resulted in death (both over a 35-year period) in a community of 100,000 is $535 Cdn per year taken from the responses to the first WTP question in Version 1. This WTP translates into a VSC (a case being the above mortality/morbidity combination) of $14.4 million.[5]

Because the Weibull outperforms the lognormal distribution in a variety of ways and because we believe the yea-saying observations should be deleted, refer to the last two columns but one of the table. The most reliable comparisons are for Cancer asked first (Cancer V1), Microbial V2 (microbial asked first in Version 2), and Both Cancer and Microbial Pooled. The mean WTP are $182, $200 and $294, respectively. Using the pooled versions for added power, mean WTP for reductions in cancer is $157 per household per year, while that for microbial cases is $211 and that for these changes combined is $294. Median WTP is about half that of the mean.

Are any of these differences statistically significant? Table 5 presents results of both Wald tests and likelihood tests. The tests of whether question order matters, or alternatively, whether the answers to the cancer questions can be pooled (and the same for microbial questions and "both" questions) show that they can be pooled by both types of tests. The next relevant comparison is whether the WTP for cancer risk reduction and that for microbial risk reduction are statistically different. Comparing WTP Cancer Pooled to WTP Microbial Pooled we find that the Wald statistics is 3.685, slightly lower than the 95% Chi-squared value of 3.84. Thus, we barely reject the hypothesis that the microbial WTP is larger. Finally, there are two types of "adding up tests" -- what may be termed the weak and the strong adding up tests. The weak test asks whether the WTP for both risk reductions when asked together (in Question 3) is greater than that for either risk reduction separately. Comparing the Cancer Pooled to Both Pooled, we see that mean WTP value for both risk reduction changes exceeds that for cancer alone. However, comparing Microbial Pooled to Both Pooled, we reject this symmetric finding (barely). However, if we compare the more reliable WTP for microbial risk reduction, i.e., when it is the first question asked, to the WTP for the third question (Both) pooled, we find that the Wald statistic exceeds

the target value and that therefore the WTP for both changes exceeds that for microbial risk reductions alone.

The strong test asks whether there is a summation relationship, i.e., whether the sum of the risk reductions for cancer (pooled) and microbial disease (pooled) is significantly the same as the combined risk reductions asked in Question 3 (pooled). In fact, this hypothesis cannot be rejected. In an absolute sense, however, the sum of mean WTPs ($157 + $211) exceeds the WTP for both risk reductions ($294), which could indicate declining marginal utility of health improvements.

CVM Regression Results

Table 6 presents regression results assuming a Weibull distribution (the lognormal results are similar) explaining variables affecting the pooled responses to the cancer risk reduction question, the microbial risk reduction question and the question with both reductions. These results are representative of results with many other specifications and with many variables tried. In all regressions, whether respondents believe in the health information we give them is a robust variable, where those who believe are willing to pay more than those who do not. Household income is *negative and significant for cancer,* but positive and insignificant in the other regressions. This result is somewhat surprising and may arise from correlation between income and other factors in the model, including education and/or belief in the scientific information. Those from larger households and who are older, who have a college education, and who live in more rural areas (but are served by municipal water supplies) are willing to pay more. Interestingly, those who do not engage in averting behaviour are willing to pay *less* than those who do not. This variable could be hypothesized to take either sign. On the one hand, those who do not engage in averting behaviour may feel tap water has few risks, so might be willing to pay less. On the other hand, those who do not engage in averting behavior may have stronger preferences for good water quality, so would be more willing to pay for improvements. The former hypothesis is the one that appears to be closer to the mark.

Finally, for cancer only, there appears to be an ordering effect, where those who answered the cancer question first were willing to pay more than those who answered the cancer question second. This is shown by the significant coefficient on the variable V1Q1 in Table 6. This is a

dummy variable coded 1 for respondents who answered Version 1 questions (cancer risk reduction followed by microbial risk reduction). This is in contrast to the findings in table 5 using the Wald and Likelihood tests. However, the other two comparisons show no difference in responses across versions. The insignificant coefficient on V2Q1 (dummy for version two responses where microbial risk reduction is asked first) indicates no ordering effect for microbials. Similarly, the insignificant coefficient for the dummy variable V1Q3 indicates that there is no significant difference in responses to the third question (microbial plus cancer risk reductions) in either Version 1 or Version 2 responses. Other variables, such as for health status, were not significant.

Empirical Results: ABSCM Method

We estimated six models from the Proportional data (versions 3 and 4), each version independently, and a pooled model, for each of a full sample and a smaller sample that removed yea-saying observations. The parameter estimates are presented in Table 7. These models include only the attributes and status quo constant. All parameters are highly significant and of the expected sign. There are significant status quo effects as illustrated by the positive status quo constant. Tests of pooling Versions 3 and 4 (likelihood ratio tests) indicate that these versions can be pooled and the joint model used for further analysis.

Table 8 provides estimates of the WTP for Microbial Deaths and Illnesses, Cancer Deaths and Illnesses, and the sum of these two WTP values. Since these are proportional models a single WTP amount is presented for both the mortality and morbidity reduction similar to the

presentation of the CVM cases. The removal of yea-saying observations generally reduces the size of WTP. For example, a reduction in microbial risks in the full sample pooled model are valued at $219 while, for the yea-saying removed sample, they are valued at $175. Overall the yea-saying removed sample exhibits WTP reductions of approximately 20 to 35% relative to the full sample. Table 8 also illustrates the effect of the status quo parameter. When excluded the WTP measures are significantly higher (on the order of 50%). Welfare measures with the status quo effect excluded rely on the attributes in the model to capture all of the welfare effect of the change while welfare measures with the status quo included discount changes from the status quo (or changes in attribute levels) by the amount of the status quo preference parameter. There is little guidance in the literature on how to treat this difference, thus we present both measures. If a more conservative measure is desired the WTP with status quo effect included is appropriate. A further finding in Table 8 is that the microbial programs appear to be more highly valued than the cancer programs. This is a finding similar to that obtained in the CVM responses. This policy relevant result carries through many of our findings.

Table 9 presents the parameter estimates for the Non-Proportional versions. As in the Proportional case the parameters are highly significant and the signs are as expected. In this case, however, the test of pooling is rejected. The version with one alternative (Version 6) is statistically different than the version with two alternatives (Version 5). The most significant difference can be seen in the size of the status quo effect. A much lower proportion of respondents chose the status quo in the two alternative version. The reason for this difference is unclear, but shows up in other work by the authors and is a topic for further research. The inability to pool results from these Non-Proportional versions results in our conducting many of

the tests discussed below on Version 5 and 6 individually, as well as on the joint version for comparison.

Willingness to pay measures for the Non-Proportional versions are presented in Table 10. Values are provided for microbial deaths and illness reduction programs (jointly) and cancer deaths and illness reduction programs (jointly) to parallel the Proportional versions and the CVM analysis. In addition, the marginal values per cancer and microbial death and illness case are presented. Separate measures of the mortality and morbidity values are made possible by the Non-Proportional design. In Table 10, the yea-saying removed results are generally lower than the full sample results for the programs, but the difference is not as pronounced as in the Proportional versions. For example, the WTP for microbial deaths and illnesses, with the status quo effect, in Versions 5 is $306 in the full sample and $288 in the removed sample. The cancer program provides values of $110 for the full sample and $80 for the yea-saying removed sample for this same version. The status quo effect, however, is of the same magnitude as in the Proportional case. The difference between status quo included and excluded for a microbial program is $443 versus $306 in Version 5, and $336 versus $163 for Version 6. WTP amounts from Version 6 are generally smaller than those from version 5, although the size of the difference varies. The two alternative version provides smaller WTP values, regardless of whether the status quo effect is included or not. Table 11 provides Wald test statistics examining the differences in the WTP measures in the Proportional and Non-Proportional models. Examining the tests of differences within versions we find that in the Proportional version there is no significant difference between cancer and microbial WTP, but there are differences between WTP for other programs (the critical value for a 5% level of significant is

approximately 5). Examining the Non-Proportional versions, only in Version 6 is the difference between Cancer and Microbial WTP not significantly different. In all cases the WTP values between Cancer and Both are significantly different. The differences between Microbial and Both, however, are generally not statistically different, although they are close to the critical value.

The middle panel of Table 11 tests difference across versions. Interestingly the tests of WTP across version are all insignificant. That is, the WTP value for the Cancer and Microbial programs are not different across the Proportional and the Non-Proportional versions. This is a very powerful result suggesting that the different elicitation strategies do not generate widely different WTP values.

Finally, the bottom panel of 11 provides tests of the adding-up of deaths and illness WTP in the Non-Proportional version (where these two effects are separated) against the WTP in the Proportional version. In all cases the null hypothesis is accepted, indicated that response format did not significantly alter WTP.

Table 12 presents measures of the value of statistical life (VSL) and the value of statistical illness (VSI) for cancer and microbial deaths and illnesses. This summary table is based on the Non-Proportional versions. Several pieces of information emerge from the table. First, in most cases VSLs for microbial mortality are higher than for cancer mortality. This is particularly the case when we use data that has removed yea-saying observations. Second, the VSL values themselves are "high" relative to those in the published literature, however, they are not too far outside the

range of accepted values. Recall that these are values for public reductions in mortality and thus one would expect them to be larger that private WTP values.

The values of statistical cases of cancer are in the $2M to $4M per case range, and range between 20% to 50% of the value of cancer mortality reduction. The value of a statistical case of microbial disease is in the vicinity of $20,000, which is the product of an estimated WTP of $0.018 per case per household and 100,000 people (38,500 households) in the community over 35 years. The value per case appears to be quite high and in our view results from the inability of respondents to register preferences in a choice format that would lead to WTP estimates of a fraction of a cent.

To provide an analysis of the effect of demographic factors on WTP, Table 14 presents two sample sets of parameter estimates that included interactions with demographic factors. These are examples of similar models for the various versions of the Proportional and Non-Proportional models. In general the most robust findings are significant impacts of income (higher income respondents are more likely to choose an alternative program to the status quo and higher income respondents are less sensitive to cancer), Male (less likely to choose a program), Urban (less likely to choose program), and those who believe scientists (are more sensitive to cancer deaths).

