Willingness-to-Pay Approach

An entirely different approach to assigning value to risk reduction is to estimate what risk reduction is worth to individuals whose health might benefit. With this approach, analysts estimate consumers' willingness to pay (WTP) for reductions in health risk or improvements in health. Clearly, these values may vary among individuals because preferences are idiosyncratic. In addition, such values usually differ from COI estimates, because, unlike the wealth-maximizing society in the COI approach, individuals may attach value to goods that are not marketed. These goods include intangibles such as pain and suffering.

The WTP approach reflects the observation that individual preferences are unique and individual demands for risk reduction vary. However, because health and safety are normal goods, a substantial portion of the variance in WTP estimates will be explained by income differences rather than preferences. So, just as in COI analysis, income and circumstance could play a role in determining the size of WTP estimates.

This chapter examines the theoretical basis for using the WTP approach for social welfare analysis and presents some of the most important criticisms of the approach. This discussion is followed by a review of the methods used to empirically estimate WTP values for health and life.

WTP is an *Ex Ante* Choice

When WTP is used in the evaluation of health and safety programs, it measures what individuals would be willing and able to pay for a reduction in the probability of encountering a hazard that might compromise their health. The WTP approach is, therefore, concerned with measuring *ex ante* valuations; valuations at the moment choices are made.

The WTP approach for estimating benefits of public health programs rests on the observation that individuals can and do make tradeoffs between health and other consumption goods and services. Proponents of the approach contend that even though individuals tend to place an infinite value on their own lives (and the lives of those they hold dear), they do not feel similarly about small changes in risk. Individuals routinely and voluntarily accept many small risks in exchange for finite benefits. For example, driving a little faster than surrounding traffic may raise the risk of injury but usually results in reaching a destination sooner. Or, a person might enjoy attending a popular movie at a crowded theater, recognizing that the activity raises the risk of contracting a contagious disease. An individual might suffer actual harm as result of the decision to speed or to sit in a crowded theater and might later regret the decision. But WTP does not measure realized damages or capture the *ex post* valuation of an individual's changed health status. COI would be more appropriate for such estimates.

WTP is most appropriate for evaluating health- or life-threatening hazards that strike with some degree of randomness, so that no one could predict exactly who will actually suffer from the hazard or benefit from the prevention. Many of the hazards addressed through publicly financed health and safety programs fit this description. In these cases, health and safety programs are not targeted at specific individuals, but at reducing hazards to which many may be exposed, reducing probabilities of risk or death or illness for many. It is hard to imagine individuals voluntarily engaging in activities involving the immediate and certain death of a participant, and WTP is not intended to be a price reflective of such exchanges.

Schelling (1966) was the first to propose WTP for valuing changes in health status. He argued that applying WTP to health and life was simply a logical extension of standard welfare economic principles—principles based on consumer sovereignty:

> The gravity of decisions about life-saving can be dispelled by letting the consumer (taxpayer, lobbyist, questionnaire respondent) express himself on the comparatively unexciting subject of small increments in small risks, acting as though he has preferences even if in fact he does not. People do it for life insurance: they could do it for life-saving. (p. 161)

As expressed by Schelling, the foundation of the WTP approach is the belief that individuals are the

best judge of their own well-being, and even in matters involving life and death, individual preferences should be held sovereign.[9]

WTP and Welfare

The usefulness of WTP estimates for cost-benefit analysis depends on the validity of these estimates as welfare measures. There is little question that WTP measures provide the best estimate of *individual* welfare available to economists. In the neoclassical economics tradition, the guiding principle in determining consumer welfare is to measure a consumer's "willingness to pay." Conceptually, these amounts are the values consumers attribute to goods they purchase, or conversely, the cost of forgone consumption opportunities. They are measured as consumer surplus derived either from a Marshallian demand curve (treating the quantity demanded as a function of prices and income, while letting utility vary) or from a Hicksian demand curve (treating quantity demanded as a function of prices and a utility level, where income adjusts to maintain the utility level). It is generally recognized that the Hicksian measures are more accurate measures of consumer welfare.

When applied to changes in mortality or morbidity risk, WTP measures the change in income, coupled with the change in the risk of mortality or morbidity, that leaves the consumer's utility unchanged. The WTP approach for calculating individual welfare changes due to changes in health status strives to estimate the theoretically correct Hicksian measures. These measures are therefore the best individual welfare measures available to economists.

Despite its usefulness as a gauge of *individual* welfare, the WTP approach is clearly less successful as a measure of social welfare. One reason for this shortcoming is that, with the WTP approach, individual valuations of life and health are aggregated to arrive at society's valuation even though such an aggregation is not usually a reliable indicator of social welfare for cost-benefit analysis. It is justifiable to sum individual utilities (WTP) only if the marginal utility of income is equal across income groups, i.e., if an extra dollar of income is equally valued by a millionaire and by someone with an income hovering above the poverty line. Only if an additional dollar is equally valuable to all groups, given the prevailing income distribution, can interpersonal comparisons of utility be made. Only in this case can individual well-being measures be aggregated to provide a basis for comparing costs and benefits across groups. Though there could be cases where the marginal utility of income was equal across income groups, it is unlikely that this condition could be met in cases of even mildly unequal income distributions. In cases where the marginal utility of income is not equal, interpersonal comparisons of utility cannot be made and money loses its value as a measure of welfare. Money is transformed into a "rubber ruler" (Friedman, 1996).

To avoid the whole issue of making interpersonal comparisons of welfare and placing values on gains to one group versus costs to another, economists, starting with Pigou's treatise on welfare economics (1952), have distinguished between efficiency and equity in welfare decisions. Welfare efficiency is concerned with maximizing the sum of individuals' welfare (whether measured as some function of net national product, consumption, or intangibles) while welfare equity is concerned with the distribution of welfare. The Kaldor-Hicks compensation principle is the decision criterion used for strict individual welfare maximization. In this role, the Kaldor-Hicks criterion has been dubbed the "fundamental principle" of cost-benefit analysis (Stokey and Zeckhauser, 1978; Gramlich, 1990).

