Chapter 7
Analyzing Benefits

The aim of an economic benefits analysis is to estimate the benefits, in monetary terms, of proposed policy changes in order to inform decision making. Estimating benefits in monetary terms allows the comparison of different types of benefits in the same units, and it allows the calculation of net benefits — the sum of all monetized benefits minus the sum of all monetized costs — so that proposed policy changes can be compared to each other and to the baseline scenario.

The discussion in this chapter focuses on methods and approaches available to monetize benefits in the context of a "typical" EPA policy, program, or regulation that reduces emissions or discharges of contaminants. This is not to say that those benefits that cannot be monetized due to lack of available values or quantification methods are not important. Chapter 11 on the "Presentation of Analysis and Results" discusses how to carry forward information on non-monetized benefits to help inform the policy-making process. In addition, this chapter includes a discussion of several alternatives to monetization that may add some context to this category of benefits. The general monetization methods and principles discussed here should apply to other types of EPA polices as well, such as those that provide regulatory relief, encourage reuse of remediated land, or provide information to the public to help people avoid environmental risks.[1]

7.1 The Benefits Analysis Process

Ideally, benefits analyses would consist of comprehensive assessments of all environmental effects attributable to the rule in question. However, it is seldom possible to analyze all effects simultaneously in an integrated fashion. In most cases, analysts will need to address each effect individually, and then aggregate the individual values to generate an estimate of the total benefits of a policy. A constant challenge in employing an effect-by-effect approach is to balance potential trade-offs between inclusion and redundancy.

Ideally, each effect will be measured once and only once. Techniques intended to bring additional effects into the analysis may run the risk of double counting effects already measured. For example, stated preference methods may be the only way to measure non-use values, but they may double count use values already reflected in hedonic or travel cost analyses. Therefore, the analyst should be careful in interpreting and combining the results of different methods.

There are of course exceptions to this "effect-by-effect" approach to benefits analysis (e.g., efforts to estimate the social benefits of reducing CO_2 emissions — see Text Box 7.1), but the remainder of the discussion below is framed with this approach in mind.

A second challenge analysts often face is the difficulty of conducting original valuation research in support

1 Other methods, such as cost-effectiveness analysis (CEA), can also be used to evaluate policies. CEA does not require monetization of benefits but rather divides the costs of a policy by a particular effect (e.g., number of lives saved). CEA can be used to compare proposed policy changes on an effect-by-effect basis, but, unlike BCA, cannot be used to calculate a single, comprehensive measure of the net effects of a policy, nor can it compare proposed policy changes to the status quo. Other methods for evaluating policies (e.g., distributional analyses) are covered in Chapter 10.

Text Box 7.1 - Estimating Benefits from Reducing Carbon Dioxide Emissions: The Social Cost of Carbon

Monetized estimates of the damages associated carbon dioxide (CO_2) emissions allows the social benefits of regulatory actions that are expected to reduce these emissions to be incorporated into BCA. One way to accomplish this is through the estimation of the "social cost of carbon" (SCC). The SCC is the present value of the stream of future economic damages associated with an incremental increase (by convention, one metric ton) in CO_2 emissions in a particular year. It is intended to be a comprehensive measure and includes economic losses due to changes in agricultural productivity, human health risks, property damages from increased flood frequencies, the loss of ecosystem services, etc. The SCC is a marginal value so it may not be accurate for valuing large changes in emissions. However, many U.S. government regulations will lead to relatively small reductions in cumulative global emissions, so for these regulations the SCC is the appropriate shadow value for estimating the economic benefits of CO_2 reductions.

Most published estimates of the SCC have been derived from "integrated assessment models" (IAMs) that combine climate processes, economic growth, and feedbacks between the two in a single modeling framework. These models include a reduced form representation of the potential economic damages from climate change. Therefore IAMs used to estimate the SCC are necessarily highly simplified and limited by the current state of the climate economics literature, which continues to expand rapidly. Despite the inherent uncertainties in models such as these, they are the best tools currently available for estimating the SCC.

The Interagency SCC Workgroup. In 2009, an interagency workgroup composed of members from six federal agencies and various White House offices was convened to improve the accuracy and consistency in how agencies value reductions in CO_2 emissions in regulatory impact analyses. The resulting range of values is based on estimates from three integrated assessment models applied to five socioeconomic and emissions scenarios, all given equal weight. To reflect differing expert opinions about discounting, the present value of the time path of global damages in each model-scenario combination was calculated using discount rates of 5 percent, 3 percent, and 2.5 percent. Finally, in a step toward more formal uncertainty analysis, all model runs employed a probabilistic representation of climate sensitivity (in addition to other parameters in two of the models).

The workgroup selected four SCC estimates from the model runs to reflect the global damages caused by CO_2 emissions: $5, $21, $35, and $65 for 2010 emission reductions (in 2007 U.S. dollars). The first three estimates are based on the average SCC across the three models and five socioeconomic and emissions scenarios for the 5 percent, 3 percent, and 2.5 percent discount rates, respectively. The fourth value, the 95th percentile of the SCC distribution at a 3 percent discount rate, was chosen to represent potential higher-than-expected impacts from temperature change. The SCC estimates grow over time at rates endogenously determined by the models. For instance, with a discount rate of 3 percent, the mean SCC estimate increases to $24 per ton of CO_2 in 2015 and $26 per ton of CO_2 in 2020.