Comparison of Results from Two Formats

Results from statistical tests on selected WTP values from the ABSCM and CVM versions of the survey are shown in Table 13. The joint models from the choice experiment are compared to the CVM WTP distributions. The tests of the joint Proportional model reject the hypothesis of equality with the CVM values except for the case of the microbial program and the status quo effect included. However, the values of the Wald tests are not that much above the critical value. The tests comparing the Non-Proportional WTP with CVM are accepted for cancer with the status quo effect excluded, and rejected for all other cases. Recall that the WTP measures with the status quo effect included are considerably lower that with the effect excluded. The WTP from the models with status quo effects includes tended to be lower than the CVM values while the WTP from the models with the status quo effects excluded tended to be higher than the CVM models. Thus, the choice experiment results appear to bracket the CVM results or provide an upper and lower bound.

Using Results to Assist in Policy Making

The values calculated using the approaches discussed in this paper can be used to inform decisions regarding drinking water infrastructure renewal and enhancement, as well as being useful for cost-benefit analysis of drinking water rulemaking. In addition, it is possible that the WTP and regression estimates can be used in various kinds of benefit transfers, to the extent that such values are insensitive to the cause of the health effects (in this case drinking water and its treatment).

The WTP estimates from the non-proportional versions of the survey can be used to derive estimates of the value of statistical life (VSL) or value of a statistical illness (VSI) both for cancer and microbial disease. These values were reported in Table 12. In order to put these results into context, Viscusi and Aldy (2003) examine a number of studies that have produced estimates for the VSL. They report that the range of values is fairly broad between $3.9 - $21.7 million US dollars (2000). While our estimates fall within the upper range of these values, we must note that the majority of studies that calculate a VSL do so using a WTP for a reduction in the risk of death to oneself (that is, a private mortality risk). In contrast, our VSL estimates are based on the WTP to avoid public mortality risks. We would expect that altruistic WTP values might be higher than private WTP values since the former would include the willingness-to-pay to avoid the deaths of members of one's community (including family members). Further, our estimates are for deaths from two specific causes. The fact that the VSL for deaths from microbial disease is somewhat greater than that for cancer is a big surprise. This may be related to previous experience with contamination of municipal water systems in Walkerton, Ontario, and North Battleford, Saskatchewan by microbial contaminants in 2000 and 2001, respectively. More research on this point is needed to verify if this is an artifact of our survey or a true representation of preferences.

In a similar fashion we can calculate the Value of a Statistical Illness as presented in Table 12. Previously, cancer morbidity costs have typically been expressed using costs of illness. Our estimates express costs for cancer in welfare terms.

In addition, we can use the estimated willingness-to-pay values from either the CVM or the ABSCM formats to obtain estimates for the composite value of a statistical case of illness, which includes deaths for a small proportion of cases. Such estimates actually integrate morbidity and mortality in one number so may prove even more useful for policy analysis than the VSLs or VSIs if the policy reduces the source of cases (such as a pollution reduction policy) rather than alters the ratio of illness to death (such as would occur with a health care policy).

Conclusions and Future Directions

This report presents findings from an Internet-based survey designed to elicit preferences relating to tap water quality and health risks. These values show that Canadians are willing to pay in order to reduce the public risks for a number of different water-related health conditions and that they may have a mild preference for reducing microbial contamination over cancer cases. The numbers pertaining to cancer appear reasonable and accord with prior work; however, there are no comparable estimates available for microbial illnesses. We would argue that respondents appear to have trouble with the large number of illnesses presented in the microbial case. This results in small values per illness per respondent ($0.02), but large values per illness when added up over the community. This is clearly an area requiring future study.

A few caveats are in order. Firstly, the numbers in this report, while typical of what we have found, are first round estimates. We are still working to incorporate respondent heterogeneity and to adopt non-linear indirect utility functions. These are the next steps. Secondly, our results show how difficult it is to collect values based on very small probability changes. This is future work. Finally, values collected in this fashion are for public goods, rather than private goods.

.With few studies of this nature and with concerns about double-counting when one adds up public values in the presence of altruism, caution is in order. Attempts to purge our estimates of altruism effects are in our plans for future research.

Table 1: Key Features of 6 Versions of Survey

Version	Question Format	Number of Questions/Tasks Per Respondent	Question Ordering	Number of programs (status quo included)	Relationship between mortality and morbidity
1	CVM	3	Cancer, microbial, both	2	Proportional
2	CVM	3	Microbial, cancer, both	2	Proportional
3	ABSCM	4	Na	2	Proportional
4	ABSCM	4	Na	3	Proportional
5	ABSCM	4	Na	3	Non-Proportional
6	ABSCM	4	Na	2	Non-Proportional

Table 2: Descriptive Statistics by Version and Sub-sample

Variables	Canadian Population Values	CVM and ABSCM (V1,2,3,5,6,7) full sample	CVM (V1 and V2) full sample	CVM (V1 and V2) yea-saying observations removed	ABSCM (V3, 5, 6, 7) full sample	ABSCM (V3, 5) yea-saying observations removed	ABSCM (V6) yea-saying observations removed	ABSCM (V7) yea-saying observations removed
INCOME	58360	57458.52	58734.17	58080.27	56819.12	54743.30	57796.83	60289.60
		(35650.89)	(35562.79)	(35501.41)	(35699.68)	(35012.21)	(35865.36)	(37069.08)
MALE	49.9%	52.75%	54.55%	54.96%	51.85%	49.86%	50.81%	56.35%
		(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
AGE	45.8	46.55	44.93	44.18	47.36	47.58	46.15	47.42
		(15.02)	(15.02)	(15.25)	(14.97)	(15.07)	(15.30)	(14.31)
HHSIZE	2.6	2.59	2.63	2.61	2.57	2.65	2.52	2.53
		(1.31)	(1.26)	(1.26)	(1.34)	(1.35)	(1.37)	(1.25)
EDUCATION	55 %	61.77%	61.18%	62.04%	62.07%	61.77%	64.86%	62.98%
		(0.49)	(0.49)	(0.49)	(0.49)	(0.49)	(0.48)	(0.49)
ENGLISH	73%	76.13%	75.92%	75.64%	76.23%	76.73%	75.68%	74.03%
		(0.43)	(0.43)	(0.43)	(0.43)	(0.42)	(0.43)	(0.44)
URBAN	80% [d]	65.14%	61.67%	63.17%	66.87%	64.82%	70.27%	71.27%
		(0.48)	(0.49)	(0.48)	(0.47)	(0.48)	(0.46)	(0.45)
ASSETS	na	89417.54	79677.19	79464.06	94483.83	93916.12	94999.78	93749.79
		(82906.93)	(74853.48)	(75809.54)	(86427.92)	(85986.44)	(85166.08)	(90464.05)
BELIEFMS	na	74.82%	73.96%	71.39%	75.25%	72.85%	80.00%	70.17%
		(0.43)	(0.44)	(0.45)	(0.43)	(0.45)	(0.40)	(0.46)
NOAVERT	na	45.37%	45.95%	46.46%	45.07%	42.38%	49.19%	45.86%
		(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.5)	(0.50)
N		1219(906*)	407(310*)	353(266*)	812(596*)	361(263*)	185(135*)	181(130*)

Notes: [a] Standard deviations are in brackets.

[b] Yea-saying data is identified in CVM samples when YY for all three CVM questions and respondents indicate are willing to pay anything for health risk reductions. Yea-saying data is identified in CE samples when latter condition is true.

[c]* denotes number of observations for ASSETS.

[d] The Census **definition is more encompassing than ours. It includes an individual as being in a rural area if the population is less than 1000. We used 10,000 to better capture locations with municipally supplied water.**

Table3: Definition of variables

Variables	Definition
INCOME	Average household income in Canadian $
MALE	Percentage of respondents who are men
AGE	Average in years
HHSIZE	Average size of a household
EDUCATION	Percentage of respondents with more than Some Community College/CEGEP/Trade School
ENGLISH	Percentage of respondents whose first language is English (This is indicated by whether respondents completed the survey in their choice of English or French)
URBAN	Percentage of respondents live in a city in which the population is over 10,000
ASSETS	Total value of household's financial assets in Canadian $
BELIEFMS	Percentage of respondents who believe scientists are certain about microbial illnesses arising from drinking tap water. (Highly correlated with other belief variables relating to certainty of scientific community about risks associated with cancer and microbial deaths and cancer illnesses.)
NOAVERT	Percentage of respondents who undertake no averting behavior against drinking water related health risks

Table 4. Mean and Median WTP Estimates by Endpoint, by Assumed Distribution, Full and Clean Sample

| | Lognormal | | | | Weibull | | | | |
| | Full Sample | | Yea-saying Observations Removed | | Full Sample | | Yea-saying Observations Removed | | Weibull Mean Test |
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	
WTP Cancer V1	535 (199)	110 (16)	289 (82)	84 (11)	266 (46)	119 (44)	182 (27)	91 (27)	2.458
WTP Cancer V2	393 (149)	72 (11)	201 (55)	55 (8)	200 (33)	79 (31)	133 (19)	60 (19)	3.097
WTP Microbial V1	1075 (567)	130 (22)	532 (220)	95 (15)	332 (67)	137 (61)	226 (39)	104 (38)	1.883
WTP Microbial V2	552 (194)	118 (17)	344 (100)	92 (13)	265 (43)	127 (43)	200 (29)	101 (30)	1.548
WTP Both Cancer and Microbial V1	1187 (597)	198 (35)	667 (273)	149 (24)	404 (86)	200 (82)	293 (54)	156 (55)	1.191
WTP Both Cancer and Microbial V2	758 (278)	179 (26)	514 (162)	143 (20)	345 (60)	186 (60)	276 (44)	153 (45)	0.849
WTP Cancer Pooled	464 (124)	90 (9)	244 (48)	68 (7)	232 (28)	97 (27)	157 (16)	74 (16)	5.480
WTP Microbial Pooled	739 (223)	123 (14)	416 (101)	94 (10)	294 (37)	132 (36)	211 (24)	103 (24)	3.453
WTP Both Cancer and Microbial Pooled	922 (277)	187 (21)	739 (223)	123 (14)	370 (50)	192 (49)	294 (37)	132 (36)	1.478

Note: [a] Standard errors in parentheses.