The Kaldor-Hicks criterion states that a proposed policy change is desirable on social welfare grounds if everyone's welfare can potentially improve (Kaldor, 1939 and Hicks, 1940). The Kaldor-Hicks criterion means that a program may be desirable even if it makes some worse off and others better off. That is, a program where some pay yet receive no benefits while others receive benefits without paying could be acceptable under the Kaldor-Hicks criterion. If the

---

[9] Robinson (1986, p. 139) argues that the fundamental concepts represented by the WTP approach are inexorably linked to neoclassical economic philosophy: "Any conceptual strength possessed by the willingness-to-pay approach stems solely from its compatibility with the subjectivist orientation of the welfare economics of the postwar period. Analysis of the philosophical origins of the school of economics reveals that the path it took is not the only one possible, and that for some purposes others may be better."

value of the benefits generated by the program exceeded the payments, it would be possible for gainers to compensate losers. As long as there are positive net gains to society as a whole, the Kaldor-Hicks compensation principle is met. The Kaldor-Hicks criterion not only avoids the difficult question of how to compare costs and benefits accruing to different segments of the population, it also effectively avoids considering the distribution of costs and benefits at all. As succinctly put by Gramlich (1990), "Who these gainers and losers are, and how much they gain or lose are questions that simply do not matter under the Kaldor-Hicks standard" (p. 115).

The efficiency-first, equity-second approach of the Kaldor-Hicks criterion is defended on two fronts. The first defense is that any attempt to incorporate equity considerations in the welfare maximization equations, say through the introduction of weights, will result in inefficiencies. It is argued that any equity-enhancing redistribution should be achieved through lump-sum transfers after welfare maximization has taken place (Harberger, 1978). The second defense of the efficiency-first approach is that the role of economists should be restricted to enhancing efficiency and that equity considerations are best left to the political sphere (Kaldor, 1939). Others have argued that equity and efficiency must be attacked simultaneously and that lump sum transfers are mythical beasts (Layard and Glaister, 1994).

Despite the ongoing debate concerning equity and efficiency in determining social welfare from individual welfare estimates, the WTP approach vigorously applies the standard tools of neo-classical welfare economics to issues concerning health and life. At the theoretical level, the WTP approach to valuing human life is a faithful application of the principles of standard applied welfare economics: it builds up from individual valuations, it does not make interpersonal comparisons of utility, and it adopts the Kaldor-Hicks compensation criterion as its *modus operandi*.

Is Efficiency Sufficient for Health Policy?

As illustrated above, the WTP approach is a consistent application of modern applied welfare economics to policy with health ramifications. Proponents of the approach argue that if economic valuation principles are fundamentally sound, they must be equally applicable to every commodity, including health and life.

Criticism of the WTP approach usually centers on the assertion that health and life are not like other commodities and that there is no reason to suppose that standard economic techniques are adequate for the task of valuing life and limb. Broome (1978) questioned the validity of employing standard cost-benefit techniques to matters of life and death. The debate between Broome and his detractors is mirrored in the ongoing debate in the literature concerning the validity of the WTP approach for valuing change in health status.

Broome's first criticism of valuing life based on individual preferences concerns the compensation criterion embedded in most cost-benefit analyses. Broome argued that even a compensation scheme designed to fully compensate those who would otherwise be harmed by a public decision, would be inoperable with respect to life and death decisions. He noted "no finite amount of money could compensate a person for the loss of his life, simply because money is no good to him when he is dead (p. 92)." Broome also rejected the device adopted by analysts to circumvent this problem—the practice of "veiling" the identity of the victims in statistics and probabilities. Broome argued that ignorance of the identities of the victims does not mitigate the fact that real people with names and faces will actually die and that there is no ethical reason for valuing the life of an identified stranger more or less than that of an unidentified stranger.

Furthermore, Broome argued that when people make trade-offs involving risks to life and limb, they are ignorant of the actual outcome and are therefore not accurate judges of their own best interests:

> Consider any project in which an unknown person will die. Because whoever it is does not know it will be him, because of his ignorance, he is prepared to accept a ridiculously low compensation for letting the project go forward. The government does not know who will be killed either, but it knows it will be someone, and it knows that, whoever it is, no finite amount of compensation would be adequate for him. The cost of the project must therefore be infinite, and it is only the ignorance of the per-

> son destined to die that prevents his demanding an infinite compensation. It may be true that sometimes we are forced to make decisions based on imperfect knowledge if nothing better can be done. But this is one case where the problems of imperfect knowledge can easily be eliminated. If there is to be a death, we know at once that the cost, defined as the compensation required for the loss, is infinite. Any other conclusion is a deliberate and unfair use of people's ignorance. (p. 95)

Broome's critique was amply counter-critiqued (Buchanan and Faith (1979), Jones-Lee (1979), Williams (1979), and Mishan (1981)). The central element of the critique forwarded by Buchanan and Faith was that Broome included an incorrect characterization of "costs" in his calculations. In particular they argued that Broome confused costs that influence choices (costs individuals believe, *ex ante*, they will incur from a choice) with damages (cost individuals actually incur, *ex post*, from a chosen action). They argued that Broome erroneously equated costs with damages and as a result, incorrectly compared the infinite cost of loss of life with finite benefits of expenditures on general goods and services. Buchanan and Faith maintain that the costs that influence an individual's decisions are rejected alternatives:

> To say that 'costs' are infinite for the person who loses his life in the draw of a lottery in which he rationally chooses to participate is to say nothing at all about the *value* that such an individual placed on life in the moment at which the choice was made. (p. 246)

Buchanan and Faith contend that at the instant at which individuals make risky choices, the costs they perceive are those goods and services they must sacrifice to achieve small reductions in risk. Costs are therefore of finite value. For example, the cost to the driver who chooses a speed greater than the surrounding traffic is a small increase in the likelihood of injury in an accident. The cost that influences his choice is not death, but a change in risks incurred. The cost of choosing a speed consistent with surrounding traffic is arriving later than he would by driving faster. Because the driver is willing and able to trade one alternative for another, there is no question that the alternatives the driver rejects are of finite value to him.

The rest of the critiques of Broome's paper amount to reaffirmations of the central tenet of applied welfare economics, that each person knows his or her own interest best and that public decisions should be based on these private valuations. Mishan (1981) described Broome's rejoinder (1979) as an "attack on my proposal arising not from a belief that it is inconsistent with the standard procedure but that, in some sense, it is illegitimate to extend to life and limb the standard procedure that is appropriate for other goods and bads" (p. 136). Mishan contended that to be consistent, economics must apply standard valuation procedures to all goods and bads, including life and limb and that "once he [either 'the economist' or Broome] accepts that the distinctive characteristic of economic evaluation is recourse to the individual's own valuations of the change in question, he has no choice but to go along with their responses, 'paradoxical,' perverse, or otherwise" (p. 137).