Going Forward. The Interagency SCC Workgroup presented the SCC estimates with a clear acknowledgement of the many uncertainties involved and the final report outlined a number of limitations to the analysis. The Interagency SCC Workgroup is committed to re-visiting these estimates on a regular basis and revising them as needed to reflect the growing body of scientific knowledge regarding climate change impacts and the potential economic damages from those impacts.

Further Reading: U.S. Interagency Working Group on Social Cost of Carbon (2010). Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866, www.epa.gov/otaq/climate/regulations/scc-tsd.pdf.

of specific policy actions. Because budgetary and time constraints often make performing original research infeasible, analysts regularly need to draw upon existing value estimates for use in benefits analysis. The process of applying values estimated in previous studies to new policy cases is called *benefit transfer*. The benefit transfer method is discussed in detail in Section 7.4, but much of this chapter is written with benefit transfer in mind. In particular, the descriptions of revealed and stated preference valuation methods in Sections 7.3.1 and 7.3.2 include recommendations for evaluating the quality and suitability of published studies for use in benefit transfer.

A general "effect-by-effect" approach to benefits analysis

This approach consists of separately evaluating the major effects of a given policy, and then summing these individual estimates to arrive at an overall estimate of total benefits. The effect-by-effect approach for benefits analysis requires three fundamental steps:

1. Identify benefit categories potentially affected by the policies under consideration;
2. Quantify significant endpoints to the extent possible by working with managers, risk assessors, ecologists, physical scientists, and other experts; and
3. Estimate the values of these effects using appropriate valuation methods for new studies or existing value estimates from previous studies that focus on the same or sufficiently similar endpoints.

Each step in this approach is discussed in more detail below. Analysts also should consider whether this general framework is appropriate for assessing a specific policy or whether a more integrated approach that incorporates all of the relevant effects simultaneously can be applied. When applying the effect-by-effect approach it is important to avoid double counting benefits across effects as much as possible. Collaboration with appropriate experts will be necessary to execute these steps meaningfully.

Step1: Identify potentially affected benefit categories

The first step in a benefits analysis is to determine the types of benefits associated with the policy options under consideration. More information on benefits categories can be found in Section 7.2. To identify benefit categories, analysts should, to the extent feasible:

Develop an initial understanding of policy options of interest by working with other analysts and policy makers. Initially, the range of options considered may be very broad. Resources should be focused on benefit categories that are likely to influence policy decisions. Collaboration between all parties involved in the policy analysis can help ensure that all potential effects are recognized and that the necessary and appropriate information and endpoints are collected and evaluated at each step in the process. Analysts should take care to think through potential secondary or indirect effects of the policy options as well, as these may prove to be important.

Research the physical effects of the pollutants on human health and the environment by reviewing the literature and consulting with other experts. This step requires considering the transport of the pollutants through the environment along a variety of pathways, including movement through the air, surface water, and groundwater, deposition in soils, and ingestion or uptake by plants and animals (including humans). Along these pathways, the pollutants can have detrimental effects on natural resources, such as affecting oxygen availability in surface water or reducing crop yields. Pollutants can also have direct or indirect effects on human health, for example affecting cancer incidence through direct inhalation or through ingestion of contaminated food.

Consider the potential change in these effects as a result of each policy option. If policy options differ only in their level of stringency then each option may have an impact on all identified physical effects. In other cases, however, some effects may be reduced while others are increased or remain unchanged. Evaluating how physical effects change under each policy option requires evaluation of how the pathways differ in the "post-policy" world.

Determine which benefit categories to include in the overall benefits analysis using at least the following three criteria:

1. Which benefit categories are likely to differ across policy options, including the baseline option? Analysts should conduct an assessment of how the physical effects of each policy option will differ and how each physical effect will impact each benefit category.
2. Which benefit categories are likely to account for the bulk of the total benefits of the policy? The cutoff point here should be based on an assessment of the magnitude and precision of the estimates of each benefit category, the total social costs of each policy option, and the costs of gathering further information on each benefit category. A benefit category should not be included if the cost of gathering the information necessary to include it is greater than the expected increase in the value of the policy owing to its inclusion. The analyst should make these preliminary assessments using the best quantitative information that is readily available, but as a practical matter these decisions may often have to be based on professional judgments.
3. Which benefit categories are especially salient to particular stakeholders? Monetized benefits in this category are not necessarily large and so may not be captured by the first two criteria.[2]

The outcome of this initial step in the benefits analysis can be summarized in a list or matrix that describes the physical effects of the pollutant(s), identifies the benefit categories associated with these effects, and identifies the effects that warrant further investigation.

The list of physical effects under each benefit category may be lengthy at first, encompassing all of those that reasonably can be associated with the policy options under consideration. