June 1999

VALUING HEALTH FOR ENVIRONMENTAL POLICY WITH SPECIAL EMPHASIS ON CHILDREN'S HEALTH ISSUES

PROCEEDINGS OF THE SECOND WORKSHOP IN THE ENVIRONMENTAL POLICY AND ECONOMICS WORKSHOP SERIES

-- Session Two--

A Workshop sponsored by the
US Environmental Protection Agency's Office of Children's Health Protection, Office of Economy and Environment, and Office of Research and Development

Edited by Shi-Ling Hsu
Environmental Law Institute
1616 P St. NW, Washington, D.C. 20036

Silver Spring Holiday Inn
Silver Spring, Maryland
March 24-25, 1999

Acknowledgements

Sections of this report indicated as "summarizations" were prepared by the Environmental Law Institute with funding from the Office of Research and Development, U.S. Environmental Protection Agency. ELI wishes to thank Ainsley Caldwell and Stephen Lingle of EPA's Office of Research and Development, and project officer Alan Carlin together with Nicole Owens and Tracey Wolff of EPA's Office of Policy. ELI would also like to thank the researchers who provided the material for the report.

Disclaimer

Although the information in this document has been funded in part by the United States Environmental Protection Agency under Cooperative Agreement CR-822795-01 to ELI, it may not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Proceedings for Session Two, Wednesday, March 24, 1999

--Table of Contents--

Introductory Remarks for Session II

by Chris Dockins, US EPA Office of Economy and Environment

I intend to make these introductory remarks rather brief. I have never been to a workshop, conference session, or seminar where people have left saying to themselves: "That was good, but I wish they had spent more time on the introductory comments."

My first thought when I was asked to introduce this session and to serve as moderator was "why me?" You may be wondering the same thing. I confess to not being heavily published in this field. I have, however, spent a substantial portion of my time over the last two years or so writing on the subject of health valuation for environmental policy. Melonie Williams – who offered the introduction and policy discussion for the morning session – Melonie and I work in the same office in EPA's Office of Policy. Our jobs, and those of our colleagues involve providing support to analysts elsewhere in the agency as they perform economic analyses. This is the context in which I and others at EPA have done a great deal of our writing: memoranda, explanatory notes, evaluating or proposing approaches for benefit analysis, reviews (by request) of ongoing benefits analyses. A major emphasis in this work is helping to apply existing valuation literature to the analysis of environmental policy actions.

There is no need for me to remind everyone here that EPA is facing new challenges in the economic analysis of health benefits, due in part to statutory requirements and other mandates. But a brief description of some of the Agency's efforts in support of the economic analysis of non-fatal health effects might be productive. My focus is on using values from revealed and stated preference methods rather than cost-of-illness figures. Cost-of-illness estimates still play a prominent role in our economic analyses for a number of reasons, but we recognize that estimates of willingness-to-pay are generally preferred. My purpose here is to provide some context for the "for Environmental Policy" portion of the workshop title, not to review many of the more technical issues surrounding morbidity risk valuation.

To begin with, I want to provide one illustration of how the agency has responded to increasingly available literature on willingness to pay values for non-fatal health effects. I admit that I am providing a time series sample of one, but the illustration, I think, is useful.

In 1983, EPA performed an extensive benefit analysis of the health effects from proposed new National Ambient Air Quality Standards. The benefit analysis looked at acute and chronic morbidity effects in addition to premature fatalities. Although the analysis clearly indicated, a preference for WTP value estimates for these effects, Volume II of the report notes specifically that there were no existing applicable values available in the literature. This limited the analysis to a second-best approach based on lost earnings, lost non-work time, and direct medical expenditures. The authors recognized that cost-of-illness estimates provide only a rough lower bound of benefits, and they performed plausibility checks to better understand the implications of this approach.

Compare this to work used in the 1997 EPA report *The Costs and Benefits of the Clean Air Act.* (Similar work was used in the benefit analysis of the 1997 NAAQS for particulate matter and ozone). Setting aside differences in available concentration-response and risk data, this report monetizes changes from nine distinct non-fatal health endpoints, not including a suite of hospital admissions. These conditions include chronic bronchitis, acute upper and lower respiratory symptoms, acute bronchitis, shortness of breath, and minor restricted activity days. The final report indicates that of these nine endpoints, all but one (work-loss days) utilized a value either directly from WTP studies or derived from WTP studies.

True, many of these symptoms are relatively minor – chronic bronchitis is the most severe – and the value of reduced mortality risks dominated the empirical results. But this is not always the case for EPA benefit analysis and, anyway, it misses the point. The fact is that as literature available to support practical economic analysis has grown EPA analysts have made earnest attempts to employ this literature. At least, we can conclude that from this limited sample.

I don't wish to overstate this conclusion. Some of the studies from which values were obtained are considered by some to be dated. There are also still problems with matching health effect values to endpoints affected by environmental polices, generally. This is not simply a matter of lacking estimates to work with – policy analysts always want more data. There are questions of how to apply existing work:

- to assess the suitability of existing estimates for use in benefits analysis
- to combine multiple valuation estimates for the same or similar effects
- to interpolate or extrapolate to health effects that differ from those in the original studies.

The economics literature is not ignoring these questions. Neither is EPA. A few other activities might provide additional context and illustrate the importance of this topic for EPA analysts.

Last year our office conducted an in-house survey to assess opinions regarding economic research priorities for the Agency. The survey collected information from a large number of EPA analysts involved in economic work. Aggregating across Agency offices, the greatest research needs appeared to be (1) improving methods to value changes in ecosystem form and function, and (2) valuing changes in morbidity risks.

Prior to this survey, EPA's Social Science Discussion Group (an internal workgroup) initiated an effort to produce a practical reference document for Agency analysts. This document, the *Handbook for Non-Cancer Health Effects Valuation*, now exists in a final draft form that has been subject to some external peer review. It should soon be available to analysts in a web-based format. Among many other things, the Handbook includes a summary of existing valuation estimates and devotes quite a bit of effort to issues associated with applying these estimates to environmental policies. The focus is on practical considerations in this benefit transfer exercise.

The increase in available WTP estimates for non-fatal effects – and in the need to incorporate these values into benefit analysis – has created a demand for work to support more careful consideration of how to apply them.

Historically, one constraint to fully utilizing valuation estimates for non-fatal health effects has been a lack of risk estimates commensurate with those often available for cancer risks. The use of "reference doses" approaches, for example, does not provide estimates of probability changes that economists can use directly. However, there have been some efforts at EPA to develop quantitative risk estimates for non-cancer health effects. Although the methods being explored cannot be implemented on a large scale right now, they may offer more to work with in the future.

I should note here that even in the absence of quantitative risk estimates, WTP values for morbidity effects may be valuable for performing analyses that supplement benefits assessments. Arsenic in drinking water, for example, is associated with a number of non-cancerous health effects (including kidney and liver damage, and vascular changes leading to hypertension), but the data do not support quantitative estimates of the number of potential cases. Estimates of value for these effects would at least allow for some breakeven or switch point analysis to inform decision makers.

Finally, analysts from across the agency are working with the newly-created Office of Children's Health Protection to explore how economics would estimate values for changes in the health status of children. A major concern for this group – and what may be a major concern for the agency – is related to chronic effects in children, including developmental and cognitive impairments. Of course, tomorrow's session will discuss some of these issues in more detail. I note this now simply to indicate the importance of valuation for other morbidity risks.

The papers presented in this session begin to address these, and other, concerns. The first paper, presented by Dr. William Schulze, serves as our segue from the previous session to this one, as well as a link to tomorrow's discussion. The paper also provides a break in the discussion of stated preference methods, obtaining valuation estimates from actual behaviors. This work, it should be noted was supported by EPA's Office of Air and Office of Policy.

The second paper, presented by Dr. Reed Johnson, assesses alternative – and combined – stated preference approaches to estimate WTP for components of health impairments. As an EPA economist I can appreciate that the work appears to be developed with benefit transfer applications to environmental policies clearly in mind.

Valuing Reduced Risk for Households with Children or the Retired

--Working paper*--

PRESENTED BY:
William Schulze
Cornell University

CO-AUTHORS:
Laurie Chestnut (Stratus Consulting)
Timothy Mount (Cornell University)
Weifeng Weng (Cornell University)
Hong Kim (Cornell University)

* This is a working paper developed for the US Environmental Protection Agency Office of Children's Health Protection, Office of Economy and Environment, and Office of Research and Development's workshop, "Valuing Health for Environmental Policy with Special Emphasis on Children's Health Issues," held on March 24-25, 1999, at the Silver Spring Holiday Inn in Silver Spring, Maryland.

Valuing Reduced Risk for Households with Children or the Retired[1]

Timothy Mount
William Schulze
Weifeng Weng
Hong Kim
(Cornell University, Ithaca, NY)

Laurie Chestnut
(Stratus Consulting, Boulder CO),

Paper prepared for the workshop on "Valuing Health for Environmental Policy with Special Emphasis on Children's Health Issues," March 24-26, 1999 in Silver Spring, Maryland.

Presenter: William Schulze, Department of Agricultural, Resource, and Managerial Economics, Cornell University, Ithaca, NY, 14853;
e-mail: wds3@cornell.edu; phone (607) 255-9611.

Section 1: Introduction

Little work has been done either theoretically or empirically to value morbidity and mortality either for children or retired adults. This paper addresses both of these issues by first presenting a theoretical model of how families value risk and then examining family automobile purchases. In particular, we show that parents may value risks to their children's lives and health (the model assumes two altruistic parents) through Nash cooperative bargaining to determine how much money to invest in the health and safety of their children. To allow empirical estimation of values, automobile safety is then shown to be a family public good, where the marginal investment cost of purchasing a safer automobile is set equal to the usage- weighted sum of the values of statistical life (VSL) of family members. We use data on automobile purchases to estimate how much families with children spend on automobile safety and how much families with retired members and no children spend on safety, for comparison to families without children or retired members. This allows estimation of an average value of a statistical life (VSL) for each type of family.

[1] We would like to thank Sharon Sandlan for her assistance in preparing this manuscript. We also would like to thank Tom Walton and Alan Carlin of the USEPA for their contributions to the research. All conclusions and remaining errors are the sole responsibility of the authors. This research was supported by USEPA Cooperative Agreement Number CR824393-01-1.

Our research using secondary data is a preliminary effort to determine the feasibility of collecting a national data set to allow estimation of separate values for mortality and possibly morbidity for different family members from choices made concerning both the type of vehicle and usage pattern by family members. A major limitation of the secondary data we use here is that only the usage of weighted average statistical values of life per family member can be estimated for single car families. We examine families with different compositions to attempt to see if differences exist.

The paper is organized as follows: Section 2 presents a simplified theoretical model of family automobile purchase decisions focusing on safety and how safety values for each individual are determined in a family setting. Section 3 describes our empirical work estimating a hedonic price function for automobiles showing a positive association between risk of fatal accident and price as well as our estimates of average implied values of life for different family groups. Section 4 discusses anomalies in the hedonic price function that have been estimated here and by others. In addition, we present descriptive information concerning automobile safety that will become important for improving value estimates in future research. Finally, we summarize our findings and implications for future research in Section 5.

Section 2. Theoretical Issues

How willingness to pay (WTP) for health and safety may vary with the age of the person at risk is a very important policy question for which we have little well-established empirical data. Cropper and Freeman (1991) address this question with a life-cycle consumption-saving model that they apply with a quantitative example to examine how WTP for a risk reduction in the current time period can be theoretically expected to change over a person's lifetime. This model is based on the premise that a person makes consumption and saving decisions over time to maximize personal utility. Because this model is based on the premise that utility is a function of consumption, the authors note that if there is additional utility derived from survival per se, then the life-cycle model provides a lower bound estimate of WTP. The quantitative example depends on assumptions regarding a lifetime pattern of earnings, endowed wealth, the rate of individual time preference, and other parameters of the model. These will all vary for different individuals, and uncertainty exists empirically about population averages for many of these factors. However, using reasonable values to calibrate the model is illustrative. Cropper and Freeman note that if consumption is constrained by income early in life, the model predicts that VSL increases with age until age 40 to 45, and declines thereafter. Shepard and Zeckhauser (1982) also illustrate this point with numerical examples for the life-cycle model. When they estimate the model with reasonably realistic parameters and assume no ability to borrow against future earnings or to purchase insurance, they find a distinct hump in the VSL function with a peak at around 40 years and dropping to about 50% of the peak by 60 years. When they allow more ability to borrow against future earnings and to purchase insurance, the function flattens and at 60 years drops only to 72% of the VSL at age 40. However, the hump shape to the VSL over a person's lifetime remains.

The conclusions reached by these theoretical analyses of the effect of age on WTP for mortality risk reduction using the life-cycle model are somewhat consistent with the empirical findings obtained by Jones-Lee et al. (1985). However, the empirical findings show that WTP varies with age much less than would be predicted by the life-cycle models. In this stated preference study, respondents gave WTP estimates for reductions in highway accident mortality risk and the answers showed a fairly flat hump-shaped relationship between VSL and age, peaking at about age 40. Although the directions of the changes in WTP with age are consistent with what the life-cycle models predict, the magnitudes of the changes are smaller. The Jones-Lee et al. results show that at age 65 the VSL is about 90% of the VSL of a 40-year-old person.

It is often suggested that WTP will be lower for the elderly than for the average adult because expected remaining years of life are fewer. This expectation is based on the presumption that WTP for one's own safety declines in proportion to the remaining life expectancy. Some analysts have suggested that effects of age on WTP might be introduced by dividing average WTP per statistical life by average expected years of life remaining (either discounted or not) to obtain WTP per year of life (Miller, 1989; Harrison and Nichols, 1990). Such a calculation implies very strong assumptions about the relationship between life expectancy and the utility a person derives from life; namely, that utility is a linear function of life expectancy. Although this might be correct, it is also plausible that this calculation will result in significant understatement of

WTP for the elderly. An understatement could result for a number of reasons. One is that there may be a value to being alive that is independent of the amount of time one expects to live. Another is that as one ages, the remaining time may be more highly valued than it was in midlife. In fact, the retired are now often characterized as "healthy and wealthy."

Determining appropriate WTP values for changes in mortality risks to children poses some particular analytical challenges. Children are not the economic decision makers whose preferences can be analyzed to determine an efficient allocation of society's resources regarding their own health and safety, so both revealed and stated preference approaches must rely on parental decisions to show what WTP for children's health and safety might be. Based on the expected relationship between WTP and expected life-years lost, it may be reasonable to assume that reductions in risks to children are valued equal to or greater than risks to adults. On the other hand, the life-cycle consumption-saving models show increasing WTP for risk reductions between the ages of 20 and 40, reflecting the typical pattern of increasing income and productivity during this stage of life. Extending this to children might suggest lower WTP for reducing risks to children, however, this pushes beyond the theoretical constructs of the life-cycle model regarding an individual as an economic decision maker. The only theoretical model which addresses these concerns, with respect to dependent children, has been developed by Chestnut and Schulze (1998). Their work treats the case of a family with non-paternalistic altruistic parents who engage in Nash cooperative bargaining to determine health and safety expenditures on their children and the implied VSL. We use this model as a starting point for our analysis.[2]

[2] It should be pointed out that some interesting revealed preference empirical approaches based on a household production function framework to analyze household expenditure decisions as they relate to children's health have been attempted (Agee and Crocker, 1996; Joyce et al. 1989). These analyses infer implicit WTP for changes in children's health as revealed by expenditure decisions of the household. Limitations in available data and analytical difficulties in properly specifying and verifying modeled relationships pose challenges for this approach; however, its basis in actual household decisions and behavior is an important strength. Estimates of WTP for changes in mortality risk for children are not directly available from these two studies, but similar approaches might be applied to obtain such WTP estimates.

Given the state of existing research, our first task is to develop a model that can potentially explain the behavior of households with dependent children. This model is developed in the context of automobile safety to allow empirical estimation of an appropriate family VSL, since the existing theoretical literature only considers individuals rather than families, with the exception of the work by Chestnut and Schulze mentioned above. Our work here paraphrases this earlier work and adds a hedonic market for automobile safety.

We begin by considering the case of a single individual with no family who may, or may not, survive for a single period. The following notation will be useful:

c = consumption,

w = wage income,

r = risk of a fatal automobile accident,

Π = probability of survival without automobile fatality risk,

Π-r = probability of survival with automobile fatality risk,

H(Π) = health expenditures (increasing in Π),

P(r) = automobile price (decreasing in r), and

U(c) = strictly concave utility function.