Table 5: Likelihood Test and Wald Tests

	Lognormal				Weibull			
	Full Sample		Yea-saying Observations Removed		Full Sample		Yea-saying Observations Removed	
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
Likelihood Ratio Tests								
Cancer V1, V2 vs. Cancer Pooled	4.217		4.632		4.205		4.663	
Microbial V1, V2 vs. Microbial Pooled	1.221		0.887		0.662		0.294	
Both V1, V2 vs. Both Pooled	0.552		0.288		0.304		0.059	
Wald Tests								
Cancer V1 vs. V2	**0.327**	**3.815**	**0.787**	**4.325**	**1.362**	**0.555**	**2.232**	**0.894**
Microbial V1 vs. V2	**0.760**	**0.170**	**0.609**	**0.022**	**0.714**	**0.019**	**0.292**	**0.004**
Both V1 vs. V2	**0.426**	**0.192**	**0.234**	**0.041**	**0.316**	**0.018**	**0.059**	**0.002**
Cancer Pooled vs. Microbial Pooled	**1.171**	**4.161**	**2.387**	**4.744**	**1.710**	**0.590**	**3.685**	**0.964**
Internal Consistency Tests								
Weak Adding Up Test: Cancer Pooled vs. Both Pooled	**2.290**	**17.812**	**4.727**	**13.398**	**5.697**	**2.861**	**11.269**	**2.138**
Weak Adding Up Test: Microbial Pooled vs. Both Pooled	**0.264**	**6.472**	**1.745**	**3.167**	**1.478**	**0.991**	**3.453**	**0.465**
Microb V2 vs. Both Pooled	**1.198**	**6.510**	**2.627**	**2.790**	**2.499**	**0.996**	**3.897**	**0.434**
Strong Adding Up Test: (Cancer pooled + Microbial pooled) = Both Pooled	**0.558**	**0.938**	**0.102**	**4.592**	**5.161**	**0.309**	**2.507**	**0.971**
Related Tests:								
Cancer V1 & Both V1	**1.074**	**5.257**	**1.766**	**6.267**	**1.983**	**0.739**	**3.389**	**1.121**

Cancer V2 & Both V2	1.337	13.907	3.315	16.818	4.451	2.482	9.032	3.635
Cancer V1 & Both Pooled	1.288	8.340	3.603	5.071	2.312	1.211	5.879	0.803
Microb V1 & Both V1	0.019	2.726	0.149	3.661	0.434	0.373	1.001	0.605
Microb V2 & Both V2	0.367	3.750	0.796	4.605	1.149	0.639	2.086	0.927

External Scope Test

Cancer V1 & Both V2	0.423	4.902	1.530	6.726	1.070	0.799	3.355	1.372
Microb V2 & Both V1	1.024	4.286	1.242	4.494	2.069	0.613	2.291	0.782

Table 6: Regression Results – Weibull Distribution (Using Data that Removes Yea-saying Observations)

Variable	Cancer		Microbial		Cancer plus Microbial	
Constant	**4.348**	***	**4.252**	***	**4.564**	***
	(0.394)		(0.398)		(0.379)	
Household Income	**-5.870E-06**	**	**3.640E-06**		**1.600E-06**	
	(2.630E-		(2.770E-		(2.770E-	
Male	**-0.302**	*	**-0.260**		**-0.355**	*
	(0.182)		(0.187)		(0.190)	
Household Size	**0.135**	*	**0.108**		**0.070**	
	(0.074)		(0.078)		(0.075)	
Age 65 or Older	**0.514**	*	**0.256**		**0.061**	
	(0.272)		(0.267)		(0.264)	
English	**0.302**		**0.143**		**0.248**	
	(0.204)		(0.209)		(0.206)	
College	**0.164**		**0.540**	***	**0.525**	***
	(0.182)		(0.188)		(0.188)	
Urban	**-0.407**	**	**-0.397**	**	**-0.189**	
	(0.186)		(0.191)		(0.190)	
BelieveInformation	**0.577**	***	**0.710**	***	**0.774**	***
	(0.196)		(0.199)		(0.196)	
Do Not Avoid Tap Water	**-0.285**		**-0.337**	*	**-0.266**	
	(0.180)		(0.184)		(0.185)	
V1Q1	**0.380**	**				
	(0.176)					
V2Q1			**-0.132**			
			(0.180)			
V1Q3					**0.051**	
					(0.180)	
Scale	**1.343**		**1.330**		**1.232**	
	(0.096)		(0.095)		(0.092)	
N	363		363		363	
Log Likelihood	-432.83		-460.80		-440.59	

Notes: [a] Standard errors in parentheses.
[b] * = significant at 10% level, ** is 5% and *** is 1%.

Table 7: ABSCM: Estimated Parameters Proportional Versions

	Full Sample			Yea-saying Observations Removed		
Variable	Version 3 (proportional)	Version 4 (proportional)	Pooled Version 3 and 4	Version 3 (proportional)	Version 4 (proportional)	Pooled Version 3 and 4
Status Quo Constant	0.837* (.175)	0.618 (.129)	0.720* (.102)	0.850* (.190)	0.664* (.136)	0.759 (.109)
Microbial deaths	-0.163* (.020)	-0.156 (.013)	-0.159* (.011)	-0.154* (.022)	-0.164* (.014)	-0.163 (.012)
Cancer deaths	-0.125* (.017)	-0.122 (.013)	-0.121* (.010)	-0.121* (.018)	-0.122* (.014)	-0.120 (.011)
Program cost	-0.004* (.001)	-0.004 (.001)	-0.004* (.000)	-0.005* (.001)	-0.005* (.001)	-0.005 (.001)
Observations (choice sets)	824	800	1624	712	732	1444
Log-likelihood	-507.22	-780.19	-1288.76	-428.70	-700.68	-1131.35

Notes

[a] Microbial deaths and illnesses are proportional, statistics are reported for deaths in the model. Similarly, cancer deaths and illnesses are proportional.

[b] Standard errors in parentheses. Asterisk indicates significance at the.01 level.

[c] Test of pooling version 3 and 4, chi-squared 2.70, critical value 11.07.

[d] Test of pooling version 3 and 4, chi-squared 3.95, critical value 11.07.

Table 8: ABSCM: Estimated Mean Willingness to Pay Values

Variable	Full Sample			Yea-saying Observations Removed		
	Version 3 (proportional)	Version 4 (proportional)	Pooled Version 3 and 4	Version 3 (proportional)	Version 4 (proportional)	Pooled Version 3 and 4
Microbial deaths and illnesses (including SQ effect)[a]	202.420 (40.69)	237.820 (29.63)	219.270 (21.48)	142.580 (31.59)	196.830 (21.89)	175.400 (17.17)
Cancer deaths and illnesses (including SQ effect)[a]	101.720 (26.90)	149.290 (26.03)	123.550 (18.26)	69.401 (26.66)	113.150 (22.16)	88.427 (16.31)
Microbial deaths and illnesses (excluding SQ effect)[b]	429.660 (107.29)	396.310 (65.87)	404.230 (53.59)	321.090 (75.18)	334.740 (51.75)	332.590 (38.44)
Cancer deaths and illnesses (excluding SQ effect)[b]	326.080 (81.03)	309.800 (51.36)	306.920 (40.33)	250.300 (58.21)	247.690 (38.87)	244.150 (29.07)
Cancer deaths and illnesses (including SQ effect)[a]	528.350 (99.03)	548.450 (68.81)	524.760 (53.677)	393.590 (64.87)	445.430 (47.296)	419.130 (36.612)
Microbial and cancer deaths and illnesses (excluding SQ effect)[b]	752.56 (176.13)	709.80 (111.14)	707.100 (92.11)	575.20 (123.89)	580.88 (81.14)	574.22 (65.40)

Notes:

[a] Welfare calculations include consideration of the status quo constant.

[b] Welfare calculations do not include consideration of the status quo constant.

[c] Standard errors in parentheses, based on Krinsky Robb simulation using 1000 draws.

45

Table 9: ABSCM: Estimated Parameters Proportional Versions

Variable	Full Sample			Yea-saying Observations Removed		
	Version 5 (non-proportional)	Version 6 (non-proportional)	Pooled Version 5 and 6	Version 5 (non-proportional)	Version 6 (non-proportional)	Pooled Version 5 and 6
Status Quo Constant	0.523* (.123)	1.067* (.172)	0.728* (.097)	0.528* (.127)	1.132* (.183)	0.753* (.102)
Microbial deaths	-0.054* (.011)	-0.074* (.017)	-0.056* (.009)	-0.053* (.011)	-0.081* (.018)	-0.058* (.009)
Microbial illness	-7.621E-05* (.000)	-8.662E-05* (.000)	-8.040E-05* (.000)	-7.552E-05* (.000)	-8.769E-05* (.000)	-7.974E-05* (.000)
Cancer deaths	-0.058* (.011)	-0.046* (.015)	-0.055* (.009)	-0.048* (.011)	-0.041* (.016)	-0.045* (.009)
Cancer illness	-0.008* (.002)	-0.022* (.003)	-0.011* (.002)	-0.008* (.002)	-0.020* (.004)	-0.010* (.002)
Program cost	-0.004* (.001)	-0.006* (.001)	-0.004* (.000)	-0.004* (.001)	-0.006* (.001)	-0.005* (.001)
Observations (choice sets)	812	812	1624	740	724	1464
Log-likelihood	-786.502	-458.291	-1262.62	-716.16	-398.46	-1130.95

Notes:

[a] Microbial and cancer deaths and illnesses are non-proportional.
[b] Standard errors in parentheses. Asterisk indicates significance at the .01 level.
[c] Test of pooling version 3 and 4, chi-squared 32.62, critical value 14.45.
[d] Test of pooling version 3 and 4, chi-squared 35.65, critical value 14.45.