Supporters of the WTP approach to valuing life and health contend that it is a logical and consistent application of the primary tenets of standard applied welfare economics and that unease with the results simply reflects an underlying unease with the foibles of human nature. Fuchs and Zeckhauser (1987) suggest that failure to apply standard economic tools to life and health is a result of myth maintenance as opposed to economic efficiency and cost containment. They contend that myths regarding the valuation of life and health persist in our society and give us comfort but that as a result of our myths, "many mechanisms of cost containment must work in the shadows" (p. 267).

As proof of their unflagging support of the right of individuals to determine the value of life-saving or health-enhancing policy, proponents of the WTP approach have often argued for policy prescriptions that appear starkly unfair in a life and death context. For example, Viscusi (1991) supports Schelling's suggestion that the fact that the Titanic carried only enough lifeboats for first-class passengers could be a logical and valid conclusion of a properly executed WTP study (though he concludes that such lifeboat contracting could not hold up in practice because once the ship started to sink it would be impossible to deny access to the lifeboats).

The reasoning behind this conclusion hinges on the conviction that preferences are adequately revealed through consumer choices. Just as spending five

times more on bubble gum reveals five times the preference for bubble gum, spending five times more on health care reveals five times the preference for the health. Likewise, the fact that lower income groups tend not to drive new cars with up-to-the minute safety features indicates that lower income groups place a lower value on health and safety than wealthier new car drivers. Schelling (1966) clarifies this view:

> A special matter of policy is bound to arise here. If a government is to initiate programs that may save the lives of the poor or the rich, is it worth more to save the rich than to save the poor? The answer is evidently yes if the question means is it worth more to the rich to reduce the risk to their own lives than it is to the poor to reduce the risk to their own lives. Just as the rich will pay more to avoid wasting an hour in traffic or five hours on a train, it is worth more to them to reduce the risk of their own death or the death of somebody they care about. It is worth more because they are richer than the poor. (p. 157)

The reasoning implied by statements like those above is incomplete. Interpersonal comparisons of utility of this type are invalid unless the marginal utility of income is equal between groups.[10] In actuality, there is no reason to assume that an extra dollar was of equal value to the steerage and first-class passengers on the Titanic and there is very little reason to assume that the marginal utility of income is equal for a family with an income near the poverty line and one that purchases a new car every year. Like all consumption choices, the purchasing decisions of both the steerage passengers and the drivers of rusty, old cars are constrained by income as well as by preferences. If the marginal utility of income is greater for the old car-driver than the wealthier new-car owner, then the used child seat in the back of the old car could entail a larger sacrifice and reveal a stronger preference for safety than that revealed by the new car-owner's more expensive purchase of air bags, anti-lock brakes and impact resistant side bars. As stressed by Deaton and Muellbauer (1980), consumption decisions are conditioned by preference *and* possibility.

Though the theoretical rationale for using unweighted individual welfare valuations in cost-benefit analysis is not based on the assumption of equal marginal utility of income across socio-economic groups, the results are similarly influenced by the current distribution of preference and possibility. The theoretical linchpin of WTP studies is the Kaldor-Hicks principle. The efficiency-first, equity-second criterion embedded in the Kaldor-Hicks principle results in policy prescriptions that favor wealthier segments of society (at least initially). The efficiency-first criterion requires that the policymaker maximize the unweighted aggregation of individual valuations. Only after maximum efficiency is achieved does the policymaker address equity concerns for real social welfare maximization.

An efficiency-first criterion would indicate that safety policy be directed to those sectors of the society that place the highest value on safety. In cases where individual WTP amounts are influenced by income as well as by preference, higher income groups would most often exhibit greater safety purchases and as a result would be the beneficiaries of safety policy. A cost-benefit analysis incorporating these results would indicate that the government should target safety improvements to upper income groups with equity concerns addressed through redistributive policy after efficiency maximization.

The usual defense of the efficiency-first, equity-second approach of the Kaldor-Hicks criterion rings a bit hollow when applied to issues of life and limb. It might be reasonable to argue that equity-enhancing redistribution should be achieved through lump-sum transfers after welfare maximization has taken place for those cases where a redistribution of income would be sufficient to leave everyone as well off as before the policy change. However, in cases where policy results in a particular distribution of premature death, disability, or ill-health, it might be difficult to adequately compensate the "losers" with any amount of lump-sum transfers. Layard and Walters (1994) argue that "there is no ethical justification for the Hicks-Kaldor criterion; where compensation will not be paid there seems no alternative to interpersonal comparisons of the value of each person's gains and

[10] In making statements of this sort, Schelling and Viscusi are most likely victims of semantics. They probably did not intend to compare "utility" when discussing comparisons of "worth."

losses" (p. 6). For policy influencing the distribution of life and death, potential compensation will always remain just potential. The Kaldor-Hicks principle could be a valid operating criterion for most goods and services, but the fact that those needing to be compensated might be dead or dying seems to invalidate the logic of the criterion for health policy: to echo Broome, it is strictly impossible to redistribute between those in this world and those in the next.

Empirical Results

In the theoretical discussion presented above, it was implicitly assumed that WTP amounts can be measured. The WTP theory was critiqued because of the efficiency-first, equity-second criterion that is adopted with an unweighted aggregation of WTP amounts. However, because WTP amounts are subjective, the task of deriving them is very difficult. The very subjectiveness that makes them so theoretically appealing is also what makes them empirically challenging. So, whether unweighted or weighted, aggregated or individual, WTP amounts are extremely difficult to estimate over a whole population.