Note: subscripts denote derivatives where appropriate.

The individual must make two choices. First, the baseline probability of survival, Π, is chosen subject to the constraint that increasing Π increases health expenditures, H(Π), consequently reducing both consumption, c, and money available for purchasing a car, P. Similarly, the individual chooses how risky a car to drive, r, taking into account that lower r implies that the price of the car, P(r), is greater. Investments in health, Π, and automobile safety, reducing r, are chosen prior to realizing whether or not the individual will survive. The individual is assumed to maximize expected utility,

(1) $$(\Pi - r)U(c),$$

where the death state provides no utility because the individual has no family, subject to the budget constraint,

(2) $$(\Pi - r)(w-c) - P(r) - H(\Pi) = 0.$$

This budget constraint assumes that costless insurance (available for expected value) is available both to cover the purchase price of the automobile, P, and initial health and other safety investments, H. Most car loans, in fact, carry life insurance for the amount of the loan, and life insurance could presumably cover the costs of other health and safety investments. The optimal choice of Π is then determined by

(3) $$H_{\Pi} = VSL,$$

and, the optimal choice of r is determined by

(4) $$-P_r = VSL,$$

where,

(5) $$VSL \equiv (U/U_c) + w - c.$$

Equation (3) sets the marginal health cost of increasing the odds of survival equal to the value of the individuals statistical life (VSL) while equation (4) sets the marginal increase in price for purchasing a safer car equal to the VSL as well. The VSL is defined in (5) for the case of perfect insurance markets and is equal to the monetized value of utility, (U/U_c), which is lost in death, plus the excess of earnings over consumption. The interpretation of this relationship is much clearer in the family setting that we treat below, so we will defer discussion.

The model developed above can readily be extended to a family setting by using the Nash cooperative bargaining between parents approach employed by McElroy and Horney (1981).

Following our previous work (Chestnut and Schulze, 1998), we modify the notation used above, again considering a single car family (the case we analyze empirically), as follows:

n = the size of the family,

i = 1, 2,….,n denotes individual family members,

i = m = 1 denotes the mother,

i = f = 2 denotes the father,

i= k = 3, …..,n denotes children,

c_i = consumption of the ith family member,

w_i = wage of family member i,

r= automobile fatality risk, the same for all family members,

Π_i = probability of survival, excluding automobile fatality risk, of i,

$H(\Pi_1,\ldots\ldots, \Pi_n)$ = family health expenditures (increasing in Π_i),

P(nr) = automobile price (decreasing in total family risk, nr),

$U^k (c_k)$ = child's utility function,

$U^i(c_i ;\ldots,(\Pi_k\text{-}r)U^k(c_k),\ldots)$ = parent's utility function (i = m, f), and

E^i = bargaining threat point of expected utility in divorce (i = m, f).

The family must decide how much to allocate to each family member for consumption, spending on the health of each (and in so doing select survival probabilities), and on the safety level of the single automobile they purchase for all. Note that the demand for driving is inelastic in this model, since the only driving choice is over the risk of the chosen automobile. The hedonic price function for the automobile is now taken as P(nr) so that the total family risk level determines the price of the car. All of the existing hedonic price analyses of automobile safety use total fatalities per year for a vehicle model divided by the total number of that model on the road as the risk variable. Thus, the risk measure is not divided by occupancy (n in this theoretical model). It is, in fact, plausible to suppose that it is more expensive to increase the safety for each of four passengers than for one, so this assumption may be reasonable.

The utility functions of both the father and mother are assumed to depend not only on their own consumption, but also on the expected utilities of each of their children. The children's utility is assumed to be solely a function of their own consumption.

Investment in the safety and health of their children is a public good to the parents, which is the subject of negotiation, as is the level of consumption of each. The Nash cooperative bargaining model assumes that the solution maximizes the multiplication of the increase in the expected utility of the outcome over the threat point expected utility in divorce for the mother and the father. The threat points are assumed, in models of the family, to be a function of divorce laws, job opportunities, etc. Thus, in the Nash cooperative bargaining solution,

(6) $$[(\Pi_m\text{-}r)U^m(c_m;..,(\Pi_k\text{-}r)U^k(c_k),..)\text{-}E^m]\ [(\Pi_f\text{-}r)U^f(c_f;..,(\Pi_k\text{-}r)U^k(c_k),..)\text{-}E^f],$$

is maximized with respect to c_i, Π_i, and r, subject to the budget constraint,

(7) $$\sum_{i=1}^{n} \Pi_i(w_i - c_i) - P(nr) - H(\Pi_1,\ldots\ldots, \Pi_n) = 0.$$

The resulting conditions for allocating health expenditures and survival probabilities take the form:

(8) $$H_i = U^i/U^i_c + w_i - c_i \equiv VSL_i \qquad i = 1,\ldots\ldots\ldots,n.$$

The remarkable fact is, that, in spite of the complicated structure of the problem specified above, the implied VSL_i for each family member shown in (8) is identical in form to that for the single individual shown in (5) above. The interpretation of the VSL_i can be illustrated with the following examples. Imagine that the mother is the sole breadwinner with a stay-at-home father. In this case, assuming that the children are young, $w_i - c_i < 0$ for the other family members and $w_m - c_m > 0$ for the mother. Thus, if the mother were to die, this would be a severe financial blow to the rest of the family and the mother's VSL would reflect this relative to the VSL of other family members. For young children it is clear that $w_k - c_k < 0$ in the short run. However, in the inter-temporal version of the model, $w_k - c_k$ is replaced by its discounted present value, which may be positive. U^i/U^i_c depends solely on c_i in the single period model and on the lifetime consumption pattern in the full inter-temporal model. The important point is that the child's consumption depends in youth on the parents' income and wealth. Further, if parents find the value of their child's smile to be high enough, the child's consumption will be maintained, by them, at a high level, leading to a high VSL. A young child's utility may also be large from relatively small levels of financial consumption, also leading to a high VSL. These arguments suggest that the VSL of children is a purely empirical question and depends not only on their own life cycle wealth but also on their family's wealth.

Finally, the choice of automobile risk, r, is determined by

$$(9) \qquad -nP_r = \sum_{i=1}^{n} VSL_i.$$

Thus, the safety of the shared family vehicle is determined by a public good condition which sets the marginal cost of obtaining a safer vehicle for each individual equal to the sum of the VSLs of individual family members. The slope of an estimated hedonic price function for automobile safety is $-P_r$, which, by (7), is equal to the average VSL for the family, $\sum_{i=1}^{n} VSL_i /n$.

Thus, if we examine P_r for different households with a single car, we can obtain estimates of the average value of life for those households. However, the average is a weighted average where the weights are determined by each family member's use of the vehicle.

Section 3. The Hedonic Model of Implied Average Values of Statistical Life for Different Families.

The econometric model used here is based on the work of Rosen (1974), Atkinson and Halvorsen (1990), and Dreyfus and Viscusi (1995) on hedonic pricing.

Atkinson and Halvorsen (1990) use the data for 112 models of new 1978 automobiles to obtain estimates of the VSL. Since the available fatality data is a function of both the inherent risk of the vehicle and the driver's characteristics, the drivers' characteristics are included in the regression as control variables. The VSL is calculated based on WTP. Their estimated VSL for the sample as a whole is $3.357 million 1986 dollars.

The data used in Dreyfus and Viscusi (1995) differ from those used in earlier studies in that they reflect actual consumer automobile holdings. Dreyfus and Viscusi (1995) use the 1988 Residential Transportation Energy Consumption Survey together with data from industry sources. They generalize the standard hedonic models to recognize the role of discounting on fuel efficiency and safety. The estimates of the implicit value of life range from $2.6 to $3.7 million and the estimates of discount rate range from 11 to 17 percent.

We utilized both the industry data source of the vehicle attributes and the households' choice of automobiles to estimated the willingness-to-pay for changes in the risks of mortality and to derive the per capita value of statistical life for different types of households. The hedonic price equation can be written, following Atkinson and Halvorsen (1990):

$$(10) \qquad P_{auto} = f(R, A),$$

where P_{auto} is the price of automobile, R is the inherent mortality risk associated with the automobile, and A is a vector of other characteristics. The available mortality data, F, is a function of both R and a vector of the involved driver's characteristics D. Assuming that F is monotonic in R, equation (1) can also be written as:

$$(11) \qquad P_{auto} = g(F, A, D),$$

The function form used for the estimation is

$$(12) \qquad \log(P_{auto}) = \alpha_0 + \sum_k \alpha_k \log(X_k) + e,$$

where X_k is a representative regressor and e is an unobserved residual.

The 1995 National Personal Transportation Survey (NPTS) is used to obtain information on the household's choice of automobiles. 1995 NPTS was conducted by the Research Triangle Institute (RTI) under the sponsorship of the U.S. Department of Transportation (DOT). The survey covers 42,033 sampled households. A sub-data set of 4036 one-car households holding a 1990-1995 model year vehicle were merged with vehicle attribute data collected from industry and other sources for the same years. The vehicle price data is gathered from *NADA Official Used Car Guide*, and other attribute data is collected from *NADA Official Used Car Guide*, *Ward's Automotive Yearbook*, and *Consumer Reports*, respectively. The vehicle mortality rate is measured by the number of fatalities occurring in each make/model/year vehicle per 1000 of that vehicle sold. The number of fatalities is based on the information from the U.S. Department of Transportation's Fatality Analysis Reporting System (FARS) for calendar year 1995-1997. For model year 1990-1994 vehicles, the average of the FARS 1995 and 1996 fatalities is used, while for model year 1995 vehicles, the average of the FARS 1996 and 1997 fatalities is used. Since the mortality rate is jointly determined by the inherent mortality risk associated with the automobile and the driver's characteristics, a vector of driver's characteristics is also included in the model to provide control variables. The variables of driver's characteristics are gathered from FARS 1995-1997. The variables used are summarized in Table 1, while Table 2 shows the descriptive statistics of selected vehicle attributes. The selection of vehicle attributes and driver's characteristics is similar to Dreyfus and Viscusi (1995) and Atkinson and Halvorsen (1990).

Least square estimates of the log linear model are presented in Table 3. Two equations are estimated separately. The first equation omits fuel economy, while

Table 1. Variable Definition

Variable Name	Definition
Price	Vehicle price as of end-of-year 1995.
Value Retained	Original sales value retained, as of end-of-year 1995.
Mortality Rate	Number of fatalities occurring in that make/model/year vehicle per 1000 of that vehicle sold.
City Fuel-efficiency	Miles per gallon in city area.
Reliable Rating	A discrete variable coded from 1 to 5, 5 is the highest while 1 is the lowest.
Acceleration	The horsepower-to-weight ratio.
Traditional Styling	Length plus width divided by height.
ClassX	Discrete variables coded as 1 for the appropriate class. Class1 to class7 represent small, middle, large, luxury, SUV, van, and pick-up truck, respectively.
YearXX	Discrete variables coded as 1 for the vehicle model year.
Young Driver	Proportion of fatalities in this make/model/year vehicle in which the driver was younger than 25 years.
Older Driver	Proportion of fatalities in this make/model/year vehicle in which the driver was 45 or older.
Alcohol	Proportion of fatalities in this make/model/year vehicle in which the alcohol involvement was reported.
Gender of Driver	Proportion of fatalities in this make/model/year vehicle in which the driver was male.
Seat Belt	Proportion of fatalities in this make/model/year vehicle in which the driver was wearing a seat belt.
Previous Offenses	Proportion of fatalities in this make/model/year vehicle in which the driver had no previous offense.
Late Night	Proportion of fatalities in this make/model/year vehicle which occurred between 12:00am to 5:59am.
One-car Accident	Proportion of fatalities in this make/model/year vehicle in which only one vehicle was involved.

Table 2. Descriptive Summary of Selective Variables

Variable	Mean	Standard Deviation
Price	15703.53	9371.57
Value Retained	0.7720	0.1753
Mortality Rate	0.1345	0.0994
City Fuel-efficiency	20.26	4.82
Reliable Rating	3.019	1.321
Acceleration	0.0475	0.0102
Traditional Styling	4.451	0.519

Table 3. Parameter Estimates

Variable	Model 1 Estimated Coeffiecent	Model 1 Standard Error	Model 2 Estimated Coeffiecent	Model 2 Standard Error
Constant	8.7815	0.2268	12.0281	0.2557
Value Retained	0.4974	0.0530	0.5224	0.0456
Mortality Rate	-0.0525	0.0084	-0.0348	0.0072
City Fuel-efficiency	…	…	-0.9421	0.0480
Reliable Rating	0.0855	0.0128	0.1388	0.0113
Acceleration	0.3227	0.0421	0.2158	0.0366
Traditional Styling	0.6176	0.0943	0.3036	0.0827
Class2	0.2624	0.0200	0.0754	0.0197
Class3	0.4112	0.0349	0.1312	0.0332
Class4	0.8388	0.0272	0.5354	0.0281
Class5	0.7817	0.0318	0.1935	0.0406
Class6	0.6518	0.0314	0.1878	0.0359
Class7	0.2802	0.0312	-0.1093	0.0334
Year91	0.1076	0.0232	0.1032	0.0199
Year92	0.2095	0.0258	0.2063	0.0222
Year93	0.3099	0.0293	0.3122	0.0252
Year94	0.3859	0.0325	0.3909	0.0279
Year95	0.4618	0.0353	0.4675	0.0303
Young Driver	0.0237	0.0373	0.0471	0.0321
Older Driver	-0.0223	0.0306	-0.0369	0.0263
Alcohol	0.0585	0.0364	0.0421	0.0313
Gender of Driver	0.0520	0.0284	0.0076	0.0245
Seat Belt	-0.0054	0.0276	0.0015	0.0238
Previous Offenses	-0.0048	0.0293	-0.0179	0.0252
Late Night	0.0362	0.0406	0.0216	0.0349
One-car Accident	0.0217	0.0287	0.0065	0.0246
R^2	0.8311		0.8752	

the second equation includes that variable. All coefficients of vehicle characteristics are significant at the 99 percent significance level. None of the coefficients of driver's characteristics is significant at the 95 percent significance level. Since driver's characteristics should be included as control variables based on the model structure, they are kept in the model.

Model 2 shows the negative sign for the coefficient of fuel economy. Since keeping all other vehicle characteristics constant, it should cost more to produce a more fuel-efficient engine, and since consumers would be less willing to pay for a car with poor fuel economy, the sign of the coefficient is wrong. Comparison between Model 1 and Model 2 also shows that the coefficient on mortality rate is not robust. In fact, adding more regressors into the model drives the coefficient of mortality rate down. Since our main purpose is to compare the differences of per capita value of statistical life (VSL) among different types of households, both models presented in Table 3 are used to illustrate the effects of the different specifications.

The estimated VSL is:

(13) $$VSL = \frac{\beta_m \text{price}}{\text{mortality rate}} / \left(\sum_{t=1}^{L} \left(\frac{1}{1+r}\right)^t\right),$$

based on the NPTS data for the household's choice of automobiles, where r is the discount rate, set to 10 percent, L is the remaining vehicle life. The expected life of vehicle is set as 6 years. The means of estimated per capita VSL for different types of households, based on the estimates of Model 1 and 2, respectively, are presented in Table 4. From Table 4, we can see that the means per capita VSL are very similar for different types of households, which leads to the conclusion that the value of a statistical life for children and seniors may not differ appreciably from that of other age groups.

Table 4. The Mean of Estimated Per Capita Value of Statistical Life (VSL) for Different Types of Households

Household Category	VSL based on Model 1 (million)	VSL based on Model 2 (million)
Grand Mean	2.730	1.809
Household No One Retired/No Kid	2.729	1.808
Household With Kids	2.643	1.751
Household With Retired Member/No Kid	2.811	1.862

Since the estimated coefficient for safety in the hedonic price equation is not robust to changes in the specification of the model, it is important to consider how the model can be improved. One of the main limitations of the standard single-equation specification is that the effects of the vehicle's physical characteristics are confounded with the characteristics and behavior of the family or individual that uses the vehicle. Although it is quite reasonable to consider that a safer vehicle would cost more to buy, using actual fatality data to measure safety is dependent on driving behavior as well as the physical characteristics. This is illustrated in Table 5 which reports relative death rates per vehicle year for different types of vehicle (values < 100 indicate fewer fatalities than average). In most categories of vehicle, larger vehicles have lower fatality rates. For luxury vehicles, the rates for the medium and large sizes are essentially identical, but for sports cars, the rate for medium sports cars is substantially larger than it is for small sports cars. This is likely to reflect the behavior of the drivers of sports cars as much as it does the physical safety of the vehicles.