Table 10: ABSCM: Estimated Willingness to Pay – Non-Proportional Models

Variable	Full Sample			Yea-saying Observations Removed		
	Version 5 (non-proportional)	*Version 6 (non-proportional)*	*Pooled Version 5 and 6*	*Version 5 (non-proportional)*	*Version 6 (non-proportional)*	*Pooled Version 5 and 6*
Microbial deaths and illnesses (including SQ effect)[a]	306.180 (44.14)	163.870 (31.53)	246.110 (27.46)	288.190 (41.81)	161.780 (31.36)	237.190 (26.88)
Cancer deaths and illnesses (including SQ effect)[a]	110.440 (37.30)	81.127 (20.21)	81.927 (20.13)	80.136 (38.35)	43.977 (23.89)	40.421 (24.30)
Microbial and cancer deaths and illnesses (including SQ effect)[b]	554.580 (74.78)	422.660 (52.42)	496.570 (47.02)	501.150 (67.49)	386.240 (49.83)	447.040 (44.43)
Microbial deaths and illnesses (excluding SQ effect)[b]	443.490 (73.25)	336.750 (55.28)	415.200 (51.24)	426.580 (70.95)	343.660 (58.59)	404.61 (53.58)
Cancer deaths and illnesses (excluding SQ effect)[b]	246.539 (52.26)	255.637 (42.25)	248.650 (35.31)	214.596 (47.99)	226.663 (43.86)	209.57 (33.82)
Microbial and cancer deaths and illnesses (excluding SQ effect)[b]	690.029 (111.04)	592.387 (89.77)	663.850 (79.21)	641.176 (103.50)	570.323 (94.22)	614.180 (78.96)
Marginal value of microbial death	13.616 (3.37)	11.789 (2.99)	12.659 (2.44)	12.940 (3.32)	12.601 (3.17)	12.825 (2.50)
Marginal value of microbial illness	0.019 (0.00)	0.014 (0.00)	0.018 (0.00)	0.018 (0.00)	0.014 (0.00)	0.018 (.002)
Marginal value of cancer death	14.661 (3.02)	7.403 (2.38)	12.525 (2.10)	11.763 (2.86)	6.354 (2.46)	10.011 (2.06)
Marginal value of cancer illness	1.944 (0.62)	3.581 (0.69)	2.442 (0.47)	1.864 (0.62)	3.170 (0.68)	2.176 (0.47)

Notes:

[a] Welfare calculations include consideration of the status quo constant. Welfare measure for 10 fewer cases MICD and CAND, and 15500 cases fewer MICI, and 50 cases fewer CANI)

[b] Welfare calculations do not include consideration of the status quo constant. Welfare measure for 10 fewer cases MICD and CAND, and 15500 cases fewer MICI, and 50 cases fewer CANI).

[c] Standard errors in parentheses, based on Krinsky Robb simulation using 1000 draws.

Table 11: Wald Tests of Differences in Willingness to Pay In ABSCM Models

	Tests of Difference within Version		
	Cancer vs Microbial	*Cancer vs Both*	*Microbial vs Both*
Joint Proportional (JP)	3.367	34.924	15.687
Joint Non-Proportional (JNP)	9.47	22.19	4.82
Version 5 (V5)	6.12	13.98	2.92
Version 6 (V6)	2.55	10.93	4.17
	Tests of Difference Across Versions		
	JP vs JNP	JP vs V5	JP vs V6
Cancer	0.601	0.277	0.110
Microbial	1.193	1.357	0.025
Both	0.152	0.299	0.001
	Test of Non-proportional version sum of death and illness vs Proportional effect		
	Including status quo effect	*Excluding status quo effect*	
Microbial	3.750934	1.193	
Cancer	2.691023	0.601	

Note: Using willingness to pay measures that remove yea-saying observations and exclude status quo effects unless otherwise noted.

Table 12: Value of Statistical Life and Case Calculations from Non-Proportional Versions Based on Marginal Values (no status quo effect)

	Full Sample			Yea-saying Excluded		
	Version 5 (non-proportional)	Version 6 (non-proportional)	Pooled Version 5 and 6	Version 5 (non-proportional)	Version 6 (non-proportional)	Pooled Version 5 and 6
Microbial death	18,887,000 (4,684,700)	16,352,000 (4,198,300)	17,359,000 (3,326,200)	17,498,000 (4,510,100)	17,135,000 (4,333,800)	17,634,000 (3,585,000)
Microbial illness	26,567 (4,519)	19,103 (3,404)	24,815 (3,036)	25,188 (4,291)	18,591 (3,322)	24,013 (3,124)
Cancer death	20,157,000 (4,281,100)	10,092,000 (3,334,200)	16,980,000 (2,917,000)	16,021,000 (4,057,400)	8,538,000 (3,261,100)	13,559,000 (2,785,800)
Cancer illness	2,676,000 (876,650)	4,933,000 (992,250)	3,335,000 (652,050)	2,539,900 (903,860)	4,330,900 (943,830)	2,952,400 (624,130)

Note: Standard deviations are in brackets. Results are generated using 1,000 draws in a Krinsky-Robb procedure.

Table 13: Wald Tests of Differences in Willingness to Pay

	Tests of Difference CVM versus ABSCM			
	JP – SQ excluded vs CVM Pooled	*JP – SQ included vs CVM Pooled*	*JNP – SQ excluded vs CVM Pooled*	*JNP – SQ included vs CVM Pooled*
Cancer	6.92	9.01	1.97	16.05
Microbial	7.20	1.45	10.87	14.45

Note: Weibull distribution forms used. All measures use data with yea-saying observations removed.

Table 14: Sample Models with Demographic Interactions and Attributes

Variable	Non-Proportional, Pooled Version 5 and 6 Yea-saying Observations Removed		Non-Proportional, Pooled Version 5 and 6 Full Sample	
	Coefficient	Coefficient/St.Er.	Coefficient	Coefficient/St.Er.
SQ	0.301	0.725	0.290	0.727
MICD	-0.026	-0.605	-0.026	-0.631
MICI	-7.017E-05	-2.388	-6.232E-05	-2.215
CAND	-0.022	-0.550	-0.030	-0.757
CANI	-0.028	-3.247	-0.025	-2.971
SQ*INCM	-5.284E-06	-1.949	-5.852E-06	-2.238
MD*INCM	-2.002E-07	-0.674	7.512E-08	0.266
MI*INCM	5.286E-11	0.261	-5.176E-11	-0.269
CD*INCM	1.855E-07	0.691	1.838E-07	0.720
CI*INCM	1.141E-07	1.997	1.075E-07	1.982
SQ*MALE	0.528	2.766	0.542	2.972
MD*MALE	0.027	1.336	0.025	1.316
MI*MALE	-1.264E-06	-0.092	-4.300E-06	-0.334
CD*MALE	0.003	0.158	-0.005	-0.287
CI*MALE	-0.002	-0.434	-0.001	-0.377
SQ*HHSZ	0.086	1.074	0.110	1.479
MD*HHSZ	-0.007	-0.814	-0.011	-1.409
MI*HHSZ	-5.894E-06	-1.023	-8.080E-06	-1.523
CD*HHSZ	-0.010	-1.351	-0.005	-0.727
CI*HHSZ	0.002	1.365	0.001	0.977
SQ*ENGL	-0.233	-1.003	-0.188	-0.836
MD*ENGL	0.009	0.370	0.015	0.618
MI*ENGL	2.083E-05	1.237	2.193E-05	1.355
CD*ENGL	0.019	0.845	0.010	0.452
CI*ENGL	0.003	0.643	0.000	0.089
SQ*EDU	0.094	0.460	0.237	1.222
MD*EDU	0.023	1.055	0.006	0.313
MI*EDU	7.282E-06	0.493	4.903E-06	0.355
CD*EDU	0.013	0.684	0.007	0.368
CI*EDU	0.006	1.510	0.005	1.342
SQ*ILL	-0.018	-0.322	-0.038	-0.714

*MD*ILL*	-0.003	-0.519	-0.003	-0.543
*MI*ILL*	4.375E-06	1.099	5.065E-06	1.368
*CD*ILL*	0.002	0.347	0.004	0.763
*CI*ILL*	3.477E-04	0.301	5.473E-04	0.509
*SQ*URBAN*	0.443	2.144	0.385	1.966
*MD*URBAN*	-0.032	-1.454	-0.029	-1.376
*MI*URBAN*	-1.757E-05	-1.185	-1.370E-05	-0.986
*CD*URBAN*	-0.002	-0.083	0.003	0.139
*CI*URBAN*	0.005	1.178	0.005	1.290
*SQ*BLIEF*	0.131	0.578	-0.007	-0.029
*MD*BLIEF*	-0.018	-0.743	-0.018	-0.783
*MI*BLIEF*	-3.107E-05	-1.887	-2.753E-05	-1.732
*CD*BLIEF*	-0.038	-1.706	-0.049	-2.290
*CI*BLIEF*	-0.004	-0.845	-0.003	-0.676
*SQ*AVERT*	0.116	0.595	0.156	0.839
*MD*AVERT*	-0.008	-0.397	-0.006	-0.297
*MI*AVERT*	1.565E-05	1.123	1.308E-05	0.990
*CD*AVERT*	-0.014	-0.724	-0.012	-0.690
*CI*AVERT*	-0.001	-0.326	-0.003	-0.840
*SQ*AGE65*	-0.278	-0.963	-0.342	-1.246
*MD*AGE65*	0.025	0.812	0.026	0.902
*MI*AGE65*	0.000	-0.349	0.000	-0.131
*CD*AGE65*	-0.010	-0.350	0.004	0.157
*CI*AGE65*	-0.005	-0.718	-0.004	-0.676
BILL	-0.005	-9.068	-0.005	-9.313
Number of observations	1464		1624	
Log likelihood	-1082.634		-1206.9	