This point was stressed by Buchanan and Faith (1979) in their observation that the value an individual places on a commodity is best measured by opportunity costs, defined as what an individual believes he is giving up by choosing one way rather than another. These individually assessed opportunity costs exist only at the moment a decision is made, and only in the mind of the choicemaker. Opportunity costs need not bear any relation to objectively measurable costs, like realized damages. Because opportunity costs are inherently subjective and unobservable, Buchanan and Faith argue that external observers, including analysts conducting a cost-benefit study, cannot discern the value an individual places on life (or, more precisely, on changes in life-threatening risk). Thus, even if benefits of a program were large enough to compensate all those made worse off, it would be impossible to do so, because appropriate compensation levels would elude measurement.[11]

There is no way to overcome the Buchanan and Faith argument; no one can know exactly what is in the mind of another. But, public sector decisionmakers have to choose which programs to fund and which activities to regulate. Health and safety policies will be made even if decisionmakers have only incomplete knowledge of costs and benefits. Decisions will be made even if there is nothing to guide program selection toward those that are inexpensive and offer large benefits. The real question Buchanan and Faith raise for health and safety policy is whether economists can estimate the value of health benefits well enough so that the results of cost-benefits analyses serve as good guides toward efficient program selection. In practice, economists routinely assign prices to non-marketed goods through a variety of methodologies, including the contingent valuation method, the hedonic pricing method, and the travel cost method. Some of these price estimates are quite speculative while others are more certain.

One of the most straightforward methods of assigning value involves deriving a price from associated marketed commodities (having observable prices) and a set of behavioral assumptions. That is, there may be marketed commodities for which demand characteristics are arguably similar to the non-marketed commodity. For example, consider assigning a value to irrigation water in the Southwest. In some States, water or water rights may not be traded separately from land. Yet economists can confidently assign value to new irrigation water and thereby estimate benefits of a construction project that would provide irrigation water. A simple method for assigning a value to water would be to calculate the price differential for land sold with and without irrigation water, based on recent sales prices. That price differential should represent the present discounted value of profits earned through the extra productivity of irrigation water and, equivalently, the WTP for water.

The above example uses an observable characteristic of real estate sales with the assumption of profit maximization to assign a value. So far, valuing risks to life and health has proved to be more difficult than valuing other non-marketed commodities. Finding associated marketed commodities and behavioral assumptions that allow analysts to derive a price for risk reduction is not a trivial task. As a consequence, estimating the value of risk reduction requires more heroic assumptions and leads to less robust results

[11] The Buchanan and Faith argument that values are subjective is not specific to life and health. The same argument could be made for any commodity. Thus, one could argue that economists cannot assign value to any non-marketed commodity.

than estimating the value of other non-marketed commodities.

In the health economics literature, analysts have used four primary methods for empirical estimation of willingness-to-pay measures.

- The compensating-wage method
- The contingent valuation survey method
- The household production function method
- The hedonic price method

Each method provides a means of deriving Hicksian willingness-to-pay estimates for individuals making tradeoffs between risks to life and health and other consumption goods and services. Each of these methods is examined below

Compensating Wage Differentials

The dominant empirical approach to assessing WTP risk tradeoffs uses labor-market data on wage differentials for jobs with health risks. This approach assumes that workplace risks are well understood by workers and that the additional wages workers receive when they undertake more risky occupations reflect risk choices. The underpinnings of the compensating wage approach have been traced to Adam Smith and his observation that risky or otherwise unpleasant jobs will command a compensating wage differential (pp. 99-100). The compensating differential approach relies on the assumption that workers will accept exposure to some level of job-related risk in return for some level of compensation. For example, suppose jobs A and B are identical except that, on average, there is one more job-related death per year for every 10,000 workers in job A than in job B, and workers in job A earn $500 more per year than those in job B. The implied value of a statistical life revealed by the willingness of workers in job B to forgo an extra $500 per year for a 1-in-10,000 lower annual risk is calculated at $5 million (example from Fisher et al., 1989).[12]

The basic approach in the compensating-wage literature is to estimate a hedonic wage equation where wages are specified as a function of personal characteristics of the worker and characteristics of the job. Individual worker characteristics can include wealth, age, sex, education, experience, and health. These variables are particularly important as they affect the firm's demand for the individual worker, the workers' preferences, and other labor opportunities available to the worker.[13] Job characteristics that influence costs of providing particular safety levels can include the fatality risk of the job, the nonfatal risk of the job, worker compensation benefits that are payable in case of injury on the job, and annuity benefits in the event of a fatal accident. Compensating wage differential models are consistent with WTP theory in that they recognize that individuals have unique preferences over risky alternatives and that their opportunities to reduce risk vary, often depending on the marketability of their labor skills. Compensating wage differential models postulate that a large share of the differences in risk preferences are systematic, depending on objective and measurable individual characteristics.

Accurate and consistent measurement of the risk variables and worker characteristics has been a major stumbling block to empirical estimation of compensating wage premiums, especially for early studies. Ideal risk measures should reflect subjective assessments of the risks associated with each job by both workers and employers. In fact, most studies have relied on information from national data sets that typically provide information on several thousand workers and their occupations (for a through discussion of this point, see Viscusi, 1993).

---

[12] In this example, emphasis is on the amount that workers are willing to forgo to reduce risk, i.e., willingness to pay is calculated. In much of the compensating wage literature emphasis is placed on the increase in compensation that workers require in order to assume more risk, i.e., willingness to accept is calculated. Experimental evidence has routinely shown that willingness-to-accept is greater than willingness to pay: individuals require a larger financial inducement to accept a risk than they are willing to pay to avoid a risk (Morrison (1998)). Viscusi (1993) argues that for small changes in risk, willingness to pay and willingness to accept should be the same.

[13] To include these characteristics, empirical studies must have access to micro-level data sets, something which proved problematic in the early compensating-wage literature.

Wage premiums observed in the market are a result of the interaction of labor supply and labor demand, as conditioned by the characteristics of the job and individual worker preferences. The willingness-to-pay or willingness-to-accept measures are the result of holding the expected utility with respect to risk and income constant for the individual worker, while varying risk levels.

There is wide variation in the empirical estimates generated by the compensating wage technique, particularly when the earlier econometric studies are included in the comparison. For non-fatal job risk, empirical studies have encountered two difficulties. The first arises in untangling premiums for non-fatal and fatal risk in those cases where the two types of risk are correlated. Failure to account for non-fatal risk leads to bias in many fatality risk studies. The second difficulty arises because of data discrepancies: there is currently no up-to-date government data base that covers both fatal and nonfatal injuries (Viscusi (1993) discusses this point). Viscusi (1993) surveys 24 labor-market studies covering diverse populations and diverse types of injuries. He finds that, in general, empirical studies find statistically significant wage premiums for job injury risk. Most of the estimates based on data for all injuries regardless of severity are clustered in the $25,000-$50,000 range, with the wage-risk trade-off tending to be greater for more severe types of injuries.