Table 5: Relative Annual Death Rates of Drivers and Passengers by Type of Vehicle, 1991-5

	Sports	2 door	4 door	Luxury	Wagon, Minivan	Sports Utility	Pickup
Small	146	154	135	-	112	174	-
Medium	191	120	88	62	63	81	153
Large	-	81	74	65	52	60	106

100 is average
< 100 implies fewer fatalities than average
> 100 implies more fatalities than average
Source: Insurance Institute for Highway Safety, September 1997

To clarify this situation, one should distinguish between 1) the probability of having an accident that involves at least one fatality and 2) the probability of surviving in such an accident. The probability of having an accident is definitely influenced by driving behavior and is probably influenced by the vehicle's characteristics as well. The probability of survival, on the other hand, is definitely influenced by the physical characteristics of the vehicles involved in the accident, and possibly by the characteristics of the driver and the occupants (e.g. very young children and very old people may be less able to recover from injuries). These issues are illustrated in Table 6 by the average survival rates for different ages of the occupants of vehicles in which at least one fatality occurred.

Table 6: The Average Survival Rates for Occupants of Vehicles by Year and Age Group (Percent)

Age	All Accidents			≥ 1 Occupant is		
	1995	1996	1997	1995	1996	1997
≤5	67	66	67	67	66	67
6-15	67	68	69	77	77	77
16-21	48	49	48	71	69	69
22-24	42	42	43	66	67	64
25-64	31	31	31	58	59	58
≥ 65	24	23	22	26	31	35

Every observation has at least one fatality in the vehicle.

Source: Fatality Analysis Reporting System, US Department of Transportation

In general, the survival rates by age group in Table 6 are highly consistent across years. For All Accidents, the survival rates for young children (≤5 years old) are much higher than they are for other age groups. For the oldest age group (≥ 65 years old), the survival rate (22%-24%) is only one third of the rate for young children (66%-67%). However, most accidents do not involve young children. To provide an alternative comparison across age groups, the corresponding survival rates for accidents with at least one very young child as an occupant are also shown in Table 6. Hence, these alternative rates are more representative of the survival rates of different age groups relative to the survival rates for children ≤5 years old. Using this measure, children from 6-15 have the highest survival rates, and the rates decline with age after that, but they are higher than the corresponding values for All Accidents. The rates for people ≥ 65 years old (26%-35%) are substantially lower than all other age groups. These results provide some evidence that people ≥ 65 years old are more vulnerable to fatalities in an accident. It should be pointed out, however, that these differences in survival rates may reflect the choice of seat (e.g. children are more likely to travel in the rear seat than adults) as well as the medical susceptibility of the individuals to injuries.

The information in Table 7 provides some evidence about the relative safety of different seats in a vehicle. Using data on accidents with a fatality, vehicles were selected in which there was a driver, at least one front seat passenger and at least one rear seat passenger. Hence, the values in Table 7 represent the survival rates when all three types of occupant were involved in an accident simultaneously. Once again, the survival rates are consistent across years. Passengers in the rear seats have substantially higher survival rates than people in the front seats, and passengers in the front seat have slightly lower survival rates than the drivers.

Table 7: Survival Rates in Accidents with a Fatality*
by the Location of the Seat of the Occupant

Type of Occupant	1995	1996	1997
Driver	62	62	61
Front Seat Passenger	58	61	61
Rear Seat Passenger	70	70	69

*At least one of every type of occupant is in each vehicle.

Source: Fatality Analysis Reporting System, U.S. Department of Transportation

The evidence about the effects of the characteristics of the driver on safety is more complicated to interpret. The average survival rates for different age groups (the same as Table 6) from 1995-1997 are broken down by the age and gender of the driver in Table 8. The expectation was that the survival rates for occupants would be lower when young male drivers were involved, for example. However, there are no consistent differences between male and female drivers in Table 8. Nevertheless, the survival rates are substantially lower for children and young adults when the drivers are also young (25 years old) and highest when the drivers are old (≥ 65 years old). For older occupants, the lowest survival rates correspond to the same age group as the driver, presumably because many of the accidents involve only a single occupant/driver.

Table 8: Survival Rates of Occupants by the Age and the Gender of Driver, 1995-1997 (Percent)

		Age and Gender of Driver					
		Male			Female		
Age	All drivers	24	25-64	≥65	24	25-64	≥65
5	67	62	69	80	61	68	83
6-15	68	61	72	84	63	72	86
16-21	49	47	70	78	44	72	88*
21-24	42	37	61	87*	39	67	43*
25-64	31	48	29	59	47	31	66
≥65	23	20	28	23	24	25	22

*Less than ten fatalities reported
Source: Fatality Analysis Reporting System, U.S. Department of Transportation

The previous discussion shows that actual fatality rates vary by the age of the occupant, the location of the seat of an occupant and the age of the driver. The first attempt to develop an objective measure of vehicle safety, that was independent of driving behavior, was to use standardized crash data from the Consumers Union. These data simulate the severity of the injuries to occupants from hitting a wall at 35 mph. While data show the expected higher safety for larger cars, they also give a relatively low rating for Sports Utility Vehicles (SUV) even though these vehicles are generally larger than most cars. The standard crash tests correspond to head-on collisions between two identical vehicles, and consequently, the accidents simulated in the crashes are more severe when the vehicles are heavier. One reason for driving a large SUV is to gain a safety advantage over other vehicles in an accident. A good measure of safety for different vehicles should reflect the characteristics of a "typical" accident, and in this respect, an SUV should have a relatively high safety rating.

A more promising approach to developing an objective measure of safety is to try to isolate the effects of the physical characteristics of the vehicles from the driving behavior using regression techniques with the actual fatality data. The first step was to select accidents that involved two vehicles (from the seven categories of private vehicles discussed in the previous section). The objective of the model was to predict the survival rate of the occupants in one vehicle, using the physical characteristics of that vehicle and of the "other" vehicle as explanatory variables. In addition, the characteristics of the driver and of the occupants are also likely to be important as explanatory variables. The general form of the regression equation can be written as follows for an accident involving vehicles i and j:

$$S_i = f(W_i / W_j, G_i, G_j, O_i)S_{ij} + e$$

where S_i is the survival rate for vehicle i

S_{ij} is the survival rate for both vehicles combined

W is the weight

G is a vector of other physical characteristics

O is a vector of occupant characteristics

e is an unobserved residual

The multiplicative form of S_{ij} in the equation is designed to reflect the severity of each accident. If all occupants in both vehicles in an accident are killed, then all survival rates are zero and $S_{ij} = S_i = S_j = 0$.

Preliminary results using this type of model are encouraging. The signs of coefficients are logical and the magnitudes are estimated relatively accurately. Weight is clearly an important factor, and the survival rates of occupants in small vehicles are much lower if they hit a large vehicle rather than another small vehicle. When the final form of the equation has been selected, it will be used to predict the survival rates for individual vehicles hitting an average vehicle in a typical accident. Under this scheme, small vehicles will be at a relative disadvantage in the simulated accident, but large vehicles, like an SUV, will be heavier than the average vehicle, and therefore, the corresponding predicted survival rate will be relatively high. This will provide a clear link between safety and the weight of a vehicle. Hence, the ambivalent role of the weight of a vehicle in a standard model of the hedonic price of a vehicle can be clarified. Although heavier vehicles are safer, they are likely to be more expensive to buy and more expensive to operate, because they are less fuel-efficient. The implications of these different relationships will be explored in more detail as part of the next component of our research.

Section 4. Conclusions

Our analysis in the preceding sections, while encouraging for our proposed national survey of automobile usage, points out some important potential difficulties.

First, from the theoretical model of Section 2, it is apparent that we must collect data on usage by individuals, by automobile type, to estimate fraction of usage by age (child, adult, or senior) for multiple car families.

Second, since risk differs depending on seating position, these data must be collected as well.

Third, the role of weight in vehicle safety must be explored further. Since such attributes as safety, interior room, cargo space, and fuel economy are all correlated with vehicle weight, no reliable estimate of marginal cost of safety can be identified from the hedonic price function until a likely simultaneous equation bias problem is solved.

Finally, although considerable theoretical speculation exists that the value of a statistical life should differ by age, we find little support for this hypothesis in our preliminary analysis.

References:

Agee, M.D., and T.D. Crocker. 1996. "Parental Altruism and Child Lead Exposure: Inferences from the Demand for Chelation Therapy." The Journal of Human Resources 31:677-691.

Atkinson, Scott E and Robert Halvorsen (1990). "The Valuation of Risks to Life: Evidence from the Market for Automobiles." The Review of Economics and Statistics 72 (1): 133-136.

Becker, G. S. (1974) "A Theory of Social Interactions," Journal of Political Economy, 82, 1095-1117.

Becker, G. S. (1991) A Treatise on the Family, Harvard University Press, enlarged edn.

Bergstrom, T. C. (1996) "Economics in a Family Way," Journal of Economic Literature, 34(4), 1903-1934.

Chestnut, L. and W. Schulze. 1998. "Valuing the Long Term Health Risks From Wartime Toxic Exposures," Proceedings of the First International Conference on Addressing Environmental Consequences of War.

Cropper, M.L. and A.M. Freeman III. 1991. "Environmental Health Effects." Measuring the Demand for Environmental Quality. J.B. Braden and C.D. Kolstad (ed.) North-Holland. New York.

Dreyfus, Mark K. and W. Kip Viscusi (1995). "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency." Journal of Law and Economics, vol XXXVIII (April 1995): 79-105.

Harrison, D. and A.L. Nichols. 1990. Benefits of the 1989 Air Quality Management Plan for the South Coast Air Basin: A Reassessment. Prepared for California Council for Environmental and Economic Balance by National Economic Research Associates, Inc., Cambridge, Massachusetts.

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

Joyce, T.J., M. Grossman, and F. Goldman. 1989. "An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health." Journal of Urban Economics 25:32-51.

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

Miller, T.R. 1989. "Willingness to Pay Comes of Age: Will the System Survive?" Northwestern University Law Review 83:876-907.

NADA official used car guide (1995) Vol. 62. McLean, Va., National Automobile Dealers Used Car Guide Co.

Rosen, Sherwin (1974). “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.” Journal of Political Economy 82 (1): 34-55.

Shepard, D.S. and R.J. Zeckhauser. 1982. "Life-Cycle Consumption and Willingness to Pay for Increased Survival." in M.W. Jones-Lee (ed.). The Value of Life and Safety. New York: North-Holland.

Ward’s Automotive Yearbook (1990-1996). Vol. 52-58.

Willingness to Pay for Air-Quality Related Health Improvements: A Multiple-Format Stated-Preference Approach

--Working Paper*--

PRESENTED BY:
F. Reed Johnson
Triangle Economic Research

CO-AUTHORS:
Melissa Ruby (University of North Carolina at Chapel Hill)
William Desvousges (Triangle Economic Research)

* This is a working paper developed for the US Environmental Protection Agency Office of Children's Health Protection, Office of Economy and Environment, and Office of Research and Development's workshop, "Valuing Health for Environmental Policy with Special Emphasis on Children's Health Issues," held on March 24-25, 1999, at the Silver Spring Holiday Inn in Silver Spring, Maryland.

TRIANGLE
ECONOMIC
RESEARCH

Willingness to Pay for Air-Quality Related

Health Improvements: A Multiple-Format Stated-Preference Approach

F. Reed Johnson*

Melissa C. Ruby**

William H. Desvousges*

* Triangle Economic Research
** University of North Carolina, Chapel Hill

This research was carried out in part under Public Works and Government Services Canada contract #H4078-5-C174/01-SS. Funding was provided by Health Canada, Environment Canada, Ontario Hydro, the Ontario Ministry of Environment and Energy, and Environnement et de la Faune Québec. However, the views expressed in this paper do not necessarily represent those of any of the sponsoring organizations. The authors wish to thank the project Scientific Authorities, Dave Stieb and Paul De Civita, as well as the study advisory panel: Mike Fortin, Magnus Johannesson, Kerry Smith, and David Streiner. Alicia Gable, Sara Hudson, and Matthew Bingham of TER provided invaluable assistance at various points in the study. Any errors are, of course, solely the responsibility of the authors.

1000 Park Forty Plaza
Suite 200
Durham, NC 27713

•

Phone: 919-544-2244
Fax: 919-544-3935
E-mail: frj@ter.com

Abstract

This study uses stated preference (SP) analysis to measure willingness to pay to reduce acute episodes of respiratory and cardiovascular ill health. The SP survey employs a modified version of the health-state descriptions used in the Quality of Well Being index. The four health-state attributes are symptom, episode duration, activity restrictions, and cost. Preferences are elicited using two different SP formats: graded-pair and discrete-choice. The different formats cause subjects to focus on different evaluation strategies. Combining two elicitation formats yields more valid and robust estimates than using only one approach.

We obtain estimates of indirect utility function parameters using advanced panel econometrics for each format separately and jointly. Socioeconomic differences in health preferences are modeled by allowing the marginal utility of money relative to health attributes to vary across respondents. Because the joint model captures the combined preference information provided by both elicitation formats, we use these model estimates to calculate willingness to pay.

The results demonstrate the feasibility of estimating meaningful WTP values for policy-relevant respiratory and cardiac symptoms, even from subjects who never have personally experienced these conditions. Furthermore, because WTP estimates are for individual components of health improvements, estimates can be aggregated in various ways depending upon policy needs. Thus using generic health attributes facilitates transferring willingness-to-pay estimates for benefit-cost analysis of a variety of potential health interventions.

Willingness to Pay for Air-Quality Related Health Improvements: A Multiple-Format Stated-Preference Approach

The economic analysis of many health-intervention and regulatory programs often requires evaluating the benefits of improved health. Obtaining credible measures of the economic value of morbidity, however, is one of the more difficult problems facing health economists. The existence of insurance, universal health-care systems, and market participants that are unrepresentative of the population of policy interest often obscure essential supply-and-demand relationships. Thus, revealed-preference information generally has proven to be an insufficient basis for obtaining policy-relevant values of human health. In addition, studies which use contingent valuation (CV) are not well suited to valuing multiple health-state attributes. Health economists increasingly are turning to stated-preference (SP) approaches as an alternative.[3]

This study demonstrates the feasibility of applying SP techniques to elicit values for health conditions described in terms of symptom, activity restriction, and duration. Furthermore, by combining two SP elicitation methods, graded-pair and discrete-choice, the estimates presented in this study are more valid and robust measures of benefits than could be obtained from a single format. Our results demonstrate the sensitivity of results to the symptoms being considered as well as to the activity restrictions associated with those symptoms. Subjects clearly indicate systematic preferences for milder morbidity effects over more severe ones. These preferences translate into WTP estimates that vary logically with severity.

SP methods evolved as market-research tools for evaluating consumer behavior and predicting sales of new products (Cattin and Wittink, 1982; Wittink and Cattin, 1989). SP recently has been applied in environmental and health economics as an alternative to CV methods.[4] Viscusi, Magat, and Huber (1991) and Krupnick and Cropper (1992) (using the Viscusi data) use SP analysis to elicit a value from subjects for reducing chronic health risks. Other SP studies have elicited preferences for other health and health-care attributes, including Ryan, McIntosh, and Shackley (1998), Chakraborty, Gaeth, and Cunningham (1993), Ryan and Hughes (1997), Bryan et al. (1997), and Van der Pol and Cairns (1998), and Propper (1995).

In this study, specified health states consist of multiple attributes. We presume that people have preferences for different levels within these attributes and are willing to accept some trade-offs among them. The utility obtained under different health states is derived from the revealed trade-

---

[3] Terminology has not been standardized among various disciplines. In this study, we use the term "stated preference" to refer to a group of techniques used primarily in market-research studies to measure consumer preferences. The term "conjoint analysis" also has been used to describe some of these techniques. Although contingent valuation also could be called a SP technique, CV was developed independently by environmental economists and generally relies on a different set of elicitation formats and analytical approaches.

[4] In the resource economics literature Gan and Luzar (1993) use SP to value hunting trips in Louisiana. Mackenzie (1993) values hunting trips in Delaware using SP analysis. Opaluch et al. (1993) also use SP to describe public preferences for siting a noxious facility. Adamowicz, Louviere, and Williams (1994) and Adamowicz et al. (1997) use SP to explain recreational site choice selection. Johnson et al. (1995) use SP to estimate electric customers' willingness to pay for environmental and other attributes of electricity generation. Roe, Boyle, and Teisl (1996) use SP to value the effects on sport fishing of implementing alternative management plans to restore runs of Atlantic salmon in Maine.

offs.[5] By including a cost attribute, we can use the implicit marginal utility of money to scale changes in health-state utility in monetary units.