Figure 1. Percentage of "Yes" Responses by Bid Value (Pooled versions)
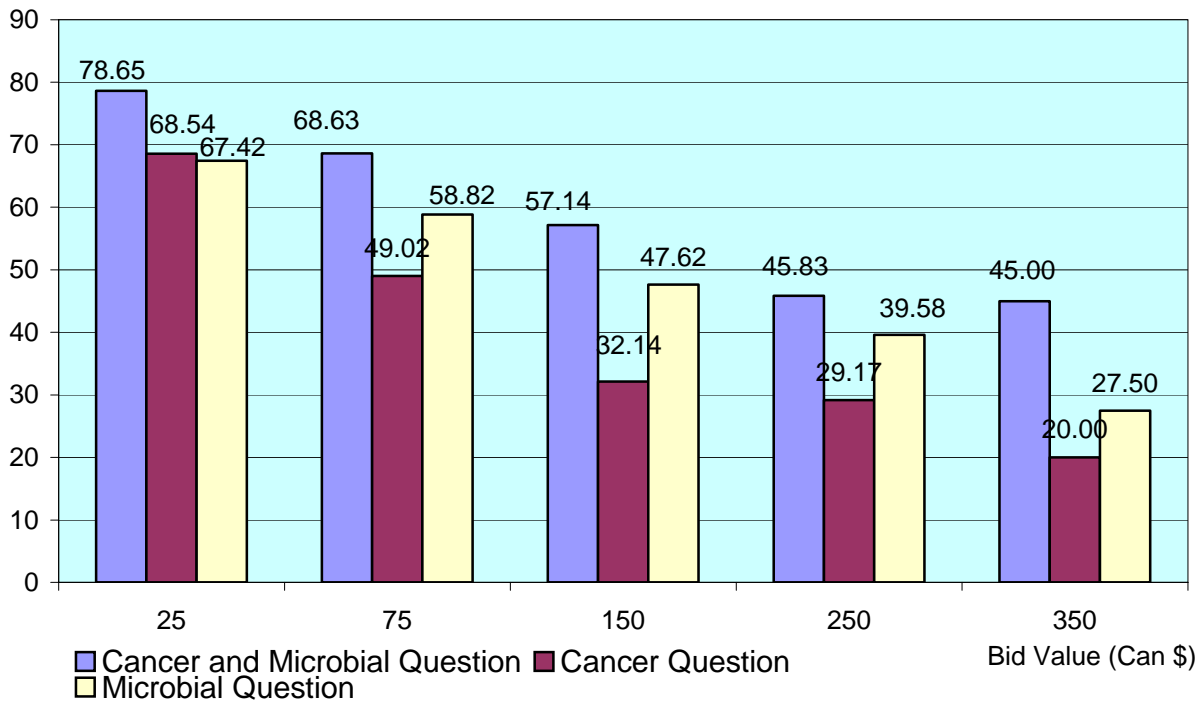
References

Abt Associates Inc. *Arsenic In Drinking Water Rule Economic Analysis* Office of Ground Water and Drinking Water. U.S. Environmental Protection Agency. USEPA 815-R-00-026. (December 2000). 257 pp.

Adamowicz, W., P. Boxall, M. Williams and J. Louviere (1998) Stated preference approaches to measuring passive use values. American Journal of Agricultural Economics (80):64-75.

Adamowicz, W., J. Louviere, J. Swait. *Introduction to Attribute-Based Stated Choice Methods*. Final Report to Resource Valuation Branch, Damage Assessment Center, NOAA, U.S. Department of Commerce. (January 1998) 44pp.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner and H. Schuman (1993)"Report of the NOAA Panel on contingent valuation" Federal Register. US Department of Commerce. 58(10), 4602-4614.

Barbeau, B.; P. Payment; J. Coallier; B. Clement, and M. Prevost "Evaluating the Risk of Infection from the Presence of Giardia and Cryptosporidium in Drinking Water" *Quantitative Microbiology* 2 (1): 37-54, March 2000.

Carson, Richard T.; Mitchell, Robert C.; Hanemann, W. Michael; Kopp, Raymond J.; Presser, Stanley; and Ruud, Paul A., *A Contingent Valuation Study of Lost Passive Use Values Resulting From the Exxon Valdez Oil Spill* (Anchorage: Attorney General of the State of Alaska, November 1992).

Carson, R. T. and R. C. Mitchell "Public Preference Toward Environmental Risks: The Case of Trihalomethanes" Department of Economics, University of California, San Diego Discussion Paper 2000-21. (August 2000). 37 pp.

Cochran, W. G., and G. M. Cox. *Experimental Designs*. New York: John Wiley & Sons. 1957.

Corso, P., J. Hammitt, and J. Graham, "Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation." *Journal of Risk and Uncertainty*. 23 (2) (2001): 165-184.

de Wit, M; M. Koopmans; L. Kortbeek; W. Wannet; J.Vinjé; F. van Leusden; A. Bartelds and Y. T. van Duynhoven "Sensor, a Population-based Cohort Study on Gastroenteritis in the Netherlands: Incidence and Etiology" *American Journal of Epidemiology* Vol. 154, No. 7 : 666-674. 2001.

Dillman, D.A. *Mail and Telephone Surveys: The Total Design Method*. John Wiley and Sons, New York. 1978.

Dillman, D.A. *Mail and Internet Surveys: The Tailored Design Method.* John Wiley and Sons, New York. 2000.

Fabian, Robert. "The Qualy Approach." In *Valuing Health for Policy: An Economic Approach*. Edited by G. Tolley, D. Kenkel, and R. Fabian.. Chicago, University of Chicago Press. 1994.

Federal-Provincial-Territorial Committee on Drinking Water. *Summary of Guidelines for Canadian Drinking Water Quality (April 2004).* Safe Environments Program, Health Canada. Government of Canada publication. http://www.hc-sc.gc.ca/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/water-eau/doc-sup-appui/sum_guide-res_recom/summary-sommaire_e.pdf.

Garber, A. C., et al. "Theoretical Foundations of Cost Effectiveness Analysis." In *Cost-Effectiveness in Health and Medicine: Report to U.S. Public Health Service by the Panel on Cost-Effectiveness in Health and Medicine.* Edited by Marthe R. Gold, et al., U.S. Department of Health and Human Services. 1996.

Haab, T. C. and McConnell, K. E. *Valuing Environmental and Natural Resources.* Edward Elgar Publishing, London. 2002.

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

Hammitt, J.K., and J.-T. Liu, "Effects of Disease Type and Latency on the Value of Mortality Risk," *Journal of Risk and Uncertainty* 28: 73-95, 2004.

Hanley N., Mourato S. and Wright R.E. (2001) Choice modeling approaches: a superior alternative for environmental valuation? Journal of Economic Surveys 15(3).

Havelaar, A., A. De Hollander, P. Teunis, E. Evers, H. Van Kranen, J. Versteegh, J. Van Koten, and W. Slob. "Balancing the Risks and Benefits of Drinking Water Disinfection: Disability Adjusted Life-Years on the Scale." *Environmental Health Perspectives.* 108(4): 315-321. 2000.

Health Canada. Cancer Surveillance On-Line. http://cythera.ic.gc.ca/dsol/cancer/index_e.html (20 March 2003).

Holmes, T. and W. Adamowicz. "Attribute Based Methods". *A Primer on the Economic Valuation of the Environment* edited by P. Champ, T. Brown and K. Boyle. Kluwer, Amsterdam. 2002.

Jardine, C.G. N. Gibson and S. Hrudey. "Detection of Odour and Health Risk Perception of Drinking Water." *Water Science and Technology* Vol 40(6) (1999): 91–98.

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

King W. "Great Lakes Water And Your Health. A Summary Of "Great Lakes Basin Cancer Risk Assessment : A Case Control Study Of Cancers Of The Bladder, Colon And Rectum" Health Canada publication. December 1995.

King W.D., L. D. Marrett. "Case-Control Study Of Bladder Cancer And Chlorination By-Products In Treated Water (Ontario, Canada)." *Cancer Causes and Control* 7 (1996): 596-604.

Krewski, D. et al."Health Risk Perception In Canada. II: Worldviews, Attitudes, And Opinions," *Human and Ecological Risk Assessment.* Vol. 1(2) (1995): 53–70.

Krewski, D. J., Balbus, D. Butler-Jones, C. Haas, J. Isaac-Renton, K. Roberts, and M. Sinclair *Managing Health Risks from Drinking Water* Commissioned Paper 7 for The Walkerton Inquiry. Queen's Printer for Ontario, Ontario Ministry of the Attorney General. 2002.

Krupnick, A., A. Alberini, M. Cropper, N. Simon, B. O'Brien, R.Goeree, and M. Heintzelman, "Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents." *Journal of Risk and Uncertainty*. 24(2002): 161-186.

Louviere, J., D. Hensher and J. Swait. *Stated Choice Methods – Analysis and Application.* Cambridge, Cambridge University Press. 2000.

McFadden, D. "Conditional Logit Analysis Of Qualitative Choice Behavior." *Frontiers in Econometrics* edited by P. Zarembka. New York, Academic Press. 1974. 105-142.

Mead, P.; L. Slutsker; V. Dietz; L. McCaig; J. Bresee; C. Shapiro; P. Griffin, and R. Tauxe "Food-Related Illness and Death in the United States" *Emerging Infectious Diseases* 5(5):607-625. 1999

Mills, Christina J., Bull, Richard J., Cantor, Kenneth P., Reif, John , Hrudey, Steve E., Huston, Patricia and an Expert Working Group "Workshop Report Health Risks of Drinking Water Chlorination By-products: Report of an Expert Working Group" *Chronic Diseases in Canada* 19(3).

Mitchell, R.C. and R.T. Carson. *Using surveys to value public goods : the contingent valuation method*. Resources for the Future, Washington, DC. 1989.

Payment, P., Berte, A., Prévost, M., Ménard, B., and Barbeau, B. "Occurrence of pathogenic microorganisms in the Saint Lawrence River (Canada) and comparison of health risks for populations using it as their source of drinking water" *Canadian Journal of Microbiology* 46 (2000):565-576.

Payment, P., Franco, E., Siemiatycki, J., Richardson, L., Renaud, G., Franco, E., and Prévost, M. "A Prospective Epidemiological Study Of Gastrointestinal Health Effects Due To Consumption Of Drinking Water. *Int. J. Environ. Health Res*. 7 (1997): 5–31.

Payment, P., Richardson, L., Siemiatycki, J., Dewar, R., Edwardes, M., and Franco, E. "A Randomized Trial To Evaluate The Risk Of Gastrointestinal Disease Due To The Consumption Of Drinking Water Meeting Currently Accepted Microbiological Standards" *Am. J. Public Health* 81 (1991): 703–708.