Empirical studies of fatal risk tradeoffs yield results differing by a couple orders of magnitude. A fairly wide range of results is not surprising as empirical studies have focused on different populations of workers and include different measures of risk and compensation. Fisher et al. (1989) and Viscusi (1993) review the empirical literature and both conclude that the most reliable compensating-wage studies include variables detailing worker and job characteristics. In addition, the most credible of the studies are those that have been the most successful in measuring specific job-related risk (as opposed to occupation-related risk or general categories of risk). Fisher et al. (1989) surveyed 15 compensating-wage studies. In their judgment, the most defensible empirical results lie in the $1.6 to $8.5 million range (1986 dollars), with the best estimates lying at the lower end of the range (Gegax et al., 1991 with an estimate of $1.6 million and Dillingham, 1985 with an estimate of $2.5 million).[14] For the principal labor market studies surveyed by Viscusi (1993), implicit value of life estimates (deflated to 1990 dollars) are centered in the $3 million to $7 million range. Of the 24 studies, Viscusi places the most confidence in the estimates derived from wage equations, as the values derived from structural models are less robust. He favors results from his own studies (Viscusi, 1979), with an implicit value of life estimate of $4.1 million (deflated to 1990 dollars), and Moore and Viscusi (1988) with an estimate of $2.5-$7.3 million ($1990).[15]

Much of the criticism of the compensating-wage approach centers on its assumptions concerning the labor market. Many critics argue that the actual labor market bears little resemblance to the labor market described in compensating-wage models (see for example, Dorman, 1996). The compensating wage approach assumes that workers are fully cognizant of the extent and consequences of the on-the-job risks they face,[16] that labor markets are strictly competitive,[17] and that insurance markets are actuarially correct, with premiums and payouts matched to accurately assessed risks. In addition, compensating-wage models have difficulty consistently accounting for job characteristics that might substitute for wages in compensating for risk such as prestige, flexible hours, and a pleasant work environment.

---

[14] Fisher et al. examined the Gegax et al. paper before it was a journal publication.

[15] Fisher et al., commend the Moore and Viscusi study for using data from the National Institute of Occupational Safety and Health (NIOSH) instead of Bureau of Labor Statistics data. However, they argue there are costs to using (NIOSH) data. Though the NIOSH data include a complete census of all occupational fatalities and match death risk by State with compensation in each State, NIOSH data are only disaggregated to the one-digit SIC code, meaning that fairly diverse jobs are characterized by the same level of risk.

[16] Viscusi suggests that the fact that workers are not always well informed leads to the "quit effect" (Viscusi and Moore, 1991).

[17] In the compensating-wage literature, this assumption is questioned through examining the wage-risk premiums paid to union and non-union members. Fisher et al. (1989) discuss this evidence.

Another critique of the approach lies in the observation that not all risks are the same. For example, it can be argued that not all fatality risks represent the same utility loss. That is, not just the likelihood but also the manner in which a person might die makes a difference. Equiprobable risks of dying in an industrial accident or from food poisoning may not be equally undesirable. In addition, people are usually less willing to accept involuntary risk than risk that is voluntarily assumed through, say, a wage contract. As a result, studies, such as compensating-wage studies, that measure response to voluntary risk probably underestimate society's aversion to risk that is not contracted for. Implicit value-of-life estimates are quite sensitive to the level and type of risk under consideration, and values derived with respect to one sort of risk may not be accurate measures of the value of other sorts of risk.

One of the most common criticisms of the compensating-wage approach relates to the final use of the estimates rather than to the generation of the estimates. Compensating-wage studies are primarily restricted to high-risk blue-collar males, and yet even within this restricted population, the implicit value of life estimates fluctuate wildly. Age, experience, education, sex, and most significantly, wealth should influence willingness-to-pay or willingness-to-accept. The results of one compensating wage study can hardly be compared with the results of another due to heterogeneity problems. Nevertheless, these results are often applied to the general population, a population that may have very different attitudes to risk and health than the typical high-risk blue-collar male. Many government agencies have adopted Viscusi's mid-range estimates as official policy, requiring that these estimates be used in all analyses, regardless of the type of hazard and who is at risk. The Food and Drug Administration (Food and Drug Administration, 1995) and the Consumer Product Safety Commission (Miller et al., 1997) currently use Viscusi's midpoint value of $5 million for each life saved. The Department of Transportation used a value of $2.2 million for many years (Viscusi and Hamilton, 1996), but has recently raised the value to $2.7 million.

The most striking observation that emerges from the compensating wage literature is the sensitivity of value-of-life estimates to the characteristics of the study population and to the level and type of risk. As a result, the general applicability of these estimates is questionable. "The value of life is not a universal constant, but reflects the wage-risk trade-off pertinent to the preferences of the workers in a particular sample" (Viscusi, 1993, p. 1930). At best, compensating wage studies indicate a range for implicit value-of-life measures, but caution should be exercised in making general conclusions about the value of life.

Contingent Valuation

Contingent valuation is a tool designed to allow analysts to estimate demands for goods that are not traded or only rarely traded. It is a survey method in which respondents are asked to state their preferences in hypothetical or contingent markets. The contingent-valuation method was first used to estimate the benefits of a recreation area in Maine (Davis, 1963) and continues to be widely used by environmental economists and public-good economists.

With the contingent-valuation method, analysts first draw a sample of individuals who are asked about a change in government policy governing, for example, pollution control, scenic area regulations, hunting permit allocation, or the supply of environmental amenities. Individuals usually are asked to imagine that there is a market in which they could buy such amenities. Respondents are given a detailed description of the hypothetical market and the good being evaluated. Then, they are asked the price they would pay to receive the amenity. Typically respondents do not make cash transactions, but are asked about willingness to participate in such transactions as if there were a market. Questions about the value of policy changes are hypothetical.[18]

Analysts also collect information on the demographic and socioeconomic characteristics of respondents (including age, sex, education, and income). Demographic characteristics allow analysts to draw inferences about the entire population of beneficiaries and the aggregate demand for amenities. In effect, they estimate aggregate willingness-to-pay. If analysts can show that preferences for amenities are not random, but vary systematically, conditioned by

[18] For more complete descriptions of the contingent-valuation technique, see Mitchell and Carson (1989) or Cummings et al. (1986).

observable demographic characteristics, then they can use population information on age, sex, education, and income to forecast aggregate demand.