SP Elicitation Formats

Two types of SP analysis lend themselves to valuing health effects.[6] The first, *graded pairs*, measures subjects' valuations of variations in attributes by requiring them to evaluate trade-offs among various attributes. In a graded-pair format, subjects sequentially are presented with several different pairs of bundled commodities, represented as sets of attribute levels, and asked to compare each pair. They are asked to rate the intensity of their preference for one of the pairs on a numerical scale, say from 1 to 7, where 1 indicates a strong preference for the first bundle, 7 indicates a strong preference for the second bundle, and 4 indicates indifference between the two bundles. The subject is asked to rate a series of these pairs, with each pair having different attributes or attribute levels. (See Viscusi, Magat, and Huber, 1991, for an example of this approach in valuing bronchitis risks.)

Discrete choice, in contrast, confronts subjects with several different products or programs simultaneously and simply asks them to identify the most-preferred alternative in the choice set.[7] As with the graded-pair approach, each commodity is described as a set of attributes. However, the subjects do not provide a rating of the intensity of their preferences. Typically, each subject is shown a series of these choice sets to evaluate.[8]

Each of these SP approaches has its advantages and disadvantages. The graded-pair format provides intensity-of-preference information and thus is statistically more efficient than the discrete-choice method which simply elicits commodity preferences. Graded pairs also are somewhat easier to design and provide opportunities to check on subject attentiveness and coherence of expressed preferences (Johnson and Desvousges, 1997; Johnson, MacNair, and Fries, 1997). However, graded pairs are more cognitively difficult than choices, requiring subjects to identify which profile they prefer and the degree to which they prefer it. Furthermore, graded pairs may require more complex statistical analysis to account for variations in the interpretation of the scale across subjects (Johnson and Desvousges, 1997). Finally, when graded pairs do not include a constant alternative across repetitions, the data may not be clearly linked to an unambiguous welfare reference point needed for benefit-cost analysis.[9]

Choice-format questions avoid these problems at the cost of statistical efficiency and perhaps less depth in preference searching. Discrete choices can be less burdensome to subjects, depending upon the number of alternatives and the number of attributes in a given choice set. Also, discrete-choice formats conventionally include a status-quo, or "opt-out", alternative in every choice set (Olsen and Swait, 1997). This status-quo option provides an unambiguous reference point for

---

[5] Defining the properties of such preferences has been explored by multiattribute utility theory (Keeney and Raiffa, 1978). For an application to health-state utility, see Torrance, Boyle, and Harwood (1982).

[6] A third format, *ranking*, has also been used in market research. For a more extensive review of experimental design possibilities, see Green, Tull, and Albaum (1988).

[7] For a review of choice experiments in marketing applications, see Carson et al. (1994).

[8] This approach has been used in modeling travel behavior, evaluating new products, and estimating recreation demands. See, for example, Louviere and Hensher (1982). Health studies which use this approach include Bryan et al. (1997), Propper (1995), and Ryan and Hughes (1997).

[9] Roe, Boyle, and Teisl (1996) use a graded-pair format with a constant reference alternative. However, including such an alternative diminishes the statistical efficiency of the graded-pair format.

welfare changes. Finally, some researchers (Louviere, 1988; Olsen 1992) have argued that a discrete-choice task, in contrast to graded pairs, corresponds more closely to how real-world decisions are made and thus is a better predictor of choice. However, other researchers disagree (Huber et al., 1993) and suggest that the graded-pair approach, because it provides a more intense searching of preferences, can predict choice better than discrete-choice formats. Elrod, Louviere, and Davey (1992) found that both graded-pair and choice formats predicted choices equally well for a conventional market good.

Huber et al. (1993) argue that each format has relative merits because each relies on a different cognitive process for eliciting preferences.[10] The different formats cause subjects to focus on different evaluation strategies. The graded pairs encourage thinking about the value of marginal tradeoffs among attribute levels. In contrast, the choice task encourages subjects to eliminate alternatives that are unacceptable for a particular attribute. Real-world behavior exhibits both of these strategies. Indeed, Huber et al. have found that employing more than one elicitation format may predict choice better than using one format alone, simply because each format requires different heuristics and may provide only a partial picture of preferences. In this study, we have adopted this approach and apply both graded-pair and discrete-choice formats to estimate joint WTP estimates.

Table 1 shows the levels associated with each attribute in the two SP formats. The experimental design consists of main-effects, nearly orthogonal arrays of 40 graded pairs and 40 choice sets using the attribute levels in Table 1. Given subjects' practical time and attention constraints, we administered eight graded pairs and eight choice sets to each subject, each set of which was drawn randomly from five design blocks.

Graded-pair and discrete-choice sp questions

The survey was administered by computer, and Figure 1 shows an example of a graded-pair screen. In this example, the price is expressed as illness-related costs.[11] The subjects indicate their preferences for Condition A versus Condition B. The complete SP exercise presents a series of these graded pairs to subjects and records their ratings. Because there is no unambiguous baseline reference point in the graded-pair questions and no health conditions with zero cost, the graded-pair questions are designed to obtain information on marginal trade-offs among health attributes and costs.

[10] See also Huber (1997), Tversky, Sattath, and Slovic (1988), Payne (1976, 1982), and Huber and Klein (1991) for discussion of the different cognitive processes involved in graded-pair versus choice tasks.

[11] The payment vehicle for both formats is described as illness-related costs that are not covered by the government health system or a company insurance plan. These costs are associated with items that reduce discomfort or the length of illness (such as vitamins, medicines, air filters or humidifiers, special foods or liquids, or other optional treatments). These costs also may include such costs as child care while sick or transportation to the doctor. Subjects were instructed to assume that any missed time from work would be covered by paid sick leave.

Table 1. Attribute and Attribute Levels Shown in Graded-Pair and Discrete-Choice Comparisons

ATTRIBUTE	LEVEL	DESCRIPTION		
Symptom	Nose	Stuffy/runny nose and sore throat		
	Eye	Eye irritation		
	Flutter	Fluttering in chest and feeling light-headed		
	Breath	Coughing, wheezing, shortness of breath		
	Ache	Coughing or wheezing with fever, chills, or aching all over		
	Swell	Shortness of breath, and swelling in ankles and feet		
	Pain	Pain in chest or arm		
Duration	**1**		**5**	**10**
	1-day episode		5-day episode	10-day episode
Daily Activity	NoLim	You can go to work, go to school, do housework, participate in social or recreational activities, and have no physical limitations.		
	SomeLim	You can go to work, go to school, do housework, and participate in social or recreational activities, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition.		
	NoSoc	You can go to work, go to school, do housework, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities), and cannot participate in social or recreational activities because of this health condition.		
	AtHome	You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition, but you can care for yourself.		
	NeedHelp	You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you need help caring for yourself (feeding, bathing, dressing, toilet) because of this health condition.		
	InHosp	You are in hospital and need help caring for yourself (feeding, bathing, dressing, toilet).		
Annual Costs	**Graded-Pair**		**Discrete-Choice**	
(Canadian $)	$10		$50	
	$25		$100	
	$50		$200	
	$100		$300	
	$200		$500	
	$500		$750	

Figure 1. Example of Graded-Pair Question

Category	Condition A	Condition B
Duration of Episode	5 days	1 day
Symptoms	Eye irritation (itching, burning, redness)	Pain in chest or arm
A. Daily Activities	• CAN go to work, go to school, do housework • CANNOT participate in social or recreational activities • Have SOME physical limitations • CAN care for yourself	• CANNOT leave your house, go to work, go to school, do housework, and participate in social or recreational activities • Have SOME physical limitations • CAN care for yourself
Total costs of this episode to your household	$50	$200

1	**2**	**3**	**4**	**5**	**6**	**7**
A is much better	A is somewhat better	A is slightly better	A and B are about equal	B is slightly better	B is somewhat better	B is much better

Please press a number from 1 to 7 that best reflects your rating.

Figure 2 illustrates the discrete-choice format used in this study. The discrete-choice format directly elicits total values for movements from a given diminished health state to the subject's current health state. The left-hand profile consists of a relatively severe hypothetical initial condition with zero cost. Alternatives A and B represent two courses a subject could choose if experiencing the initial condition. Alternative A portrays a condition of intermediate severity and intermediate cost. Alternative B is described as the subject's current health-state on the day of the survey with a relatively high cost. In essence, subjects can choose to remain in a relatively severe condition and not pay any additional costs for treatments outside the government health plan or insurance plan, or they can choose to pay for additional treatments to improve their health to their current health state.

Figure 2. Example of Discrete-Choice Question

Category	Initial Condition	Alternative A	Alternative B
Duration of episode	5 days	1 day	
Symptom	Shortness of breath and swelling in ankles and feet	Shortness of breath and swelling in ankles and feet	
Daily Activities	• Are in hospital • Need help caring for yourself	• CANNOT leave your house, go to work, go to school, do housework, or participate in social or recreational activities • Have SOME physical limitations • Need help caring for yourself	Your level of health as it is today. You do not experience this episode.
Additional costs to your household	$0	$300	$500

1	2	3
Prefer Initial Condition	Prefer Alternative A	Prefer Alternative B

Please press a number from 1 to 3 that best reflects your choice if faced with these options.

Unlike the graded-pairs section, the conditions that subjects evaluate must be related to obtain meaningful values. Therefore, while this format has the advantage of obtaining values for changes in health relative to an identified reference point, it potentially has the disadvantage of reducing the salience of the symptom attribute if subjects focus on changes in other attributes. Furthermore, the structure of the choice experiment assumes that subjects can exchange money for specified health improvements. For most people in the general population in good health, the choice sets include two options worse in health attributes and better in cost than current health, as intended. However, some subjects' current health may be worse than one or both of the hypothetical alternatives. In such cases, at least one of the options would be preferred, or dominant, in all of the attributes, and thus choices would reveal less information about trade-off relations.

The final survey instrument was developed through extensive pretesting, including two focus groups, three pretests, and a large-scale pilot test. The survey incorporates several state-of-the-art features including a computerized format, an information treatment with quiz questions, and detailed health-history questions. The instrument was administered to 399 randomly recruited subjects in the Toronto area between March and July 1997. Each subject answered 16 graded-pair or choice questions.

Analysis of Graded-pair Data

Graded-pair responses are ordinal ratings of utility differences between attribute-level pairs. Estimation strategy thus should account for the discrete, ordinal nature of the response variable. In this section, we describe a general model for estimating subject utility functions from such data and a procedure for calculating marginal WTP using the estimated utility functions.[12] We assume that individual indirect utility can be expressed as a function of commodity attributes and personal characteristics:

$$U^i_{st} = V^i(X_{st}, Z^i, P_{st}; \beta^i, \delta^i) + e^i_{st} \tag{1}$$

where:

- U^i_{st} is individual i's utility for commodity profile st, where s = L, R, denoting the left-side and right-side profiles for pair t, and t = 1,...,8
- $V^i(\cdot)$ is the nonstochastic part of the utility function,
- X_{st} is a vector of attribute levels in profile st,
- Z^i is a vector of personal characteristics,
- P_{st} is the cost of the commodity profile,
- β^i is a vector of attribute parameters,
- δ^i is the marginal utility of money, and
- e^i_{st} is a disturbance term.

The utility difference for profile pair t, dU^i_t, is simply:

$$dU^i_t = V^i_{Rt} - V^i_{Lt} + \varepsilon^i_t \tag{2}$$

where V^i_{Rt} and V^i_{Lt} are the indirect utilities associated with the right-side and left-side profiles, respectively, and $\varepsilon^i_t = e^i_{Rt} - e^i_{Lt}$ is the associated disturbance term. The disturbance term captures the effects of unobserved factors, including possible inherent ambiguity of subject preferences and cognitive errors.

The difference in indirect utility for commodity pair t, dU^i_t, often is specified as a simple linear function of attributes:

$$dU^i_t = V^i_{Rt} - V^i_{Lt} + \varepsilon^i_t = \left[\sum_h \beta_h \cdot X_{hRt} + \delta \cdot P_{Rt}\right] - \left[\sum_h \beta_h \cdot X_{hLt} + \delta \cdot P_{Lt}\right] + \varepsilon^i_t \tag{3}$$

where h indexes attributes. This specification assumes that attributes neither are substitutes nor complements for each other, so a change in the level of one attribute does not affect the marginal utility of any other attribute.[13]

---

[12] The first social-science application of this approach was by McKelvey and Zavoina (1975). For a recent application to environmental and health valuation, see Johnson and Desvousges (1997).
[13] See Keeney and Raiffa (1978) for an analysis of the properties of such utility functions.

The effects of personal characteristics on utility differences do not appear in Equation (3) because subject personal characteristics do not vary between the left and right sides of the screen. Controlling for such variables requires interacting them with commodity attributes or prices that do vary between profiles. For example, we can estimate the marginal utility of money as a function of individual characteristics Z^i and employ a functional form that allows for diminishing marginal utility of money.

$$\begin{aligned} dU^i_t &= V^i_{Rt} - V^i_{Lt} + \varepsilon^i_t \\ &= \left[X_{Rt}\cdot\beta + (Z^i\cdot\gamma)\cdot\sqrt{P_{Rt}}\right] - \left[X_{Lt}\cdot\beta + (Z^i\cdot\gamma)\cdot\sqrt{P_{Lt}}\right] + \varepsilon^i_t \end{aligned} \quad \textbf{(4)}$$

where γ is a vector of marginal-utility-of-money parameters.[14]

Similarly, the β coefficients on attributes also could be allowed to vary across subjects by making them functions of individual characteristics. It also is possible to employ more general nonlinear specifications of continuous attribute variables or interactions among attribute variables.

We do not observe dU^i_t directly. Instead, we observe C^i_t, which is a discrete rating category related to the unobserved dU^i_t of interest. The appropriate approach, therefore, is ordered logit or probit, which incorporates both the discreteness and the natural ordering of the data. This study uses ordered probit which assumes the error term is normally distributed. To estimate ordered-probit models, the data are sorted so that the preferred profile is on the right, making $dV^i_t = V^i_{tR} - V^i_{tL} \geq 0$.[15] We construct the rating categories by recoding responses accordingly, so that 0 indicates indifference and 3 indicates maximum difference.[16] Because probit assumes the Equation (3) error term ε^i_t is distributed $N(0,\sigma^2)$, the probability of observing response C^i_t is:

$$\text{Prob}(C^i_t = k) = \Phi\left[\mu^i\cdot\left(\alpha_k - dV^i_t\right)\right] - \Phi\left[\mu^i\cdot\left(\alpha_{k-1} - dV^i_t\right)\right] \qquad k = 0,1,\ldots,3 \quad \textbf{(5)}$$

where Φ is the cumulative standard normal distribution function, the α_k are threshold constants, and scale parameter μ^i is the inverse of the standard deviation.[17]

Because the health-valuation survey collects eight responses from each subject, it is appropriate to estimate a panel model that accounts for correlated errors in each subject's series of ratings. A random-effects model incorporates an individual-specific error term, so that

$$\varepsilon^i_t = \eta^i_t + \lambda^i \quad \textbf{(6)}$$

[14] The square root form was chosen over linear or log by estimating a Box-Cox model.

[15] This procedure assumes that subjects have no systematic preference for screen location.

[16] Because the original response scale indicates both which profile is preferred and how much it is preferred, this rearrangement maps response 3 into 5, 2 into 6, and 1 into 7. Response 7 indicates maximum utility difference and 4 indicates indifference, so C^i_t equals the recoded response minus 4.

[17] The maximum-likelihood procedure used to estimate the model parameters normalizes the α_o threshold at $-\infty$ and α_5 at $+\infty$ and does not include an intercept term.

where η_t^i is a common error term, and λ^i is an individual-specific error term distributed $N(0,\rho^2)$. Equation 5 now becomes

$$\text{Prob}(C^i = k_1^i, k_2^i, \ldots, k_8^i) = \int_{-\infty}^{\infty} \phi(\lambda^i) \cdot \prod_{t=1}^{8} \left\{ \Phi\left[\mu^i \cdot \left(\alpha_{k_t} - dV^i_t\right)\right] - \Phi\left[\mu^i \cdot \left(\alpha_{k_t - 1} - dV^i_t\right)\right]\right\} \; d\lambda^i \quad \textbf{(7)}$$

where $\phi(\lambda_i)$ is the normal probability density function for λ_i.[18]

Most ordered-category data contain no information on how scale might vary across subjects, and thus it usually is normalized uniformly to one. However, graded-pair data include multiple observations for each subject, and thus it is possible to obtain scale estimates. We can account for nonuniform scale by letting ε^i_t vary across subjects so that $\varepsilon^i_t \sim N(0,1/\mu_i^2)$.[19] Alternatively, we can make μ_i a systematically varying parameter that is a function of personal characteristics Y^i, such as income and education, so that $\varepsilon^i_t \sim N[0,1/\mu(Y^i)^2]$.