Payment, P. and M. Riley "Resolving the Global Burden of Gastrointestinal Illness: A CALL TO ACTION" American Academy of Microbiology (2002)

Ronchi, E. and S. Wald "New molecular technologies for safe drinking water" *OECD Observer* No. 215 - January 1999.

Torrance, George W."Measurement of Health State Utilities for Economic Appraisal: A Review." *Journal of Health Economics*. 5(1)1986: 1-30.

United States Environmental Protection Agency *Alternative Disinfectants and Oxidants*. April 1999.

J. Wheeler; D. Sethi; J. Cowden; P. Wall; L. Rodrigues; D. Tompkins; M. Hudson, and P. Roderick "Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance" BMJ 1999;318:1046-1050 (17 April). 1999.

Appendix 1

Below are the descriptions of the health effects from microbial illnesses and bladder cancer illnesses relating to the drinking of tap water presented in the survey.

Ipsos i-Say

E-mail: questions@i-say.com   Phone: 1-866-893-1188

Health Effects of Chlorine

When tap water is disinfected with chlorine, various by-products including Trihalomethanes (THMs) are produced. Scientists believe that THMs are an indicator for substances in the tap water that are linked to increased cases of bladder cancer when water is consumed over long periods of time.

♦ Symptoms of bladder cancer:
 - Urgent and frequent need to urinate, blood in your urine, pain during urination, and pain from the tumour.
 - Symptoms for this cancer do not occur immediately after drinking tap water, rather they take years to show since it takes years for this cancer to develop.
 - For about one in five cases, death occurs **within five years** from diagnosis.

♦ Medical Treatment of Illness:
 - Surgery, radiation, and chemotherapy are used to treat bladder cancer.
 - Side effects from surgery may include a long recuperation period and the need for colostomy (bag for body wastes).
 - Side effects of chemotherapy include loss of hair, change in taste or smell, mouth sores, possible loss of fertility, fatigue and less ability to deal with infections.

♦ Sensitive Groups:
 - Occurs most frequently in male smokers over the age of 70, but other older people can also get this cancer.

♦ Tap Water Treatment:
 - Providers of tap water can lower the chlorine levels in the water supply.
 - Less chlorine lowers cancer risks but raises microbial risks.
 - More expensive water treatment technologies are available to reduce both cancer risks and microbial risks.

Health Effects of Microbes and THMs in Tap Water

You won't need to remember these numbers. We just want to give you some idea of the risks people face.

First we list effects from all causes, then we list effects from drinking tap water only.

Microbial Health Effects in Numbers	**Cancer Health Effects in Numbers**
From all causes of microbial disease	**From all causes of cancer**
♦ Scientists estimate that for every 100,000 people:	♦ Scientists estimate that for every 100,000 people:
- Over a 35-year period, microbes from all sources (food, tap water and direct contact such as swimming), lead to 2.5 million cases of microbial infection. This means that a person may likely suffer multiple episodes of microbial illness over this period.	- Over a 35-year period, 27,000 people will contract cancer of all types.
- Over a 35-year period, about 100 deaths occur from microbes from all sources.	- Of these 27,000 people, 7,000 deaths are due to cancer of all types.
From drinking tap water	**From drinking tap water**
♦ Scientists estimate that for every 100,000 people drinking tap water:	♦ Scientists estimate that for every 100,000 people drinking tap water:
- Over a 35-year period, 23,000 people will get some sort of microbial infection.	- Over a 35-year period, 100 people will contract bladder cancer.
- Of those infected, 15 will die over the 35-year period. Death often occurs soon after infection.	- Of these, approximately 20 persons will die within 5 years as a direct consequence of the cancer.
	- Out of the 80 who do not die, some will be fully cured; others will experience cancer symptoms, and require medical interventions and drugs over their remaining lifetime.

This information is summarized in the following screen.

Sources for Health Effects Estimates

[‹‹] [››]

For a community of 100,000 people, over a 35-year period, illnesses and deaths from microbial disease and cancer will be approximately...

	MICROBIAL DISEASE		CANCER	
	Illnesses	**Deaths**	**Illnesses**	**Deaths**
From all Causes	2,500,000	100	27,000	7,000
	⬇	⬇	⬇	⬇
From Drinking Tap Water	23,000	15	100	20

On the next two screens, this situation is shown with pictures.

‹‹	››

60

Appendix 2: Example of CVM Question Format (Version 2)

Ipsos i-Say

E-mail: questions@i-say.com   Phone: 1-866-893-1188

Here's the first program we want you to vote on.

THE BENEFITS OF MUNICIPAL WATER TREATMENT PROGRAM A

Based on current water drinking patterns in your community this program would have the following benefits to every 100,000 people:

- 15,500 fewer people will develop microbial illness over a 35-year period. Another way to say this is that the average person in a community of 100,000 people will see their risk of getting microbial illness from drinking the water fall from 23,000 in 100,000 to 7,500 in 100,000
- With fewer people developing microbial illness, 10 fewer people will die from getting the disease. Another way to say this is that the average person in this community will see their risk of dying from microbial illness reduced from 15 in 100,000 to 5 in 100,000
- Bladder cancer illness and deaths will not be affected by the program

Here is a table showing these benefits:

For every 100,000 people, the NUMBER who would...	CURRENT SITUATION	PROPOSED PROGRAM A
Get **sick** from **microbial illness** in a **35-year** period	23,000	7,500
Die from **microbial illness** in a **35-year period**	15	5
Get **sick** from **bladder cancer** in a **35-year** period	100	100
Die from **bladder cancer** in a **35-year period**	20	20

Out of 100,000 people...

- ▢ People who would get microbial illness
- ▪ People who would get bladder cancer
- · People who would die from microbial illness or bladder cancer
- ▢ Remaining population

THE COST OF THE MUNICIPAL WATER TREATMENT PROGRAM A
If the majority of voters support this program your household will share in the cost starting January 2005 by paying an additional amount on your household water bill.

PLEASE VOTE NOW:

CVM21 If the estimated addition to your household's water bill was **$25 per year ($2.08 per month)** starting in January 2005, and a vote were held today, would you vote **FOR** or **AGAINST** the proposal?

◉ FOR
○ AGAINST

`‹‹` `››`

61

Appendix 3: Example of ABSCM Question Format (Version 5)

Endnotes

[1] The chlorine demand of the water is defined as: the amount of chlorine that reacts with the other chemicals in the water plus the amount required to achieve disinfection. In addition, however, utilities add extra chlorine added to the water to account for length of time in the distribution network. This is called free chlorine and is the culprit in the production of disinfection by-products such as Trihalomethanes.

[2] Two epidemiological studies suggest that drinking water from water treatment plants following standard treatment processes could be responsible for half of the cases of gastrointestinal illnesses in the receiving population (Payment et al. 1991, 1997).

[3] The typical range of annual household water bills in Canada is between $300 and $500.

[4] Attribute levels for microbial illnesses were 7500, 15000, 23000 and 30000. Attribute levels for microbial deaths were 5,10,15 and 20. Attribute levels for cancer illnesses were 50,75,100 and 125. Attribute levels for cancer deaths were 10,15,20 and 25. All were defined for a population of 100.000 and over a 35 year period. Annual increases to household water bills ranged between $25 and $350.

[5] $14.4 million = $535/50 cases*100,000*35 years/2.6 persons per household.

Trish Hall's Discussant Comments for
Session IV: Valuing Morbidity and Mortality: Drinking Water

- **Policy Implications: Adamowicz, Dupont, Krupnick**
    - ○ Preliminary Conclusion:
        - ▪ Higher VSL values for microbial illness
        - ▪ Significant value for illness avoided
        - ▪ Altruism: results can't be used at this point
    - ○ Interpretation:
        - ▪ Canadians more aware of waterborne disease outbreaks?
            - • Was also surprise by this result but agreed with the authors that Walkerton etc… may have had an impact
            - • Impacts on kids/sensitive sub-populations
                - ○ Did folks know from outbreaks that these groups are more adversely impacted by microbial illness?
            - • Bladder cancer description
                - ○ Average age of onset described as 70 years old
        - ▪ WTP values for illness very useful
            - • WTP for Mircrobial illness was significant but did not seem unreasonable but will be heavily scrutinized if we were to use
        - ▪ Combined case avoided valuations could also be beneficial
            - • Combined value for per case avoided would avoid the need to estimate mortalities and also severities
        - ▪ Altruism: Can it be sorted out
            - • But how to tease out…almost everyone drinks from a public supply at some point so how do we figure this out (Shaw brings this up)
    - ○ Points to Consider
        - ▪ Canada vs. USA
            - • Make sure tables and text clearly indicate CAD $
            - • Benefit transfer issues
        - ▪ Description on TTHM impacts could perhaps change results
            - • Would results be different with these additional descriptors?
                - ○ Routes of exposure: dermal and inhalation
                - ○ Other health impacts: other cancers and repro and developmental
                - ○ Exposure varies through out distribution system
            - • However, it could make it more difficult to conduct benefit transfer to other contaminant where this is not the case (such as arsenic).
        - ▪ Latency vs. Cessation lag (I will talk about this at the end)

- **Policy Implications: Shaw et al.**
  - Conclusion: None yet but…
    - Potentially useful for understanding this issues…even if valuation proves elusive
  - Potential implications
    - Averting behavior: does it also correlate with greater WTP?
      - Averting behavior:? If valuation can be obtained, do the results match those in the Adamowicz paper?
        -Do folks who take averting action also have higher WTPs?
    - Altruism vs. self-interest: benefits from transient water supply regulation for chronic contaminants
      - Currently, the Safe Drinking Water Act exempts transient water supplies (e.g. restaurants, truck stops) from chronic contaminant regulation (such as arsenic). Are folks concerned about these outside the home exposures even if they have their own well?
    - How can we improve risk communication?
      - Focus group shows that we need to work on putting risk in context
    - Potential valuation estimates
      - Would always be welcome