The measure elicited by a contingent-valuation survey is a Hicksian willingness-to-pay measure (compensating surplus), a dollar measure of preferences. It is equivalent to a change in income, coupled with a change in the amenity under study, that leaves the respondent's utility level unchanged. Contingent valuations do not constrain the range of prices that individuals may report. Thus, such estimates are consistent in spirit with economic notions of utility: preferences are idiosyncratic and choices depend entirely on subjective judgments.

A primary undesirable characteristic of contingent valuation is that it does not require cash transactions. Individuals may not truthfully tell interviewers their real demands. Individuals may not be sufficiently able to judge their own demands without the requirement of giving up something for their choices. Contingent-valuation practitioners have developed guidelines to minimize biases and errors arising due to the hypothetical nature of the method.

To minimize unsystematic errors and enhance a study's reliability, Mitchell and Carson (1989) stress that the key scenario elements must be understandable, meaningful, and plausible to respondents. They suggest three guidelines to encourage this result. First, the WTP questions must be clear and unambiguous. Second, respondents should be familiar with the commodity to be valued. Third, respondents should have had prior valuation and choice experience with respect to consumption levels of the commodity, thus increasing the likelihood that they will have well-formed values for the commodity.

To minimize systematic bias and increase a study's validity, potential response biases must be controlled. Mitchell and Carson (1989) argue that systematic biases commonly occur in contingent valuation studies for three main reasons. First, the scenario contains strong incentives for respondents to misrepresent their true WTP amounts, thus resulting in strategic or compliance bias. Second, the scenario contains implied value cues that help determine WTP amounts. Third, there is misspecification (or misperception) of the scenario. There is no objective test to detect systematic bias, making it incumbent on the researcher to demonstrate that bias has been minimized.[19]

The issue for contingent-valuation studies involving health is whether these studies can comply with the above guidelines, or whether the special nature of the commodity "health" makes compliance unlikely. Whether or not a health-risk contingent valuation study is reliable and valid will depend not only on the design of the survey, but also on the exact nature of the health risk being assessed. Health-risk studies on mild illnesses that affect everyone sooner or later have a greater chance of being understandable, meaningful, and plausible than studies on severe, rare diseases. Ensuring that respondents are rational and knowledgeable will be more difficult for some health-risks than others. Even if the guidelines for reliability and validity have been reasonably met, a fairly standard rule of thumb places the accuracy of contingent-valuation estimates in the range of plus or minus 50 percent (Cummings et al., 1986). Contingent-valuation estimates should be interpreted in light of this accuracy range.

The use of contingent-valuation surveys to gauge the value of health and life is linked with the environmental literature in a large number of studies valuing health and environmental quality.[20] These types of contingent-valuation studies typically result in estimates of the per-day value of reducing specific, less severe symptoms such as coughing, sneezing, or throat or sinus problems. The comparability across studies tends to be limited because they pertain to diverse symptoms and differ in their reporting of marginal versus average values and median versus mean bids. In their review of contingent valuation studies evaluating less severe symptoms, Kenkel et al. (1994) find that once they control for differences in reporting, contingent-valuation estimates are relatively consistent. This observation bolsters the con-

---

[19] Kenneth Arrow and Robert Solow co-chaired a Contingent Valuation Panel that delivered a widely quoted critique of that approach in its analysis of natural resource damage assessments under the Oil Pollution Act of 1990. The Panel produced guidelines to improve the reliability of any CV study (Federal Register, Vol. 58, No. 10).

[20] For a review of this literature see Kenkel et al., 1994.

clusion that in the case of less severe illnesses, contingent-valuation surveys are reasonably able to follow the guidelines for reliability and validity.

For severe health symptoms, the use of the contingent-valuation technique is more questionable. In these cases, respondents are probably not adequately familiar with most life-threatening illnesses, nor are they likely to be experienced in deciphering probabilities related to severe health risks. Nevertheless, because serious illness has an impact on both the probability of death and on the quality of life, the contingent-valuation technique may be the approach that is best suited to measuring serious illness. The hypothetical nature of this approach could prove valuable in deciphering the effect of quality and quantity of life on value estimates. To this end, Fabian et al. (1994) developed an approach that prepares respondents to think carefully about the probabilities of serious illness. Questions progress from those dealing with simple life-experience situations to more complicated situations involving various probabilities of serious illness and death.

The Fabian et al. approach yields life-path scenarios that are combined with probability analysis to determine one's willingness to pay to reduce the risk of undesirable scenarios. Despite the complexity of the approach and the length of the questionnaire, the validity and reliability of the results are not assured. Fabian et al. (1994) highlight two areas of concern: the inability of respondents to discriminate between one risk and another, and the sensitivity of results on the amount of information provided to the respondent. These concerns are compounded in market experiments in which risk and probability information are not carefully detailed.

Contingent-valuation studies that specifically examine willingness-to-pay for changes in life-threatening risk are among some of the earliest applications of the approach (Acton, 1973 and Jones-Lee, 1976), and as such are subject to some start-up errors. As a result, Fisher et al. (1989) focus their review of this literature on two fairly recent additions: a study by Jones-Lee et al. (1985) and one by Gegax et al. (1991).[21] Fisher et al. praise both of these earlier studies for focusing on risks that are familiar to the survey respondents and for each study's attention to creating surveys with realistic and well-defined scenarios and payment mechanisms. The Jones-Lee et al. study examined individuals' willingness-to-pay for reducing the risk of serious motor vehicle accidents in Great Britain. Their results yield value-of-life estimates between $1.6 and $4.4 million (1986 dollars). Gegax et al. examined willingness to pay for reductions in job-related risks. Their value-of-life measures ranged from $2.4 million to $3.3 million per statistical life (compared with a $1.6-million estimate from their wage-risk analysis). Again, as in the case with compensating-wage studies, the results are sensitive to the types of risk under analysis.