Estimating individual scale parameters provides a means of quantifying the coherence of individual rating patterns and of controlling for different levels of variance in subjects' error terms. Subjects who refuse to solve or have difficulty solving the utility-difference problem and who enter random or repetitive ratings will have unusually noisy ratings and thus larger estimated variance. Scale estimates thus provide a means of testing whether groups of subjects are experiencing problems with the survey design. The consistency of a subject's rating pattern can vary by degree of attentiveness, quantitative orientation, age, educational background, susceptibility to fatigue, and other factors. In addition, we can identify individual subjects who may have failed to perform the rating task properly. Finally, individual scale estimates help account for differences in how subjects interpret and use the 0-to-3 rating categories. Some subjects may be reluctant to use the entire range of ratings, which is a common problem in graded-pair elicitations (Mackenzie, 1993).

Analysis of Discrete-Choice Data

The SP survey also included a series of choice judgments with three alternatives in each choice set. The linear specification of utility analogous to Equation (4) for the three alternatives is:

$$\begin{aligned} U^i_{jt} &= V^i_{jt} + \varepsilon^i_{jt} && \equiv X_{jt} \cdot \beta + (Z^i \cdot \gamma)\sqrt{P_{jt}} + \varepsilon^i_{jt} && j = 1, 2 \\ U^i_{jt} &= V^i_0 + \delta^i_t + \varepsilon^i_{0t} && \equiv H^i \cdot \omega + (Z^i \cdot \gamma)\sqrt{P_{jt}} + \varepsilon^i_{jt} && j = 3 \end{aligned} \quad \textbf{(8)}$$

where U^i_{jt}, j = 1, 2, is the utility of each of the two alternative health profiles, and U^i_{jt}, j = 3, is the utility of the current-health alternative. V^i_0 is the subject's status-quo health-state utility, H^i is

[18] See Butler and Moffitt (1982) for derivations used in estimating this model.

[19] The way to generalize μ in a probit-type model is to retain the normalization for the base category that $\mu^i_o = 1$. For all other categories, $\mu^i_k = \mu^i$ is modeled as a function of individual characteristics. This guarantees that the estimated value of μ will be positive for every observation and that α_k adjusts to account for the normalization of the base category.

a vector of health-history variables, and ω are associated parameters. Other variables were defined previously.
Assuming ε follows a type-one extreme-value error structure, the probability that alternative j will be selected from choice set t is the standard conditional-logit expression:

$$\mathrm{Prob}(C^i_t = j) = \frac{\exp(\mu^i \cdot V^i_{jt})}{\sum_{k=1}^{3} \exp(\mu^i \cdot V^i_{kt})} \qquad \textbf{(9)}$$

where C^i_t is the selected alternative in each of 8 choice sets, μ^i is the scale parameter, and V^i_j is the determinate part of the utility of alternative j.[20] Thus, the probability that an alternative will be selected is the ratio of the utility that alternative provides relative to the sum of the utility that each alternative in the choice set provides. Note that individual characteristics fall out of this expression unless interacted with health attributes. The scale parameter μ^i also generally is not identifiable in such models and is normalized at one. However, because we have multiple observations for each subject, we can estimate $\mu^i = \mu(Z^i)$.

Like the ordered-probit analysis described above, the conditional logit model specified by Equations (6) and (7) is estimated using maximum-likelihood. That is, given the characteristics of the alternatives in the choice sets presented to subjects, the model estimates coefficients that maximize the likelihood that we would observe the actual choices in the sample. Thus, the coefficients show the relationship between the probability of selecting an alternative and the health attributes of that alternative.

Conditional logit models are known to be subject to violations of the irrelevance of independent alternatives assumption (IIA). This condition requires that the ratio of probabilities for any two alternatives be independent of the attribute levels in the third alternative. Tests of the choice data indicate that conditional logit estimates do not satisfy this requirement. Under these conditions, parameter estimates are biased. Furthermore, conditional logit does not account for correlations within each subject's series of choices. Revelt and Train (1998) recently have proposed using random-parameter logit (RPL) for SP data similar to ours. RPL is not subject to the IIA assumption, accommodates correlations among panel observations, and accounts for uncontrolled heterogeneity in tastes across subjects.
Modifying Equation (8) to introduce subject-specific stochastic components for each β:

$$U^i_{jt} = V^i_{jt} + \varepsilon^i_{jt} \qquad \equiv X_{jt} \cdot (\beta + \eta^i) + (Z^i \cdot \gamma)\sqrt{P_{jt}} + \varepsilon^i_{jt} \qquad \textbf{(10)}$$

Equation (9) now becomes:

$$\mathrm{Prob}\left[C^i = \left(C^i_{j1}, C^i_{j2}, \ldots C^i_{j8},\right)\right] = \prod_{t=1}^{8} \left[\frac{\exp\left[\mu^i \cdot V^i_{jt}(\beta *)\right]}{\sum_{k=1}^{3} \exp\left[\mu^i \cdot V^i_{kt}(\beta *)\right]} \right] \qquad \textbf{(11)}$$

[20] The basic exposition of the properties of this model can be found in McFadden (1981).

where now $\beta^* = (\beta + \eta^i)$. In contrast to conditional logit, the stochastic part of utility now may be correlated among alternatives and across the sequence of choices via the common influence of η^i. McFadden and Train (1997) show that any random-utility model can be approximated by some RPL specification.

Calculating Willingness to Pay

Estimating the parameters of the utility function enables us to quantify the value of changes in commodity attributes. The units of the estimated utility index essentially are arbitrary. However, this arbitrary measure can be converted into a dollar metric using the estimated marginal utility of money. Let X^o_j indicate the status quo vector of attribute levels. In our case, this corresponds to the subject's current health state. The cost of X^o_j is P^o_j, which we take to be zero. X^*_j indicates a changed vector of attribute levels corresponding to a given combination of symptom and activity level. The willingness to pay for a given change in commodity attributes $(X^*_j - X^o_j)$ is the amount of money $(P^*_j - P^o_j)$ that would leave subject i indifferent between the payment and the change in attribute levels, so that P^*_j satisfies

$$V^i[X^*_j, Z^i, P^*_j; \beta^i, \delta^i(P, Z^i)] = V^i[X^o_j, Z^i, P^o_j; \beta^i, \delta^i(P, Z^i)] \quad \textbf{(12)}$$

In the linear specification of Equations (3) and (9), the β coefficient for each attribute represents its constant marginal utility. The negative of the coefficient of the price attribute is interpreted as the marginal utility of money, or the utility derived from more dollars. Thus,

$$WTP^i(X^*_j - X^o_j) = \frac{\sum_h (X^*_{hj} - X^o_{hj})\, {}^{\partial V^i}\!/_{\partial X_{hj}}}{-\partial V^i / \partial P} = \frac{\sum_h (X^*_{hj} - X^o_{hj})\beta^i_h}{-(Z^i \cdot \gamma) / (2 \cdot \sqrt{P})} \quad \textbf{(13)}$$

where h indexes attributes.

Allowing the marginal utility of money to vary with personal characteristics allows for differences in tastes to affect the relative utilities of health and money, even though the β_j parameters are constant across subjects. This specification facilitates transferring WTP estimates to populations with different demographic characteristics than our sample. Thus, because WTP estimates are obtained by dividing the health coefficients by the marginal utility of money, changing values of the personal characteristics included in the utility of money function affects the WTP estimates.

Any payment less than or equal to $WTP^i(X^*_j - X^o_j)$ leaves individuals at least as well off as they would be if the change $(X^*_j - X^o_j)$ had not occurred. Because we recover the parameters for a complete utility index, WTP can be constructed for any utility difference, assuming that the marginal utility of money is constant for all utility differences of interest.[21]

Graded-Pairs estimates

Table 2 contains the definitions of the variables used in all analyses in this study. Attributes are effects-coded rather than dummy-variable coded, so the omitted categories NOSE and NOLIM are the negative sum of the included symptom, interacted with the log of duration, and activity

[21] Because utility is a nonlinear function of price, WTP is a function of some assumed price level. The WTP calculations reported below incorporate average price combinations that occur with different symptom/activity levels.

categories.[22] The daily activity restriction levels represent the additional disutility of experiencing these restrictions compared with experiencing no limitations (NOLIM). We report both nominal and normalized coefficients. The normalized coefficients rescale the health index from zero to one, with the INHOSP coefficient equal to zero (worst) and the NOLIM coefficient equal to one (best).

The graded-pairs experimental design precludes particular attribute combinations from appearing in the paired comparisons in order to make the bundles more credible to subjects. For example, the relatively mild symptoms "Stuffy/runny nose and sore throat" (NOSE) and "Eye irritation" (EYE) are never seen with activity restrictions greater than "Social and recreation limitations" (NOSOC), nor with costs greater than $100.[23] Similarly, more severe conditions are never seen with the mildest activity-restriction level NOLIM.

Table 2. Variables Used in Analysis

VARIABLES	DESCRIPTION
NOSE	Stuffy or runny nose and sore throat
EYE	Eye irritation
FLUTTER	Fluttering in chest and feeling light-headed
BREATH	Coughing, wheezing, shortness of breath
ACHE	Coughing or wheezing with fever, chills or aching all over
SWELL	Shortness or breath, and swelling in ankles and feet
PAIN	Pain in chest or arm
LNDAYS	Log of the number of days of episode (1, 5, 10) plus one
SOMELIM	Activity Level 2: You can go to work, go to school, do housework, and participate in social or recreational activities, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition.
NOSOC	Activity Level 3: You can go to work, go to school, do housework, but you have some physical limitations (trouble bending, stooping, or doing vigorous activities) and cannot participate in social or recreational activities because of this health condition.
ATHOME	Activity Level 4: You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you have some physical limitations (trouble bending, stooping, or doing vigorous activities) because of this health condition, but you can care for yourself.
NEEDHELP	Activity Level 5: You cannot leave your house, go to work, go to school, do housework, participate in social or recreational activities, and you need help caring for yourself (feeding, bathing, dressing, toilet) because of this health condition.
INHOSP	Activity Level 6: You are in hospital and need help caring for yourself (feeding, bathing, dressing, toilet).

[22] This specification facilitates comparisons between graded-pair and choice formats. Symptom was held constant within choice sets in the choice questions, so it is necessary to interact duration with symptom in the choice models. Transforming duration allows for diminishing marginal utility of duration.

[23] Thus NOSE and EYE are never seen with ATHOME, NEEDHELP, and INHOSP, as well as values of $200 and $500.

SQTCOST	Square root of the cost levels per year (Can$10–Can$750)
SCORE	Quiz score (percent correct)
AGE	The midpoint of the age category
AGESQUARED	The square of AGE
EDUCATION	Number of years of education
SYMPTOMATIC	Dummy variable = 1 if subject has ever been diagnosed with any cardiovascular or respiratory conditions or other serious illnesses
VIMPORTANT	Number of health factors subjects indicated as being very important to a person's health (0-4)
HISCORE	= 1 if subject scored 75% or better on the quiz
HIAGE	= 1 if subject is 60 years old or older
NOPAIDLEAVE	= 1 if subject is working but does not have paid sick leave
FAILCHECKS	= 1 if subject provided questionable or inconsistent answers in parts of the survey
HIINFORMATION	= 1 if subject acquires health information a few hours per week or more
ASTHMA	= 1 if subject has experienced asthma in past year
BRONCHITIS	= 1 if subject has experienced chronic bronchitis/ emphysema in past year
LUNGINFECTION	= 1 if subject has experienced lung infection (pneumonia, acute bronchitis) in past year
HEARTDISEASE	= 1 if subject has experienced heart disease symptoms in past year
INCOME	Household income before taxes in 1996
PAINRATING	Rating of pain or discomfort on a scale of 0 (no discomfort) to 8 (severe discomfort), summed over all symptoms.
FREQILL	Frequency of illness on a scale of zero (never) to four (almost all or all the time), summed over all symptoms.
FREQHOME	Frequency of staying home because of illness in past year on a scale of zero (never) to four (almost all or all the time), summed over all symptoms.
CH_A	Alternative specific dummy for the first alternative in each choice set (Initial Condition)
CH_B	Alternative specific dummy for the middle alternative in each choice set (Alternative A)
RHO	Random effects correlation coefficient for subject's eight SP questions

Table 3 reports estimates for the graded-pair, ordered-probit panel model. This model estimates and controls for the correlation among the subject's eight SP answers. The correlation coefficient, RHO, shows that there is significant correlation among the eight SP questions. Controlling for this correlation affects the significance of several variables.[24]

[24] See Johnson, et al. (1998) for comparisons with non-panel ordered-probit estimates.

Symptom attribute levels are interacted with the log of duration to facilitate comparisons with the discrete-choice estimates. There is no natural ordering of symptom disutility, so we have no specific expectations about relative magnitudes. Nevertheless, it is not surprising that the coefficients on NOSE and EYE indicate higher utilities than the other symptoms. Overall, there is relatively little variation in the normalized symptom coefficients compared to the activity-level coefficients. Nevertheless, five of the seven symptoms are significantly different than the overall mean symptom effect.

Unlike the symptom variables, the activity-restriction levels have a natural ordering with “no limitations” (NOLIM) representing the best activity-restriction level and “in hospital” (INHOSP) representing the worst activity-restriction level. All of the coefficients on the activity-restriction levels are significantly different than the mean activity effect. The coefficients of SOMELIM and NOSOC are reversed in order, but are statistically equivalent, indicating that subjects did not have strong preferences for not being able to participate in social and recreational activities and having physical limitations versus simply having physical limitations. The remaining coefficients decrease monotonically, as expected, indicating that greater activity restrictions result in greater utility losses.

Except for VIMPORTANT, the utility of money covariates are significant, suggesting that different groups of subjects have different preferences for health-state/money tradeoffs. Subjects with higher quiz scores have lower WTP for improved health, while older, more educated, and symptomatic subjects have higher WTP. The significance of the scale parameters indicates that subjects used the rating scale differently. Of the significant coefficients in the scale function, subjects with high quiz scores and no paid sick leave had less noisy responses, while elderly subjects and subjects who failed more consistency checks had more noisy responses.

Table 3. Graded-Pair Ordered-Probit Panel Model

Variable	Coefficient		Normalized	T-ratio
Health Attributes				
NOSE*LNDAYS	0.2720	***	0.6096	4.38
EYE*LNDAYS	0.1610	***	0.5432	6.33
FLUTTER*LNDAYS	-0.0862	***	0.3953	-3.38
BREATH*LNDAYS	-0.0301		0.4288	-1.34
ACHE*LNDAYS	-0.1426	***	0.3615	-5.13
SWELL*LNDAYS	-0.1559	***	0.3535	-7.39
PAIN*LNDAYS	-0.0182		0.4360	-0.63
NOLIM	0.9242	***	1.0000	6.88
SOMELIM	0.0803	***	0.4949	2.52
NOSOC	0.1243	**	0.5212	3.75
ATHOME	-0.1026	**	0.3854	-2.01
NEEDHELP	-0.2796	***	0.2795	-7.22
INHOSP	-0.7466	***	0.0000	-6.87
Utility of Money Function (*SQTCOST)				
CONSTANT	-0.021935			-1.22
SCORE	-0.0002413	**		-2.18
AGE	0.000392	**		2.13
EDUCATION	0.0004650			0.46
SYMPTOMATIC	0.009791	**		2.13
VIMPORTANT	-0.003691			-1.61
Scale Function				
HISCORE	0.1502	***		2.69
HIAGE	-0.2579	***		-4.09
VIMPORTANT	-0.0003			-0.01
NOPAIDLEAVE	0.1024	*		1.70
HIINFORMATION	-0.0423			-0.63
FAILCHECKS	-0.2967	***		-5.10
Constants				
ALPHA1	-1.2148			
ALPHA2	-0.1648			
ALPHA3	0.6591			
RHO	0.1260	***		5.95
Likelihood Ratio Chi-sq.	536***			
McFadden R-square	0.076			
Percent correctly predicted	42.4			
Number of observations	2752			

*** Significant at the 1-percent level or better

** Significant at the 5-percent level or better

* Significant at the 10-percent level or better

Discrete-choice estimates

We estimated conditional-logit choice models that are widely used in the market research literature.[25] Nevertheless, there is reason to be concerned about several sources of bias in such estimates. First, the conditional logit models assume that differences in WTP across subjects

[25] These results are not reported here. See Johnson, et al. 1998.

arise only from differences in the marginal utility of money. The utility weights for health attributes are assumed to be the same for all subjects. Second, unbiased conditional logit estimates require satisfying the IIA assumption. Tests of the choice data indicate that conditional logit estimates do not satisfy this requirement. Under these conditions, parameter estimates are biased. Finally, conditional logit estimates assume errors in each subject's series of answers are uncorrelated. The model assumes, in effect, that we sampled 2,752 subjects instead of asking eight questions of 344 subjects.