- **Points to Consider: Shaw et al.**
  - Addressing ambiguity in risk
    - Some Clarity: Do you use tap water? --very good that researchers clarified uses for cooking, making ice, etc… many people don't realize how much they actually ingest.
    - BTW: fountain sodas are also made with tap water!
  - Lots of new arsenic risk information that could address some ambiguity
    - Arsenic inhibits DNA repair
      - Arsenic may be a "promoter" of cancer and prevent the body from making repairs to damaged DNA
      - Would describing this process help people understand the risk?
    - *In utero* Arsenic exposure and lung disease
      - UC Berkeley: Arsenic Health Effects Research Program (in utero study)
  - Sources of data
    - Would not recommend using Burnett/Hahn report for benefit estimates or risk data
      - many inaccuracies regarding the Arsenic Rule
  - Latency vs. Cessation Lag
    - Same issue as with the Adamowicz paper
    - Agency prefers cessation concept
    - Some examples of it use can be found here:

- See SAB report and Stage 2 DBPR EA for more information
- http://www.epa.gov/sab/pdf/ec01008.pdf
- http://www.epa.gov/OGWDW/disinfection/stage2/regulations.html
- I'll briefly describe the differences between latency and cessation next

- **Note about Cessation Lag**
    - Outlined in EPA's Science Advisory Board's Arsenic Rule Benefits Review Panel
        - Benefits analysis based only on latency greatly underestimates actual benefits
        - A good example of this is smoking:
            - Latency: initial exposure and increase in lung cancer risk is ~ 20 years
            - Cessation: risk of lung cancer declines quickly with reduced exposure
            - Smoking probably both imitator and promoter of carcinogenic effects.
                - Promoters should see more rapid decline in risk (i.e. late stage actor not the one that started the problem)
                - Arsenic seems to be a promoter
                    - Perhaps does not cause DNA damage but inhabits DNA repair
            - The Final Stage 2 DBPR expands on the work of the SAB and includes cessation models for: smoking/lung cancer, smoking/bladder cancer, and arsenic/bladder cancer.

Valuing Reductions in Health Risks from Drinking Water: Discussion

Gregory L. Poe
Associate Professor
Department of Applied Economics and Management
Cornell University
GLP2@cornell.edu

It is a distinct pleasure to participate in this workshop, and to have the opportunity to focus my attention on a group of research efforts directed toward exploring methods of conceptualizing and measuring the economic benefits of reducing health risks from drinking water.   Individually and collectively the presentations in this session meet what I view to be the objective of EPA, and more specifically EPA STAR, funded research and collaboration: to make methodological contributions while remaining policy informative.

While I enjoyed and learned from each of the four presentations in this session, and appreciate that they are at varying levels of completion, I have been asked to center my present discussion on the presentation by Vic Adamowicz, Diane Dupont, and Alan Krupnick (hereafter ADK). Of the four research efforts comprising this session, the work by this research group is the furthest along and the only one in a position to provide a manuscript to accompany the oral presentation.

Although it is not funded through the STAR program, the ADK research is clearly in the spirit of EPA STAR objectives ascribed above. ADK does offer a methodological contribution to a contemporary debate in non-market valuation by comparing willingness-to-pay value estimates obtained from a contingent valuation (CV) study with those obtained from an Attribute Based Stated Choice Method (ABSCM). More colloquially this latter method is referred to a choice modeling or a variant of conjoint analysis. Although the research was conducted in Canada, ADK's findings are relevant to water quality policy in the United States. The tradeoff between microbial contamination and the cancer risks associated with byproducts of chlorination (i.e., Trihalomethanes) are fundamental to the Surface Water Treatment Rule, the Disinfectant/Disinfection Byproducts Rule, and the Groundwater Rule (http://www.epa.gov/safewater/dwa/electronic/ematerials.html#npdwr). The apparent high quality of this research suggests to me that KDM will make a notable and lasting contribution to both the literature on research methods and applied policy analysis.

The remainder of my comments is organized around central themes raised in ADK. With an eye toward addressing ADK's (p. 33) expressed concern that the value of statistical lives (VSL) that they find in their research "falls in the upper range of [previously estimate VSL] value," and the "fact that VSL for deaths from microbial disease is somewhat greater than that for cancer is a big surprise", the following sections discuss issues related to risk communication, the valuation of private versions public risks, and ADK's design and comparisons of stated preference methods.

Risk Communication:

Communicating drinking water risks in a manner that induces reasonable protective behavior when appropriate and reasonable inaction when exposure levels are well within safety levels is not a simple task. For instance, a recent arsenic risk communication study that endeavored to bring together concepts "of information processing, mental models and health behavior" into a single model of health behavior theory identified 45 possible variables in the path from arsenic exposure level to protective behavior (Severtson, Baumann, and Brown).

Economists, however, are more parsimonious in their characterization of risk updating with respect to new information. One such model treats an individual's subjective posterior risk assessment (R_P) as a function of prior risk perceptions (R_0) and the subjective risk associated with the information message (R_I) (see Smith and Johnson). A simple form of this relationship, which is consistent with many updating models, is a weighted linear average:

$$R_p = w_0 R_0 + (1 - w_0)R_I$$

where w_0 is the weight placed on the prior risk perceptions. In turn 'I' contains general information about contaminants and their effects and exposure information. Past research using this simple updating framework has demonstrated that in making informed risk assessment, individuals place significant weight on both prior perceptions and new information for various health risks (e.g. radon, Smith and Johnson; chemical labeling, Viscusi and O-Connor; nitrates in groundwater, Poe and Bishop).

The above relationship has implications for ADK's analysis and conclusion. Of overarching importance, it implies that $R_p \neq R_I$. Related to this is the supposition that individuals likely have, and place weight on, prior perceptions of exposure and health risks from drinking water in characterizing their reference risk. Hence the respondents' subjective assessment of how the proposed program would affect the risk that they (individually or collectively) face will typically not align with the "objective" change presented (and modeled) in the research.

I posit that these implications shed light on ADK's "upper range" VSL finding indicated previously. Specifically, prior perceptions of health effects may be artificially large because of high profile microbial contamination events in Walkerton Ontario and North Beettleford, Saskatchewan (p. 3, p. 33, ADK). If R_p is elevated relative to the exposure and risk information provided by the researchers, and supposing that respondents take the target exposure level at face value, then respondents will be valuing a larger change than indicated. Dividing this larger value by the smaller change in "objective" risk conveyed in the survey materials would engender upwardly biased VSL estimates.

Whilst I find it innovative, I worry too that the "snake in the sand" communication approach that presents both microbial and cancer risks in the same diagram could lead to disproportional focus on the change in microbial risk relative to that associated with

cancer. In examining the question formats in Appendices 2 and 3, I was taken by the fact that microbial exposure risks were, in essence, represented by an area, and cancer risks by a line. Although the changes in risks are proportional, to me the change in area associated with microbial risks loomed much larger. Should this optical "illusion" carry over to respondents, it would cause a further deviation between the change in objective risks communicated in the survey and the subjective risks utilized by the respondents.

In identifying these issues of subjective and prior risks, I do not mean to imply that ADK somehow failed in their efforts to communicate risk and risk changes. Indeed, I would argue the opposite. I am genuinely impressed with ADK's efforts to accurately understand and communicate the risks facing individuals, and would rate their work quite high relative to previous valuation work in groundwater risks. Nevertheless, I do believe that more could (can still?) be done with respect to understanding the subjective risks that individuals used as a base for formulating their willingness-to-pay values. Enhanced understanding of subjective risk, perhaps gleaned from a much smaller, shorter follow up survey or other auxiliary information, would provide an informative step toward better understanding the reported values and their relationship to prior work on groundwater and more general VSL studies. It is in this area of understanding what the respondents are valuing that I particularly commend the preliminary work presented by Douglass Shaw in this same session.

Altruism and Public Values:

ADK are correct in highlighting the fundamental difference between private and public valuation exercises and its impact on how we are to interpret value estimates, particularly with respect to comparisons with VSL estimates. Whereas groundwater quality is a public good, "best" estimates of the value of a statistical life derive largely from individual choices made in wage or market place studies (although CV and averting behavior studies have also been conducted and utilized in VSL estimates). As one moves from the private to public arena, other-regarding preferences enter into an individual's valuation equation, leading, potentially, to incomparable value estimates between public and private risk valuation exercises.

While fairness, reciprocity and other concerns are key elements of the set of other-regarding behaviors, ADK limit their concerns to "elements of altruism". That altruistic preferences are a concern in the valuation of safety is made evident in Viscusi, Magat and Forrest's work which compared willingness-to-pay values for personal risk reductions with willingness-to-pay values for programs that reduce the risk to others. They report that the sum of altruistic values for the risk reductions of other individuals are as high as six to seven times the value of reduction placed on an equivalent reduction in individual personal risk.

Economists have classified at least three types of altruistic preferences, each with a differing economic-theoretic role in benefit-cost analysis. The first is deemed "pure" altruism, reflecting the fact that I care for the well being or utility of others (Bergstrom, 2006). A second form is paternalistic altruism, which refers to the fact that I derive

utility from how you consume (eat your peas! and don't take drugs!) and derive your utility (Jones-Lee, 1992). The third is Andreoni's "impure" altruism (or warm glow giving) in which I derive egoistic utility simply from the act of giving, independent of the particular good in question (Andreoni). ADK's paper implies that they interpret economic-theoretic benefit discussions of the role of the various forms of altruism in welfare assessments to imply that it is appropriate to "purge our estimates" (p. 35) of impure and pure altruistic motives[1]. But to the extent that altruistic preferences are paternalistic or safety oriented, they should be accounted for in benefit-cost analyses of risk reductions. I concur with this assessment.[2]

I do, however, dispute ADK's interpretation that pure altruism necessarily inflates values relative to private values.   As Bergstrom (2006) reminds us "we should not forget… to count sympathetic losses each bears from the share of its costs paid by the other" (p. 339). The potential for such costs is of particular concern in the discrete choice framework employed in the stated preference elicitation formats utilized in ADK. Johannesson *et al*. argue that the coercive nature of voting and taxation raises the possibility that some people who are pure altruists will vote "no" on a project that would provide them private net benefits for risk reduction, narrowly defined, because they desire not to impose costs on others for whom costs exceed the benefits.