In some cases, contingent markets might lead to more reliable estimates of willingness-to-pay than do prices from active markets. Contingent-valuation studies may be able to eliminate biases resulting from the physician-agent relationship, insurance arrangements, and irrationality in the face of severe disease (Golan and Shechter, 1993). Viscusi (1993) argues that contingent-valuation studies may be able to avoid some of the other shortcomings of market-generated estimates in that contingent-valuation studies estimate more than one value along the respondent's constant expected utility locus while wage-based studies measure only one point. Contingent-valuation studies are able to elicit more than just a point tradeoff; they estimate a respondent's utility function. The contingent-valuation approach can therefore avoid some of the heterogeneity problems inherent in labor-based estimates by making the parameters of the utility function dependent on worker characteristics. Such an approach explicitly models a value-of-life estimate as a function of income level and nonmarginal changes in risk. Another advantage of the contingent-valuation technique is that it is not constrained by circumstance: it can investigate issues for which there are no market data, and it can circumvent income constraints to derive estimates that more truly reflect preferences and not income. Of course, these strengths could prove to be weaknesses if the process does not measure real decisions regarding scarce resources. Though the contingent-valuation approach has been used extensively in the natural resources and environmental literature for the past 20 years, the technique is only slowly being applied to health-risk questions. However, recent successes in creating valid and reliable surveys could bolster research using this approach. Early skepticism regarding the application

[21] See footnote 12.

of the contingent-valuation method to the special commodity "health" seems to be giving way to the realization that the method could prove useful in exploring health-risk tradeoffs that are obscured in market data.

Contingent valuation of food safety overcomes the problem that food is not marketed by risk levels (say, probabilities of inducing cancer) and that it is therefore difficult to assign a value to risk reduction. Contingent valuation overcomes this problem by providing survey respondents with assessments of health risk. Valuation of food safety in experimental markets attempts to go one step further—placing the good in a market-like situation where money changes hands.

Application of experimental valuation to food safety is relatively new (Hayes et al., 1995, and Fox et al., 1995). Experimenters have used auction mechanisms to establish a market-like setting under controlled conditions. Whether experimental markets elicit truthful revelations of preferences is an open question. On one hand, participants make monetary payments for goods they consume, suggesting participants are aware of the opportunity cost of their bidding behavior. On the other hand, the experimental market is still artificial and contrived; participants bid with money experimenters give them. Thus, it is not entirely clear that the opportunity cost participants incur by bidding exactly equals the cost they would realize if there were a real market for safety. For a more thorough review of this literature see Buzby et al., 1998.

Household Health Production

The household health-production function method for measuring WTP is built on the observation that households continually make decisions involving the allocation of income and time between health-enhancing goods and activities and other goods and activities. In addition to *ex post* health-care consumption items, like prescription medicines and surgeries, *ex ante* or preventive items like diet, exercise, work and leisure choices also affect health status. The household health-production approach recognizes that health is not simply an exogenous variable, but that individuals can and do make decisions attempting to influence their own health status. By maximizing a utility function that explicitly includes health expenditures, subject to an income-time budget constraint that accounts for productivity losses due to ill health, theorists using the household health-production approach are able to solve for the willingness-to-pay for health.

Grossman (1972) first modeled the trade-off between health-enhancing activities and income and leisure. Grossman's health-production model incorporates two distinct roles for good health in household demand. First, good health is a "capital" stock. Investments in health capital determine the amount of time that can be devoted to producing and consuming.[22] Second, good health is a fundamental commodity. In this distinction, Grossman adopts a conceptual separation between commodities (fundamental objects of choice) and market goods (Becker, 1965; Lancaster, 1966; Muth, 1969). Fundamental commodities, like good health and peace of mind, are not purchased but instead are produced by the individual. Purchased goods and services and the individual's time also are inputs used to produce fundamental commodities. The Grossman model incorporates these two distinct roles for health (health capital and fundamental commodity), and as a result, health is demanded by consumers in the model for two reasons: as a fundamental consumption commodity that enters directly in the utility function, and as an investment commodity determining the total amount of time available for market and non-market activities. Maximization of the Grossman model results in a WTP amount for the value of healthy time that further mirrors the two roles of good health. This WTP amount is the sum of two elements: the monetary value of the direct increase in utility associated with better health and the increased labor earnings due to better health. A primary criticism of the Grossman model is that it succeeds in endogenizing good health to such as extent that individuals in the model are able to choose their length of life.

Berger et al. (1994) develop a model that shows relations among a production function, COI, and WTP. This model includes health in three roles, as a variable in the utility function, as a determinant in the

[22] This is an extension of the human-capital model developed by Becker, 1964, and Ben-Porath, 1967.

probability of survival for the current period, and as part of the income constraint. It considers traditional cost-of-illness measures (medical expenditures and income losses) and preventive expenditures (goods and time) where cost-of-illness is a function of health characteristics, and health is a function of preventive expenditures and an exogenous shift variable (such as environmental quality). Risk is incorporated into the model through the specification of a probability density function for health. This probability density function determines the likelihood of a particular health status (given preventive expenditures and the state of the world), which in turn determines the probability of survival for the period. Through their health-production function, Berger et al. are able to solve for an individual's *ex ante* WTP for an improvement in health status. They demonstrate that WTP for combined morbidity and mortality risks is not the sum of the WTP for each individual type of risk.

Examples of health-production functions that have been empirically estimated include Cropper's (1981) study of air pollution and work-loss days; Gerking and Stanley's (1986) study of ozone reduction and morbidity; and Dickie and Gerking's (1991) study of health attributes, private goods, and air quality (see Clemmer et al., 1994, for a review of these studies). The WTP amounts generated with these studies range from $0.73 for the reduction of symptoms to $176 for a work-loss day. This variability in WTP estimates illustrates the difficulty in consistent application of the household production approach and the difficulty in comparing estimates across studies.

A complete model of health behavior that endogenizes health investment should mirror the choices people make concerning health and consumption and leisure. However, a fundamental difficulty with the health-production approach is that even at a theoretical level, it is difficult to identify all the elements that contribute to the production and maintenance of good health (Harrington and Portney, 1987; Atkinson and Crocker, 1992). Empirical measurement of these elements, once identified, is also a difficult task requiring the quantification of non-marketed and often intangible goods. In addition, the econometric estimation of health-production functions is problematic (Harrington and Portney, 1987). Mullahy and Portney (1990) highlight the difficulties of empirical estimation in cases where health inputs, not just health, are endogenous. Bockstael and McConnell (1983) demonstrate that the household health production function may be unable to easily estimate the value of non-marginal changes. As a result of all these difficulties, the household production approach is subject to serious measurement error and is restricted in its application. Berger et al. (1994) conclude that "the health production function approach to estimating WTP may be of limited usefulness" (p. 34).