As discussed above, random-parameters logit (RPL) avoids all three of these potential sources of bias. Table 4 reports the RPL results. Because each parameter includes both a systematic and random component, the model estimates a mean and standard error for each distribution. Although we can construct parameters for the omitted categories NOSE and NOLIM, there are no corresponding standard-deviation estimates for those variables.

Unlike the graded-pair format, subjects evaluated the choice alternatives relative to their current health, which always was represented by the third alternative across choice sets. In addition to the same health attribute and utility-of-money variables, the discrete-choice RPL estimates include standard-deviation estimates for the health attributes, current health covariates, and two alternative-specific constants that indicate the probability of choosing the first or second relative to the third alternative.

We again have effects-coded the health attributes, so the coefficients on the omitted attributes NOSE and NOLIM are reported as the negative sum of the included categories. We interpret statistical significance relative to mean effects for the health attributes. Symptom variables are interacted with duration because symptom was held constant within each choice set. Table 4 also reports both nominal and normalized coefficients.

Table 4. Choice Random-Parameter Logit Panel Model

Variable	Coefficient		Normalized	T-ratio
Health Attributes				
NOSE*LNDAYS	0.4522	**	0.9675	2.12
Estimated St. Dev.	---			
EYE*LNDAYS	0.1653	**	0.7995	2.07
Estimated St. Dev.	-0.0001	***		0.00
FLUTTER*LNDAYS	-0.1851		0.5944	-2.17
Estimated St. Dev.	0.0134			0.02
BREATH*LNDAYS	-0.0786		0.6568	-0.94
Estimated St. Dev.	-0.0081			-0.01
ACHE*LNDAYS	0.1101		0.7672	1.19
Estimated St. Dev.	0.0074			0.01
SWELL*LNDAYS	-0.1916	*	0.5906	-1.99
Estimated St. Dev.	-0.1016			-0.29
PAIN*LNDAYS	-0.2723	***	0.5434	-3.23
Estimated St. Dev.	-0.0273			-0.04
NOLIM	0.5078	*	1.0000	1.78
Estimated St. Dev.	---			
SOMELIM	0.4864	***	0.9875	4.87
Estimated St. Dev.	0.3133	**		2.10
NOSOC	0.5998	***	1.0538	6.06
Estimated St. Dev.	-0.4752	***		-3.57
ATHOME	0.0038		0.7050	0.04
Estimated St. Dev.	-0.5444	***		-3.82
NEEDHELP	-0.3971	***	0.4704	-3.30
Estimated St. Dev.	0.5104	***		2.59
INHOSP	-1.2007	***	0.0000	-6.22
Estimated St. Dev.	-1.4752	***		-10.01
Utility of Money Function (*SQTCOST)				
CONSTANT	-0.1967	***		-9.35
SCORE	-0.0000276	***		-3.69
AGE	0.000323			1.16
EDUCATION	0.000716	***		8.81
SYMPTOMATIC	0.008012	*		1.72
VIMPORTANT	-0.011477	***		-3.33
Current Health Function				
ASTHMA* FREQHOME	-0.0953	***		-4.39
BRONCHITIS	-0.7115	***		-3.21
LUNGINFECTION	-0.2569	***		-4.42
HEARTDISEASE* FREQHOME	-0.1335	***		-3.98
AGESQUARED	0.0000508	***		2.56
INCOME	0.002608	***		7.46
PAINRATING	0.003662	***		3.72
VIMPORTANT	0.0986	***		4.54
FREQILL	-0.1030	***		-6.94
FREQHOME	0.2059	***		7.14
Alternative-Specific Constants				
CH_A	-1.2888	***		-4.29
CH_B	0.2423			1.24
Likelihood Ratio Chi-Square	557***			
McFadden R-Square	0.093			
Percent correctly predicted	39.3			
Number of observations	2752			

*** Significant at the 1-percent level or better ** Significant at the 5-percent level or better * Significant at the 10-percent level or better

Symptom was allowed to vary within pairwise comparisons in the graded-pair format but was held constant within choice sets. Because of this difference in the experimental design, we expected to see corresponding differences in estimated symptom utility weights. In particular, we expected to see less variation among the symptom coefficients compared to the graded pairs. Instead, we observe more variation but lower statistical significance relative to the mean effect. Moreover, all the normalized choice coefficients are larger than the corresponding pairs coefficients. PAIN is now the worst condition and ACHE now ranked as one of the milder conditions rather than one of the more severe conditions, as indicated in the pairs estimates.

In contrast to the symptom attribute, all activity levels but ATHOME are statistically significant. Again, except for the normalized endpoints, the choice activity coefficients are larger than the corresponding pairs activity coefficients. In particular, we observe very little difference in the choice NOLIM and SOME activity levels, while there is a large decline in utility between NOLIM and SOME in the pairs estimates. We continue to see NOSOC rated higher than SOME, but the difference again is insignificant. We observe the same pattern of signs in the marginal-utility-of-money function as in the graded-pair model. However, the constant term, VIMPORTANT, and EDUCATION now are significant, while AGE is not. Because of differences in scaling, it is not possible to directly compare magnitudes of utility-of-money parameter differences between the two models.

All four of the respiratory and cardiovascular conditions in the current-health function are statistically significant and indicate worse health than overall average health in the sample. ASTHMA and HEARTDISEASE were insignificant when entered independently, but significant when interacted with FREQHOME. Thus ASTHMA and HEARTDISEASE are associated with lower than average health only when they are severe enough to require staying home. Of the four conditions, chronic bronchitis is by far the most serious, with disutility 2.5 times greater than the next most serious condition, while asthma is the least serious.

Age-squared also is positive and highly significant. Generally we expect that older people experience lower levels of health than younger people. It is worth noting, however, that we are modeling subjective health utility, not health per se. This result may be interpreted as saying that older people obtain higher levels of utility from the health they have. This is consistent with the positive sign on AGE in the utility-of-money function, which says that older people value money less than health compared to younger people.

Recall that the experimental design required that cost be ordered from zero to high values among alternatives A, B, and C, while activity restrictions are ordered inversely.[26] Thus some subjects may have tended to choose B simply because it had better health than A and lower cost than C. CH_B is significant in models where money and current health functions are treated as constants, but insignificant when these functions are included. This result means that including utility of money and current health functions eliminates the middle-alternative bias. However, the significant negative coefficient for A indicates that subjects chose A relative to B and C less often than the explanatory variables alone would predict.

Joint graded-pair, discrete-choice estimates

Our goal is to obtain estimates suitable for use in a wide range of benefit-cost applications. Because each SP format has strengths, as well as limitations, we pool the data and estimate a joint

[26] Note that Figure 3 labels the three alternatives that were presented to subjects as the Initial Condition, Alternative A, and Alternative B. However, we refer to these alternatives in the analysis section as Alternatives A, B, and C, respectively.

model to capture all the available information on subjects' preferences in one set of parameters. For example, the subjects paid more attention to symptoms in the pairs format than in the choice format, while they paid more attention to costs in the choice format. Estimating overall preferences in a joint model thus provides a more complete picture of subjects' utility functions for calculating WTP than does either format individually.

Although it would be interesting to estimate a joint panel model, such estimates are beyond the scope of the present study. Instead, we estimate a pooled model using the same specifications as the panel models discussed above, but do not control for intra-subject correlations. We then compare the results with the individual panel models.
From equation (1),

$$U^i_{st} = \lambda_f \cdot V^i(X_{st}, Z^i, P_{st}; \beta, \delta^i_s) + e^i_{st} \qquad \textbf{(14)}$$

where now s indexes an attribute profile in either a graded-pair or choice repetition. We include a choice-utility scale function to control for different variances in the two data sets, so $\lambda_f=\lambda_c$ for the choice format and $\lambda_f=1$ for the graded-pair format.[27] The utility parameters β are constrained to be the same for both elicitation formats. However, the utility of money function parameters are estimated separately. Thus,

$$\delta^i_f(Z^i) = \left[\gamma_{fo} + \sum_m \left(\gamma_{fm} \cdot Z^i_m\right)\right] \qquad \textbf{(15)}$$

where again f indexes the graded-pair and choice formats, respectively. We include the same utility of money variables used in previous models. The joint model also incorporates the current health condition variables used previously, so that WTP values can be estimated relative to different individual current-health reference points. Information on current-health utility comes only from the choice format. Finally, we combine SOME and NOSOC into a single category. Previous results indicated that subjects discriminated poorly between these two mild-restriction categories.

[27] The joint model omits the within-data set scale parameters, which had negligible effect on other parameters and do not enter into WTP calculations.

Figure 3. Comparison of Joint and Individual-Panel Models

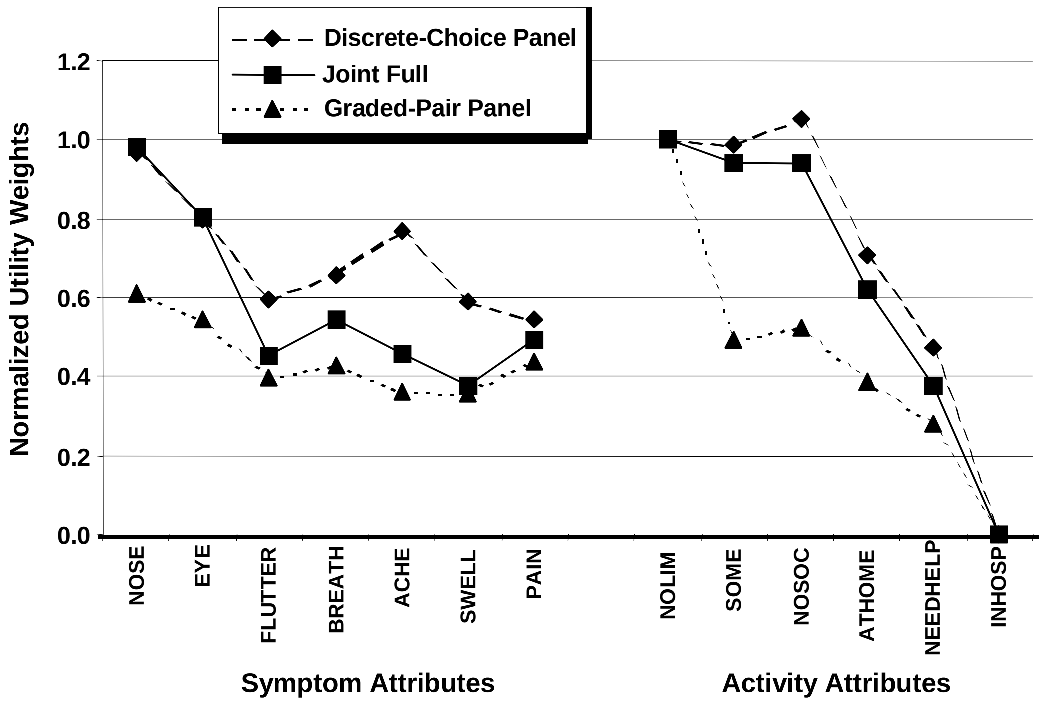

Figure 3 compares the joint estimates with the discrete-choice and graded-pair panel models. The joint estimates fall between the two panel estimates in nearly every case. The joint estimates track the graded-pair point estimates somewhat more closely than the discrete-choice estimates for the more severe symptoms. Conversely, the joint estimates track the discrete-choice estimates more closely for the activity categories. Thus the joint estimates have the desired property of giving more weight to the format with the smaller variance for each parameter.

Table 5 reports estimates for the joint model. As in previous models, we present both nominal and normalized coefficients. The pattern of normalized symptom utility weights reflects the variability in the graded-pair model rather than the discrete-choice model, as we hoped. The coefficients for NOSE and EYE are significantly different from the other symptom variables. Only BREATH is significantly different than the other conditions at the 10 percent level. Thus, the results of the symptom levels show similar variation to the pairs symptoms but some of the significant differences among the more severe symptoms have been attenuated, reflecting the influence of the discrete-choice estimates.

As in the choice model, only ATHOME is statistically insignificant relative to the mean activity effect. Furthermore, all activity levels are significantly different from each other and decrease monotonically as daily activity restrictions become more severe.

Utility of money functions are estimated separately for the graded-pair and discrete-choice data to account for systematic differences between the two formats. The implicit marginal utility of money is much smaller in the graded-pair data than the discrete-choice data. There are two possible reasons for the large difference in marginal effects. First, the nature of the discrete-choice task induces subjects to focus on the attribute levels themselves, rather than on differences among the attribute levels. Most people solve the choice valuation problem by eliminating undesirable alternatives based on one or two unattractive attribute levels. Consequently, the task itself simply may cause subjects' to pay more attention to the cost attribute in evaluating alternatives. Second, because symptoms were held constant across two of the three alternatives, subjects did not have to pay as much attention to this attribute as they did in the pairs task. Therefore, they may have had more cognitive energy to focus on the other

Table 5. Joint Model

Variable	Coefficient		Normalized	T-ratio
Health Attributes				
NOSE*LNDAYS	0.2572	***	0.9829	5.52
EYE*LNDAYS	0.1405	***	0.8035	7.04
FLUTTER*LNDAYS	-0.0871	***	0.4535	-4.59
BREATH*LNDAYS	-0.0294	*	0.5422	-1.69
ACHE*LNDAYS	-0.0831	***	0.4596	-4.08
SWELL*LNDAYS	-0.1372	***	0.3764	-7.91
PAIN*LNDAYS	-0.0609	***	0.4938	-3.05
NOLIM	0.2683	***	1.0000	4.47
SOMELIM/NOSOC	0.2299	***	0.9410	9.39
ATHOME	0.0225		0.6220	0.84
NEEDHELP	-0.1387	***	0.3741	-5.36
INHOSP	-0.3820	***	0.0000	-9.48
Utility of money function (*SQTCOST)				
CH_CONSTANT	-0.1072	***		-6.54
CH_SCORE	-0.0001819	**		-2.17
CH_AGE	0.000297	*		1.65
CH_EDUCATION	0.003968	***		5.13
CH_SYMPTOMATIC	0.003815			1.03
CH_VIMPORTANT	-0.004019	**		-2.09
GP_CONSTANT	-0.007188			-0.40
GP_SCORE	-0.0002305	**		-2.22
GP_AGE	0.000432	**		2.44
GP_EDUCATION	-0.0005507			-0.53
GP_SYMPTOMATIC	0.008877	*		1.82
GP_VIMPORTANT	-0.004918	**		-2.17
Current Health Condition				
CH_ASTHMA* FREQHOME	-0.0658	***		-2.58
CH_BRONCHITIS	-0.3523	**		-2.29
CH_LUNGINFECTION	-0.1679	***		-2.53
CH_HEARTDISEASE* FREQHOME	-0.1621	***		-2.50
CH_AGESQUARED	0.0000267	*		1.62
CH_INCOME	0.001624	***		4.39
CH_PAINRATING	0.002615	***		2.46
CH_VIMPORTANT	0.0466	***		3.27
CH_FREQILL	-0.0496	***		-2.85
CH_FREQHOME	0.1124	***		3.44
Relative Scale Function				
CH_HISCORE	0.9345	***		5.03
CH_HIAGE	-0.8042	***		-3.69
CH_VIMPORTANT	0.0718			1.20
CH_NOPAIDLEAVE	-0.1327			-0.77
CH_HIINFORMATION	-0.3505	*		-1.80
CH_FAILCHECKS	0.1819			0.89

Table 5. Joint Model (Continued)

Variable	Coefficient		Normalized	T-ratio
Alternative-Specific Constants				
CH_A	-0.5327	***		-3.28
CH_B	0.2809	**		2.14
GP_ALPHA1	-1.1330			
GP_ALPHA2	-0.1635			
GP_ALPHA3	0.6159			
Likelihood Ratio Chi-Sq.	827***			
McFadden R-square	0.063			
Graded-Pair Percent Correctly Predicted	42.1			
Choice Percent Correctly Predicted	47.5			
Number of Observations	5504			

***Significant at 1-percent level **Significant at 5-percent level *Significant at 10-percent level

attributes, including cost. The other utility of money coefficients exhibit similar patterns as in the previous two models.