> Let us assume that [an individual] is willing to pay $t for a ceteris paribus increase in his own safety. His total WTP for a uniform public risk reduction of the same magnitude will fall short of $t if he believes that others are willing to pay less than $t but will still be forced to pay that amount ($t) for the project. This is because other individuals, for whom he cares will experience a lower utility if the program is implemented. In turn, this decrease in the utility of others reduces the pure altruist's WTP for the public safety project. (p. 264)

In other words, purely altruistic behavior may in some instances *lower* the proportion of affirmative votes relative to a self-interested model. Johanneson *et*

---

[1] There is continuing debate in the economic literature regarding the role of pure altruism in benefit-cost analyses. Conventional economic wisdom suggests that the optimal provision of public goods should be based solely on selfish preferences (Bergstrom, 1982; Jones-Lee, 1991, 1992; Milgrom; Johansson) in social benefit-cost analyses for small projects evaluated close to a social welfare optimum. However, as Flores notes, public projects are rarely, if ever, financed under such conditions: most typically the funding for specific public projects imposes coercive costs that result in utility gains and losses. Moreover, projects evaluated tend to be discrete, and the initial allocation of public goods is inefficient. Under these conditions the extrapolation of Bergstrom's (1982)result for marginal changes at the optimum do not carry over to the "more modest problem [of benefit-cost analysis], determining whether a specific project can lead to a Pareto improvement" (Flores, p. 304). While Bergstrom (2006) does not dispute Flores' argument he concludes that "for a broad class of economies, a comparison of the sum of private values to the cost of a project is the appropriate test for determining whether it can lead to a Pareto Improvement" (p. 348).

[2] In the spirit of full disclosure, I should not that Professor Richard C. Bishop, an attendee at this conference, Professor Emeritus at the University of Wisconsin, longstanding leader in non-market valuation research, and the Chair of my dissertation committee indicated, after my presentation, disagreement with the exclusion of warm glow giving from benefit-cost analyses. At the time of this writing we have not yet had the opportunity to determine our point of departure on this issue.

al. demonstrate this outcome in a dichotomous choice contingent valuation study of safety. In an experimental study of willingness to pay for protection against financial risks in coercive tax settings, Messer, Poe and Schulze further demonstrate this result.

With respect to warm glow, I agree with ADK that warm glow should be removed from value estimates for use in benefit-cost analyses, but disagree with their method of doing so. To isolate warm glow respondents in the CV format, ADK "removed people who said that they would pay anything for health risk reductions and who answered Yes-Yes" (p. 20: for ABSCM they simply removed individuals who said that they would pay anything). These types of people are best categorized as yea- sayers, not warm glow respondents. Warm glow need not be large. And it could be a small or large element of every respondent's values. Hence, it appears the removal of selected yea-sayers from the data set bears little relation to removing warm glow values from the entire data set of respondents.

In sum, I concur with the intent of the last sentence in ADK's paper, "Attempts to purge our estimates of altruism effects are in our plans for future research," and heartily urge the authors to undertake this effort. In doing so, however, they must take care to do so in a manner consistent with the underlying economic-theoretic construct.

The CV and ABSCM studies:

Overall the survey implementation and the analyses seem to be, as already suggested, of high quality (as I would expect from this set of co-authors). My comments on the survey design tend to be of a more specific rather than general nature, and hence, I shall rely on a bulleted format to convey my impressions.

- Both modes: The 46% response rate is relatively low by contemporary stated preference standards for established methods of survey research such as mail, telephone or in-person contacts. Web-based survey research is still fairly nascent and it is not clear at this time what response rate expectations and non-response implications for this mode. Nevertheless, it is a concern for any policy research when the response rate falls below 50%.
- CV format:
  - Valid comparisons of adding up should account for the likely positive correlation in DC-CV responses (or more specifically error terms) across risk scenarios. Failing to account for this difference will lead to biased estimates of the significance of the difference between the parameters being compared (see Poe, Welsh and Champ)
  - There appears to be a "fat tails" problem at upper bids (45% yes to the joint treatment of Cancer and Microbial) which should be accounted for/addressed in estimating the mean WTP values.

- ABSCM format:
  - As ADK note, a fundamental question arises with the observation that the inclusion/exclusion of the status quo (a difference on the order of 50%). This is a concern, in part, because the authors provide little guidance about which of the two measures is appropriate.
- Comparing CV and ABSCM:
  - ADK note that the "WTP measures from the [ABSCM] models with the status quo effect included tended to be lower that the CVM values while the WTP form the models with the status quo effects excluded tended to be higher than the CVM models" (p. 32). Either result is of interest as well as of concern. The former result is of interest because it is not consistent with previous comparisons of CVM and ABSCM that have found that ABSCM values are not significantly different or are statistically higher than CVM (see Boyle, Morrison and Taylor). In contrast the latter results are consistent with the previous literature, but that is a concern. Here, ADK use a dichotomous choice CV format, which has been demonstrated to engender the highest deviations between hypothetical and actual values in simulated market studies (e.g., Brown *et al.*). It would thus be disappointing to find that ABSCM provides higher values than the most upwardly biased CV format.

Concluding Thoughts:

My sense is that ADK have designed a study that provides one of the fairest and competent comparisons of CV and ABSCM. I believe that this research will make a notable contribution to the stated preference literature. The research also has high potential for informing policy. As I see it the only shortcoming of this research is that, I suspect, the present statistical analyses are far from final and that there are several issues, some of which I have raised above, that merit closer consideration as this research is brought to completion. I do look forward to reading revised and updated analyses of this work, and maintain that the EPA and the other agencies that have funded this research have made a solid investment that will, in time, make a lasting contribution to the dual objective of policy and methods in the valuation of drinking water risks.

References:

Adamowicz, V. D. Dupont, and A. Krupnick, 2006. "Willingness to Pay to Reduce Community Health Risks from Municipal Drinking Water: A Stated Preference Study." Conference Draft for US EPA NCER/NCEE Workshop on Morbidity and Mortality: How Do We Value the Risk of Illness and Death? Apr. 10-12, Washington D.C.

Andreoni, J., 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving," *Economic Journal* 100: 464-467.

Bergstrom, T.C., 1982. When is a Man's Life Worth More Than His Human Capital?" in M. Jones-Lee ed. *The Value of Life and Safety*. Amsterdam:North-Holland, pp. 363-398.

Bergstrom, T.C., 2006. "Benefit-Cost in a Benevolent Society." *American Economic Review*, 96(1):339-351.

Boyle, K., M. Morrison, and L. Taylor, 2004. "Why Value Estimates Generated Using Choice Modeling Exceed Contingent Valuation: Further Experimental Evidence." Paper Presented at the Australian Agricultural and Resource Economics Society Conference, Feb. 11-13, Melbourne.

Brown, T.C., P.A. Champ, R.C. Bishop, and D.W. McCollum, 1996. "Which Response Format Reveals the Truth about Donations to a Public Good?" *Land Economics*, 72:152-66.

Flores, N.E., 2002 . "Non-Paternalistic Altruism and Welfare Economics." *Journal of Public Economics* 83(2):293-305.

Johannesson, M., P.-O. Johansson, and R.M. O'Conor, 1996. "The Value of Private Safety Versus the Value of Public Safety," *Journal of Risk and Uncertainty* 13:263-275.

Johansson, P.-O. ,1992. "Altruism in Cost-Benefit Analysis." *Environmental and Resource Economics* 2: 605-613.

Jones-Lee, M.W., 1991. "Altruism and the Value of Other People's Safety." *Journal of Risk and Uncertainty* 4(2):213-219.

Jones-Lee, M.W., 1992. "Paternalistic Altruism and the Value of Statistical Life." *Economic Journal* 102(410):80-90.

Messer, K.D., G.L. Poe, and W.D. Schulze**. 2006. "**The Value of Private Risk Versus the Value Of Public Risk: An Experimental Analysis." In K.D. Messer, L.E. Platt, G.L. Poe, D. Rondeau, W.D. Schulze, and C.A.Vossler AEM Research Bulletin – 06-02 Cornell University, *Final Report: March 31, 2006 Ecosystem Values and Surface Water Protection: Basic Research on the Contingent Valuation Method NSF GRANT SES – 0108667*.

Milgrom, P. (1993). "Is Sympathy an Economic Value? Philosophy, Economics, and the Contingent Valuation Method." In J. A. Hausman, ed., Contingent valuation: A Critical Assessment, Amsterdam, North-Holland, pp. 417-442.

Poe, G.L, and R.C. Bishop, 2001. "Chapter 3: Information and Contingent Values for Groundwater Protection: A Case Study of Wisconsin." In J.C. Bergstrom, K.J. Boyle, and G.L. Poe eds. *The Economic Value of Water Quality*, Cheltenham: Edward Elgar Publishers.

Poe, G.L., H.M. van Es, T.P. VandenBerg and R.C. Bishop, 1998. "Do Households Update Their Health Risk and Exposure Perceptions with Water Testing?" *Journal of Soil and Water Conservation* 53(4):320-325.

Poe, G.L., M.P. Welsh, and P.A. Champ, 1997. "Measuring the Differences in Mean Willingness to Pay when Dichotomous Choice Contingent Valuation Responses Are Not Independent." *Land Economics*, 73(2):255-267.

Severtson, D.J., L.C. Baumann, and R.L. Brown, 2006. "Applying a Health Behavior Theory to Explore the Influence of Information and Experience on Arsenic Risk Representations, Policy Beliefs and Protective Behavior." *Risk Analysis* 26(2): 353-368.

Smith, V.K., and F.R. Johnson, 1988. "How do Risk Perceptions Respond to Information? The Case of Radon." *Review of Economics and Statistics* 70(1):1-9.

Viscusi, W.K., W.A. Magat, and A. Forrest, 1988. "Altruistic and Private Valuations for Risk Reduction." *Journal of Policy Analysis and Management* 7(2):227-245.

Viscusi, W.K., and C.J. O'Connor, 1984. "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers." *American Economic Review* 74:942-956.

**Questions and Discussion section
was not conducted for Session VI**