Hedonic Approach—Other Market Evidence

The health-production approach incorporates the observation that many goods and services contribute to health status. The hedonic approach extends this through the observation that often only specific characteristics of a good or service contribute to health, with other characteristics serving other functions. The final price of a good or service will reflect the desirability of all its characteristics or attributes. For example, the attributes of a house include size, comfort, and location, and the price of the house will reflect all three attributes. If the attributes of the house include characteristics that affect health such as location in a polluted neighborhood or access to the purest water in the country, the price of the house should reflect the value of these health-influencing attributes. With the hedonic method, the value of each attribute of a good or service is calculated, and the WTP for each attribute, including health-related attributes is estimated.

Market studies evaluating the health risk tradeoff implicit with the purchase or use of a variety of goods and services have been accomplished. Viscusi's 1993 survey of the empirical literature includes seven value-of-life studies estimating the implicit health risk tradeoff in decisions regarding highway speed, seat belts, smoke detectors, smoking, car purchases, and property values. Viscusi (1993) argues that non-labor market studies are less direct and probably less reliable than labor market studies (compensating-wage studies), because they do not observe either the risk facing the individual or the monetary value of the attribute. Furthermore, Viscusi contends that these studies:

> . . . provide a lower bound on the value of life, but will not provide information about the consumer's total willingness to pay for safety,

> because with such discrete decisions consumers are not pushed to the point where the marginal cost of greater safety equals its marginal valuation. (p. 1936)

The implicit value-of-life estimates included in the Viscusi survey center around $1 million, a number that is quite low in comparison with other value-of-life measures.

Fisher et al. (1989) include four consumer market studies in their survey. The results of these studies are much lower than those generated by other willingness-to-pay estimation methods, with value-of-life measures ranging from $.24 to $1.4 million (1986 dollars). Fisher et al. center these estimates at about $.55 million. They believe the estimates are low because the assumptions in these studies lead to an incomplete accounting of WTP. For example, the assumption that the time spent buckling up is the only cost of putting on a seat belt leads to understatements of the implicit value of life. Many people feel uncomfortable wearing seatbelts. If this discomfort were included in the estimates, both the cost of wearing seatbelts and the implicit value of life would be higher.

Another vein of the literature using the hedonic method involves linking property values and the value of health. In this literature, investigators estimate what individuals would be willing to pay for improvements in health by observing property values in neighborhoods with varying levels of air pollution. Everything else equal, property values in neighborhoods with lower levels of air pollution should be higher than property values in more polluted neighborhoods. In cases where air pollution can be linked to adverse health effects (real or potential), differences in property values can be used to estimate health values.[23] Klemmer et al. (1994), survey the literature on hedonic pricing of housing characteristics and report the results of empirical studies estimating the value of reductions in air pollution and studies estimating the elasticities of demand for clean air. None of these studies explicitly estimate the value of health. These studies indicate that clean air (and hence the health benefits associated with clean air) is a normal good with a demand that is relatively inelastic, though negatively related to price. Though these results are rather innocuous, Klemmer et al. harshly criticize the approach taken in many of these studies and advocate caution in interpreting these estimates. In particular, they argue that studies that depend on the approach developed by Rosen (1974) suffer from inadequate exogenous price variation. Therefore the benefit estimates obtained from these studies are not very reliable. In addition, Klemmer et al. question Rosen's handling of the identification problem and simultaneity in an implicit market analysis.

In general, the hedonic methodology has yet to be refined for valuing health attributes associated with market goods. The value-of-life estimates resulting from this methodology are much lower than those estimated by other techniques. This discrepancy should be explored before these values are used in other contexts.

Conclusion

The WTP approach reflects individual preferences for risk reduction where the demand for risk reduction is derived from *ex ante*, or expected health benefits. WTP reflects the value of benefits to those whose lives are improved by policies, and the value should represent complete compensation for those who might be harmed. These quantities exist only *ex ante*, at the moment of choice. They are not equivalent to realized damages.

The WTP approach reflects the observation that individual preferences are unique, and individual demands for risk reduction vary. However, because health and safety are normal goods, some of the variance in WTP estimates will be explained by income differences rather than preferences. So, just as in COI analysis, income and circumstance could play a role in determining the size of WTP estimates.

In practice, regulatory agencies that have adopted WTP have generally adopted a single value for lives saved where the value has been derived from compensating-wage studies. Agencies apply their selected value to every health risk, regardless of the population likely to receive program benefits, the type of risk that might be mitigated, or the level of risk miti-

[23] Many contingent valuation studies also exploit this relationship.

gated. This practice not only undermines the theoretical validity of the WTP approach but also flies in the face of empirical evidence. The most striking conclusion that emerges from the literature on empirical estimation of WTP is the sensitivity of value-of-life estimates to the characteristics of the study population, the level of risk, and the type of risk. Different populations faced with different risks will place a different value on life and health. There is no universal value that can be used in every situation.

Using the contingent-valuation method for valuing health and safety allows researchers to develop a more thorough mapping of risks and preferences. But, the method relies on consumers' claims about what they would be willing to buy in an imaginary market; consumers do not have to give up anything to respond to analysts' questions. Many economists remain skeptical about applying contingent-valuation techniques to health valuation. However, recently, some have argued that the method could prove useful in exploring health risk tradeoffs that are obscured in market data. For example, Kenkel et al. (1994) portray contingent valuation as the only method for untangling morbidity and mortality issues. In some cases, contingent markets might lead to more reliable results than actual markets because contingent markets are able to eliminate biases resulting from the physician-agent relationship, insurance arrangements, and irrationality in the face of severe disease.

It is possible that with additional studies, analysts will be able to estimate the demand for risk reduction throughout the population (and to separate preferences from income constraints). At that time, analysts will be faced with exactly the same problem facing those using COI. There will be a range of values that vary demographically. Cost-benefit analysts using WTP estimates will then be back in the awkward position of assigning different values to different individuals.