The relative scale function fixes differences in variance in the two data sets. We use the same scale variables as before, but their interpretation is different in this model. The relative scale function indicates how the variance or noise varies systematically in the choice data set relative to the pairs data set. A positive coefficient indicates a source of lower variance in the choice data. Overall, this function is positive, indicating that the pairs data contains significantly more noise than the discrete-choice data. This result is not surprising because of the more cognitively burdensome nature of the pairs task.

Willingness-to-Pay Estimates

As discussed above, we can convert the parameters of the utility index shown in the previous models into a dollar metric by rescaling the index using the marginal utility of money. Employing SP techniques allows us to recover WTP estimates for all relevant combinations of symptoms and activity levels in the experimental design.[28,29]

The confidence interval is obtained by drawing 1000 times from the multivariate normal distribution of coefficients and their variance-covariance matrix. The foregoing calculation is performed for each draw, and the confidence-interval lower and upper bounds are the fifth and ninety-fifth percentile values.

Table 6 shows WTP estimates and 90-percent confidence intervals to avoid one-day, five-day, and ten-day episodes of the relevant symptom and activity level combinations, given subjects' baseline health for the past year. The first column of estimates shows WTP estimates for the MILD activity restriction, which combines SOMELIM and NOSOC. We show zero WTP for one-day episodes of NOSE, EYE, and BREATH. The point estimates for these symptoms are negative, meaning that the estimated average baseline current health in our sample actually was less than the estimated utility for these outcomes. Moving from these health states to current health would reduce welfare, hence the negative point estimate for WTP. However, the upper bound of the 90-percent confidence interval for BREATH is positive. This result is not surprising considering that approximately 36 percent of our 344 subjects have experienced cardiovascular and respiratory conditions within the past year. In addition, nonsymptomatic subjects may suffer from a variety of other ailments. Thus, the results show that on average this sample of subjects is not willing to pay to avoid an episode of these mild conditions.

---

[28] Some combinations of attribute levels, such as having a stuffy nose and sore throat and being in hospital, are not plausible and thus WTP estimates for these combinations are not presented because they fall outside the range of the experimental design.

[29] For example, the average willingness-to-pay to avoid one day of fluttering in the chest and feeling light-headed that confines subjects to their homes but allows them to care for themselves is calculated as follows:

$$\text{WTP(BREATH, ATHOME, Days)} = \frac{\text{Current Utility(Days)} - \text{Utility(BREATH, ATHOME, Days)}}{\text{Marginal Utility of Money}}$$

$$= \frac{\text{Ln}(1+1)\cdot V_{\text{Current}} - [\text{Ln}(1+1)\cdot \beta_{\text{BREATH}} + \beta_{\text{ATHOME}}]}{\frac{1}{2}\cdot\left[\frac{-dV}{d\text{Cost}^{CH}} + \frac{-dV}{d\text{Cost}^{GP}}\right]}$$

$$= \frac{0.6931\cdot 0.2955 - [0.6931\cdot(-0.0294) + (0.0225)]}{\frac{1}{2}\cdot\left[\frac{0.01975}{2\cdot\sqrt{375}} + \frac{0.05859}{2\cdot\sqrt{202}}\right]}$$

$$= 158$$

where CH and GP refer to choice and graded-pair estimates, respectively and cost is averaged over all the screens where these two health attribute levels appear for each question format. Using more precise coefficients and means, the actual value is $208 with a bootstrapped 90 percent confidence interval of $129 to $287.

Table 6. Willingness to Pay Estimates for Each Symptom/Activity Level Combination for One-Day, Five-Day, and Ten-Day Episodes (1997 Canadian $)

		Daily Activity Levels			
Symptoms	**Duration**	**MILD[a]**	**ATHOME**	**NEEDHELP**	**INHOSP**
NOSE	1 day	0[b]	3	-[c]	-[c]
		(-168/-83)	(-50/57)		
	5 days	0[b]	29	-[c]	-[c]
		(-204/4)	(-89/148)		
	10 days	0[b]	44	-[c]	-[c]
		(-225/53)	(-110/199)		
EYE	1 day	0[b]	54	-[c]	-[c]
		(-119/-32)	(-2/108)		
	5 days	29	162	-[c]	-[c]
		(-76/138)	(44/282)		
	10 days	85	222	-[c]	-[c]
		(-53/232)	(67/379)		
FLUTTER	1 day	24	171	288	427
		(-26/73)	(114/236)	(232/359)	(352/517)
	5 days	310	468	588	704
		(192/438)	(341/615)	(462/741)	(568/870)
	10 days	468	631	754	857
		(309/641)	(466/825)	(589/955)	(684/1065)
BREATH	1 day	0[b]	158	286	448
		(-60/49)	(100/225)	(224/363)	(371/536)
	5 days	266	435	566	712
		(141/405)	(299/589)	(427/732)	(566/872)
	10 days	415	589	721	857
		(249/602)	(411/789)	(537/937)	(668/1058)
ACHE	1 day	26	199	335	513
		(-36/86)	(133/275)	(267/420)	(426/623)
	5 days	362	544	682	845
		(214/523)	(382/722)	(529/866)	(677/1046)
	10 days	547	734	873	1027
		(350/762)	(524/967)	(668/1114)	(813/1278)
SWELL	1 day	56	229	365	535
		(-1/117)	(164/306)	(295/452)	(443/644)
	5 days	439	621	761	908
		(296/598)	(469/799)	(598/954)	(737/1111)
	10 days	650	837	979	1114
		(459/869)	(634/1073)	(767/1228)	(896/1375)
PAIN	1 day	14	190	329	510
		(-48/75)	(125/263)	(260/411)	(426/611)
	5 days	338	522	663	827
		(193/489)	(370/693)	(508/841)	(665/1015)
	10 days	516	705	848	1002
		(322/719)	(503/934)	(645/1084)	(791/1240)

[a] MILD combines SOMELIM and NOSOC .
[b] Negative point estimate shown as zero WTP. See text for interpretation.
[c] These combinations fall outside the scope of the experimental design.

Table 6 also shows that as activity restrictions increase for a given symptom (that is, as one moves to the right from "no limitations" to "in hospital"), WTP also increases. Differences between WTP estimates in adjacent activity levels generally are not statistically significant. Examining how WTP changes across symptoms for a given activity level reveals that differences are even less significant. Overall, WTP estimates show the most dramatic differences across activity levels, and while there are differences among the symptoms, these differences appear to be less salient to subjects for a given activity limitation.

Conclusions

Using two different SP elicitation formats is a significant improvement over previous health-valuation studies. The two formats induce subjects to employ different evaluation strategies. Thus, combining both formats allows us to capitalize on the information provided by each format and on the different cognitive processes that subjects use in each situation. The two formats also differ in their formal utility-theoretic basis. The graded-pair format elicits values for marginal tradeoffs among health-states, whereas the discrete-choice format elicits total values to avoid a given condition relative to the subject's current health. Thus, each elicitation format offers certain advantages, and we believe that employing both formats provides more robust and valid estimates.

Comparisons of WTP estimates from the joint model with the few published estimates for similar conditions based on conventional CV methodology indicates our estimates are not systematically larger or smaller (Johnson, et al., 1998). However, unlike other studies in the WTP literature, our SP estimates are derived from complete multi-attribute health-state descriptions that are clearly specified and consistent across the same sample of subjects. Thus we are able to explicitly account for the separate effects of symptoms, duration, activity restrictions, and current health on WTP.

The results demonstrate the feasibility of estimating meaningful WTP values for policy-relevant respiratory and cardiac symptoms, even from subjects who never have personally experienced these conditions. Furthermore, because WTP estimates are for individual components of health improvements, estimates can be aggregated in various ways depending upon policy needs. Thus using generic health attributes facilitates matching outcome values more accurately than previously was possible to evaluate a variety of health-care interventions and policies.

The ultimate purpose of conducting this study was to generate valuation estimates that can be applied in benefit-cost analysis. Applications in the area of environmental health could include federal and provincial programs to reduce sulfur dioxide, ozone, particulates, and other pollutants that can aggravate heart and lung conditions. In addition, valuation estimates could be used by electric utilities and other industries to assess the benefits of pollution-control equipment, evaluate alternative power-generation options, establish optimal emission caps for emissions trading, assess the dispatching of fossil-generation stations to meet excess demand, using full-cost accounting, and evaluate the costs and benefits of imports and exports of electricity from and to the United States.

References

Adamowicz, W., J. Louviere, and M. Williams. 1994. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities." *Journal of Environmental Economics and Management* 26:271–292.

Adamowicz, Wiktor, Joffre Swait, Peter Boxall, Jordan Louviere, and Michael Williams. 1997. "Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation." *Journal of Environmental Economics and Management* 32(1):65–84.

Bryan S., M. Buxton, R. Sheldon, and A. Grant. 1997. *The Use of Magnetic Resonance Imaging (MRI) for the Investigation of Knee Injuries: A Discrete-Choice Conjoint Analysis Exercise.* Paper presented at the Health Economists' Study Group Meeting, University of York, July.

Butler, J., and R. Moffitt. 1982. "A Computationally Efficient Quadrature Procedure for the One Factor Multinomical Probit Model." *Econometrica* 50(3):761-764.

Carson, Richard T, Jordan J. Louviere, Donald A. Anderson, Phipps Arabie, David S. Bunch, David A. Hensher, Richard M. Johnson, Warren F. Kuhfeld, Dan Steinberg, Joffre Swait, Harry Timmermans, and James B. Wiley. 1994. "Experimental Analysis of Choice." *Marketing Letters* 5(4):351–368.

Cattin, Philippe, and Dick R. Wittink. 1982. "Commercial Use of Conjoint Analysis: A Survey." *Journal of Marketing* 46:44–53.

Chakraborty G., G. Gaeth, and M. Cunningham. 1993. "Understanding Consumers Preferences for Dental Service." *Journal of Health Care Marketing* 21:48–58.

Elrod, Terry, Jordan J. Louviere , and Krishnakumar S. Davey. 1992. "An Empirical Comparison of Ratings-Based and Choice-Based Conjoint Models." *Journal of Marketing Research* 30(August):368–377.

Gan, Christopher, and E. Jane Luzar. 1993. "A Conjoint Analysis of Waterfowl Hunting in Louisiana." *Journal of Agricultural and Applied Economics* 25:36–45.

Green, P.E., D.S. Tull, and G. Albaum. 1988. *Research for Marketing Decisions.* Englewood Cliffs, NJ: Prentice-Hall.

Huber, Joel. 1997. "What We Have Learned From 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles or Choice Experiments." Paper prepared for the Sawtooth Software Conference, August.

Huber, Joel and Noreen M. Klein. 1991. "Adapting Cutoffs to the Choice Environment: The Effects of Attribute Correlation and Reliability." *Journal of Consumer Research* 18(December):346–357.

Huber, Joel, Dick R. Wittink, John A. Fiedler , and Richard Miller. 1993. "The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice." *Journal of Marketing Research* 30:105–114.

Johnson, F. Reed, William H. Desvousges, Erin E. Fries, and Lisa L. Wood. 1995. *Conjoint Analysis of Individual and Aggregate Environmental Preferences.* TER Technical Working Paper T-9502. Durham, NC: Triangle Economic Research.

Johnson, F. Reed, and William H. Desvousges. 1997. Estimating Stated Preferences with Rated-Pair Data: Environmental, Health, and Employment Effects of Energy Programs. *Journal of Environmental Economics and Management* 34(1):79–99.

Johnson, F. Reed, D.J. MacNair, and E.E. Fries. 1997. *Ill-Considered Values: Contingent Valuation in Light of Evidence From Multiple-Response Stated-Preference Surveys.* TER Technical Working Paper No. T-9703. Durham, NC: Triangle Economic Research.

Johnson, F. Reed, Melissa C. Ruby, William H. Desvousges, and Jonathan R. King. 1998. *Using Stated Preferences and Health-State Classifications to Estimate the Value of Health Effects of Air Pollution: Final Report.* TER Technical Working Paper No. T-9807. Durham, NC: Triangle Economic Research.

Keeney, Ralph L., and Howard Raiffa. 1978. *Decisions with Multiple Objectives: Preferences and Tradeoffs.* New York: John Wiley and Sons.

Krupnick, Alan, and Maureen L. Cropper. 1992. "The Effects of Information on Health Risk Valuations." *Journal of Risk and Uncertainty* 5:29–48.

Louviere, J.J., and D.A. Hensher. 1982. "On the Design and Analysis of Simulated Choice or Allocation Experiments in Travel Choice Modeling." *Transportation Research Record* 890:11–17.

Louviere, Jordan J. 1988. "Conjoint Analysis Modelling of State Preferences." *Journal of Transport Economics and Policy* 22(1):93–119.

Lorenzen, Thomas J., and Virgil L. Anderson. 1993. *Design of Experiments: A No-Name Approach*. New York: Marcel Dekker, Inc.

Mackenzie, John. 1993. "A Comparison of Contingent Preference Models." *American Journal of Agricultural Economics* 75(3):593–603.

McFadden, D. 1981. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data*, C. Manski and D. McFadden, eds., pp.198–272. Cambridge, MA: The MIT Press.

McFadden, D., and K. Train. 1997. "Mixed MNL Models for Discrete Response." Working Paper. Berkeley: University of California Department of Economics.

McKelvey, R. and Zavoina, W. 1975. "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *Journal of Mathematical Sociology* 4:103–20.

Olsen, G. Douglas. 1992. "The Influence of Search Requirements and the Ability to Refuse Purchase on Judgements and Choices", Dissertation, Marketing Faculty of Business, University of Alberta.

Olsen, Douglas G., and Joffre Swait. 1997. "Nothing is Important." Working paper. Alberta, Canada: Faculty of Management, University of Calgary.

Opaluch, James J., Stephen K. Swallow, Thomas Weaver, Christopher W. Wessells, and Dennis Wichelns. 1993. "Evaluating Impacts from Noxious Facilities: Including Public Preferences in Current Siting Mechanisms." *Journal of Environmental Economics and Management* 24:41–59.

Payne, John W. 1976. "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis." *Organizational Behavior and Human Performance* 16:366–387.

Payne, John W. 1982. "Contingent Decision Behavior." *Psychological Bulletin* 92(2):382–402.

Phillips, Kathryn A., Rick K. Homan, Harold S. Luft, Patricia H. Hiatt, Kent R. Olson, Thomas E. Kearney, and Stuart E. Heard. 1997. "Willingness to Pay for Poison Control Centers." *Journal of Health Economics* 16:343–357.

Propper, Carol. 1995. "The Disutility of Time Spent on the United Kingdom's National Service Waiting Lists." *The Journal of Human Resources* 30:677–700.

Revelt, D., and K. Train. 1998. "Mixed Logit With Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics.* Forthcoming.

Roe, Brian, Kevin J. Boyle, and Mario F. Teisl. 1996. "Using Conjoint Analysis to Derive Estimates of Compensating Variation." *Journal of Environmental Economics and Management* 31(2):145–159.

Ryan, M. and J. Hughes. 1997. "Using Conjoint Analysis to Value Surgical Versus Medical Management of Miscarriage." *Health Economics* 6:261–273.

Ryan, Mandy, Emma McIntosh, and Phil Shackley. 1998. "Methodological Issues in the Application of Conjoint Analysis in Health Care." *Health Economics Letters* 2(1):15–21.

Torrance, G.W, M.H. Boyle, and S.P. Harwood. 1982. "Application of Multi-Attribute Utility Theory to Measure Social Preferences for Health States." *Operations Research* 30:1043–1069.

Tversky, Amos, Shmuel Sattath , and Paul Slovic. 1988. "Contingent Weighting in Judgment and Choice." *Psychological Review* 95(3):371–384.

Van der Pol, M., and J. Cairns. 1998. "Establishing Patient Preferences For Blood Transfusion Support: An Application of Conjoint Analysis." *Journal of Health Services Research and Policy* (forthcoming).

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber. 1991. "Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Tradeoffs for Chronic Bronchitis." *Journal of Environmental Economics and Management* 21:32–51.

Wittink, Dick R., and Philippe Cattin. 1989. "Commercial Use of Conjoint Analysis: An Update." *Journal of Marketing* 53:91

Discussion of Schulze, Chestnut, Mount, Weng and Kim paper

by Clark Nardinelli, US FDA Center for Food Safety and Applied Nutrition

The authors use the relationship between prices and safety features of automobiles to estimate the value of a statistical life for children and retired persons. The work is still at an early stage, so some of my comments will no doubt apply to things that will be changed later. My comments deal with (1) the theoretical model, (2) the variables included in the regressions, (3) the sample of single car households, and (4) the interpretation of the estimated willingness to pay for safety by the households in the sample.

(1) The model is plausible and sufficiently general. I would have approached the problem differently, but I had no trouble seeing how my preferred approach could fit within the model. The model derives an implicit value of statistical life for a family with non-paternalistic altruistic parents who determine health and safety expenditures on their children through Nash cooperative bargaining. The value of a statistical life derived from the model is the sum of the monetized value of the utility lost at death and the difference between wage income and consumption expenditures. The hedonic price for automobile safety is equal to the average value of a statistical life for the family. The model implies that the value of a statistical life can be derived from hedonic regressions on automobile safety features. If the value of a statistical life varies by age, then households with children should have a different willingness to pay for automobile safety than households without children.

(2) The model is straightforward; the hard part is designing hedonic regressions that will tease out the safety premium from the other features affecting price. The authors admit that the regressions reported in the paper are their first try. In later versions, they should consider including variables for geographic differences. The estimated premium on safety could well be distorted if regional and city differences are not accounted for. For example, based on my casual observations, people in New Orleans drive differently than people in Sacramento. The quality of roads and the weather will also affect the demand for automobile safety and other features. In addition to gender, I recommend including some additional demographic variables, such as ethnicity. Finally, designing hedonic regressions does not follow strict by-the-book rules, so I urge the authors to play around with the regressions as much as possible. Finding the right combination of variables and form that will generate defensible estimates of willingness to pay for automobile safety by ages of household members will not be easy.

(3) The data present another difficulty. The results presented here come from 4,000 one-car households. This sample – as the authors point out – is not representative, so we should be careful about concluding much about the results. The finished project will be based on over 40,000 households and will not be restricted to one-car households. I look forward to seeing the results from the full sample.

(4) The tentative results of the exercise are that willingness to pay for safety does not vary much across households by age. Because this workshop is more concerned with

children than with retired people, I restrict my comments to the apparent similarity between households with children and households without children.

The results presented in the paper fail to demonstrate that the value of a statistical life is approximately the same for children and adults, all else the same. I think that a hidden income effect is at work in the regressions. In his presentation -- but not in the written version I received -- William Schulze showed some additional results that strongly indicated that an income effect was at work. In the paper's model, the value of a statistical life is inversely related to consumption expenditures. Households with children typically have higher consumption expenditures than other households. The presence of children in a household therefore lowers discretionary income, which reduces the ability to purchase safety. The same estimated values of a statistical life for households with children and households without children might well indicate a greater concern for children's safety. The concern, however, may be offset by the reduced discretionary income.

The similar average values of a statistical life for the households in the sample may also reflect other effects not picked up by the regressions. One possibility is that one-car households are relatively homogenous. Another possibility is that differences in discount rates may be affecting the results. I have not thought much about these two influences, but I want to mention them

I like the approach taken in this paper. As Professor Schulze said in his introductory remarks, the paper represents the first part of a large project aimed at using market prices for safety to generate estimates of the value of a statistical life by age. I look forward to later versions of this paper and to other products of the broader research agenda.

As a government economist, part of my job is to defend the value of a statistical life to risk managers. Risk managers want a good, clear story. A good defense of the value chosen therefore requires being able to tell a plausible story justifying the choice. I have found that the most persuasive stories use market prices and behavior; risk managers can understand revealed preferences based on market results. Research of the kind undertaken by Professor Schulze and his collaborators can make my job a lot easier.

Discussion of Johnson, Ruby and Desvouges paper

by Fred Kuchler, USDA Economic Research Service -- Summarization

Mr. Kuchler focused his remarks upon a comparison of conjoint analysis with contingent valuation. Both involve hypothetical transactions, rather than cash or arms-length transactions. Both essentially ask respondents what they think.

In 1993 NOAA convened their expert panel and came up with six possible problems with contingent valuation: (1) inconsistency with rational choice, (2) it gives rise to estimates that are implausibly large. (3) respondents do not demonstrate awareness of their budget constraint, (4) the issues presented to respondents are too complex, (5) the extent of the markets for the hypothesized goods are unknown, (6) part of respondents' bids may be due to a "warm glow" effect.

How does conjoint analysis fare on these points vis-à-vis contingent valuation? With respect to problem (3), conjoint analysis fares well, since out-of-pocket costs (in the Johnson paper, for example) are an important aspect of the scenarios presented to respondents. With respect to problem (4), the issues do not appear to be too complex, as the scenarios seemed to be within their ranges of experience. With respect to problem (6), the private goods presented to respondents (at least in the Johnson paper) give rise to no warm glow effects at all. With respect to problem (1), the models hypothesized by Johnson seemed to be based upon an indirect utility function that included many health attributes that interacted with out-of-pocket costs, which is consistent with rational choice, and furthermore in Johnson's model the marginal utility of money was diminishing.

With respect to problem (2), Mr. Kuchler calculated the value of a year of life by taking the standard value of a statistical life and annuitizing it over 36.5 years, using a discount rate of 3%. He found that value of a year of life having a value of $225,000, or a value of a day of between $600-700. The largest value for a day of illness obtained in the present paper was $535 per day, but for a person with a very bad illness. These results thus make sense, since the estimated willingness-to-pay to avoid illness (even severe illness) is less than willingness-to-pay to avoid death.

With respect to problem (5), the present paper focuses upon the willingness to pay by adults, so does not measure the willingness to pay for children, although an instrument might be designed to accomplish that.

The problem is that the health outcomes hypothesized here are certain, not probabilistic, so the good does not match the good that regulatory agencies provide. So the pertinent question is how do we translate the value into something probabilistic, so that it is usable by EPA?

Policy Discussion for Session II

by Nick Bouwes, US EPA Office of Pollution Prevention and Toxics -- Summarization

Mr. Bouwes expressed appreciation that the work being done for this workshop accounts for the benefits of environmental protection, for which estimates are still lacking. Many of the benefits estimates used by the EPA Office of Pollution Prevention and Toxics are lower-bound estimates. For example, valuing the benefits of lead protection regulation typically is accomplished by calculating the foregone income from a decrement in IQ attributable to lead poisoning, as well as medical costs and remedial education costs. Also, the office is developing a Cost of Illness handbook, containing estimates of the direct costs of 25 diseases.

Mortality provides a conceptually straightforward health endpoint: death. However, morbidity studies require the consideration of separate valuations of different health endpoints. This challenge can only be met by alternative approaches, such as contingent valuation, stated preference and cost of illness. In evaluating these different approaches, it is necessary to ask if the approach is immediately applicable, and if not, how much would it cost for EPA to undertake a valuation of a regulation? Although the cost of illness approach provides lower-bound estimates, regulations often have alternatives, so more precise estimates are necessary.

With respect to the paper by Schulze et al., Mr. Bouwes queried whether it was in the wrong session. If there are extensions to other health endpoints, they should have been presented. For example, there could have been an application to the measurement of a value of a statistical life with a latency period.

Mr. Bouwes wondered if baby seats were included in the Schulze study. Mr. Schulze replied that there was no data on the use or location of baby seats. Mr. Bouwes asked if the propensity of sport utility vehicles to roll over was considered in the safety considerations, to which Mr. Schulze replied that sport utility vehicles are still safer.

On the paper by Johnson et al., Mr. Bouwes wondered why Mr. Johnson used a joint estimator rather than the sole estimator for valuing a preferable health scenario.

With respect to both papers, Mr. Bouwes inquired as to the cost of performing these studies. Mr. Bouwes raised the issue of trading off the soundness of the theoretical foundation of the studies against the resources needed to carry out these studies.

Question and Answer Period for Session II

Johnson, Ruby, and Desvouges paper:

Bill Harbaugh, University of Oregon, asked F. Reed Johnson, Triangle Economic Research, if it is possible to check to see if the respondents in his study were providing rational responses that were consistent with economic theory. Mr. Johnson replied that it was possible to check for transitivity and monotonicity. Violations of standard microeconomic assumptions do occur, but it is probably largely due to the demanding nature of the survey, since the error is correlated with things like respondents' education, age, and the amount of time respondent had spent on the survey.

In response to the ex ante problem posed by Fred Kuchler in his discussion of Mr. Johnson's paper, i.e., that of obtaining the value of a good that is certain and applying it to a good that is probabilistic, Mr Johnson responded that there are two ways to treat this problem. One is the approach adopted by Krupnick, et al., to treat it as an ex ante problem, and simply get people to value it directly. The Triangle Economic Research approach is to obtain an ex post value, ascertain the probability of the person suffering this malady, and convert the ex post value to an ex ante one. If one knows something about a respondent's risk preferences, then in principle it should be possible to convert an ex post value to an ex ante value. Alan Krupnick, Resources for the Future, added that the existing estimates of symptom days are all ex post, and most are of low quality, so at the least, Mr. Johnson's paper is a vast improvement over existing estimates.

Lauraine Chestnut, Stratus Consulting, queried whether this was really ex post, or simply ex ante with a time lag. Mr. Johnson agreed that this was ex post only in the sense that they had eliminated the uncertainty. This study also oversampled people that actually have had the maladies that are hypothesized, and collected data on current health, and found that current health is a very strong determinant of willingness to pay. Similarly, in another Triangle Economic Research study on smokers, they found that whether the respondent was a smoker was a very strong determinant of willingness to pay.

Clay Ogg, US EPA Office of Policy, pointed out that the income variable used in Mr. Johnson's study was likely to be correlated with having kids, and therefore with willingness to pay.

An unidentified speaker asked Mr. Johnson about the plausibility of the results and wondered if the study took place in Los Angeles instead of Toronto, whether there would be many more episodes of air pollution violations, and whether this would alter the results. Mr. Johnson responded that air pollution was not mentioned, because they wanted to decouple the study from the regulatory context. Mr. Johnson added that he would be reluctant to transfer the results to non-Canadians.

William Schulze, Cornell University, cited a study of paraplegics, and pointed out that they are not necessarily any less happy than the rest of us, so even though our willingness to pay to avoid the risk of becoming paraplegic is high, the ability of people to adapt is often overlooked. Mr. Johnson agreed that this was an interesting and an important phenomenon. Analogously, Mr. Johnson cited his own study, in which he found that older people have a statistically significantly higher level of perceived health than younger ones.

Ellen Post, Abt Associates, asked why both Mr. Krupnick and Mr. Johnson made sure to obtain valuations without reference to any air pollution issues. Mr. Johnson responded that this is a split in philosophy; Richard Carson would say that the context is part of the value derived by the respondent -- if warm glow makes them feel good, why not include it in their valuation? Mr. Johnson stated his disagreement with this philosophy, and noted that is hard enough to get respondents to focus in on the attributes of the hypothesized good, without having them try to sort out their feelings. Mr. Johnson acknowledged that this is, in Mr. Krupnick's words, a "deep philosophical question." Ms. Post followed up and pointed out that when people express a higher willingness to pay for air pollution reduction, that they are not just registering their outrage, but also fear, which is a legitimate aspect of their valuation. Mr. Johnson responded that he would like to allow people's broader preferences to count, but did not want to be in the position of judging which fears were justifiably part of their preferences, and which were not, e.g., psychosomatic headaches. Mr. Schulze also responded to Ms. Post's question by pointing out that staying with private good scenarios keeps the valuations free of warm glow problems.

Thomas Crocker, University of Wyoming, pointed out that while the expert knows more about the subject, individuals know more about their preferences, and while there is an argument that perhaps more weight should be accorded to experts, the importance of understanding individual preferences should not be underestimated.

Nick Bouwes, US EPA Office of Pollution Prevention and Toxics, stated that EPA is currently working with estimates of the cost of illness, and wondered if it would be worthwhile comparing estimates. EPA's cost of illness estimates should be a lower bound for the willingness to pay to avoid these serious illnesses, however.

Bryan Hubble, US EPA Office of Air Quality Planning and Standards, inquired as to whether scenarios included multiple symptoms. Mr. Johnson responded that they did not.

Schulze, Chestnut, Mount, Weng and Kim paper:

Richard Belzer, Washington University, inquired as to the multicollinearity problems inherent in using vehicle weight as a proxy variable for safety. Mr. Schulze responded that some more econometric adjustments would be made. Mr. Belzer pointed out that another attribute that could be captured by the weight variable is comfort. Mr. Schulze responded that this data can not be obtained because data on vehicle model and year are not available. This data might be obtained by primary data collection.

Mr. Krupnick asked about actual injury rates, which are probably correlated with fatality rates. Mr. Schulze acknowledged that this was correct, and that his studies results should be considered as including both injuries and fatalities. Mr. Krupnick then pointed out that the result should not be called a "value of a statistical life" because of the joint production problem of some vehicles reducing both injuries and fatalities.

Steve Crutchfield, USDA Economic Research Service, suggested that some other overall index of operation costs should be used, rather than gasoline mileage and suggested that insurance costs may be a good index, except that Virginia captures much of its revenue through a property tax. So Mr. Crutchfield suggested that location-specific cost variables are called for. Mr. Schulze agreed, but pointed out that these cost variables also capture some of the benefits of the vehicle, as gas mileage is likely to be correlated with safety.

Mahesh Podar, EPA Office of Water, asked if crash-worthiness test results were considered, and Mr. Schulze replied that they were, but dismissed because these tests are conducted using similar-sized vehicles, so cannot measure the relative safety of different vehicles.

Kenneth Acks, Costs and Benefits Newsletter, reiterated the problem of collinearity between the comfort and safety variables, to which Mr. Schulze acknowledged that the good needed to be treated as a joint products problem.

Mr. Ogg asked if there was a problem with income being correlated with many other variables that might determine the value of the statistical life in question, to which Mr. Schulze replied that in the hedonic formulation, the income effect can be separated out.

Mr. Bouwes pointed out that even with the same income, one wouldn't necessarily expect the same expenditures with children, since children have many expenses other than just vehicle safety. Mr. Weifeng Weng, Cornell University, co-researcher with Mr. Schulze, responded that income was adjusted for family size. Mr. Crocker complemented Mr. Schulze's study, remarking that this is one of the few studies that have endogenously modeled risk. Robin Jenkins, US EPA Office of Economy and Environment, asked if the incomes of the respondent families matched the incomes of the U.S. population generally, especially given that 30% of the U.S. population lives in poverty. Mr. Schulze responded that even the poorest respondent families in his sample were still spending a substantial amount of money on child safety. This was one of the most significant results of the study -- that the amount of money being spent on child safety was relatively income-invariant.

Brian Caulkins, Washington State Department of Ecology, asked if one could impute the value of an adult life for one who was working inside the home, to which Mr. Schulze replied that a use-weighted mean was used in the analysis, so that the value of a housewife or househusband was accurately accounted for.

Mr. Crocker asked if individuals might have different views of the efficacy of safety technology, and wondered if this might affect their willingness to pay for automobile safety features. A lower confidence in safety features may cause the respondent to reveal a value that is lower than the individual's actual willingness to pay for risk reduction.

Mr.Belzer posed a problem regarding vehicle choice in multiple-car households and driving teenagers. Most parents will choose the heaviest, safest vehicles for their teenagers. In this sense, it is the parents that are valuing safety on behalf of their teenage children. Yet the fact that these children have risky driving habits must affect vehicle choice and willingness to pay for safety features.

Jim Neuman, Industrial Economics, Inc., asked if Mr. Schulze had considered asking automobile manufacturers for marketing studies on preferences for automobile safety features, to which Mr. Schulze replied that he suspected that lawyers would probably not permit the release of such data. Mr. Neuman continued by asking whether conjoint analysis would be more probative. Mr. Schulze said he believed that this was still essentially a contingent valuation problem.

Peter Negelhout, US EPA Office of Economy and Environment, pointed out that while accidents are random, drivers can take defensive measures to minimize risk, such as choosing where and when to drive. Mr Schulze agreed, and emphasized that automobile safety is only one of the things people do to reduce risk.

Elyce Biddle, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, followed up on Mr. Neuman's point by encouraging Mr. Schulze to seek out marketing studies from automobile manufacturers, since they may not be as reluctant to release such information as Mr. Schulze might think.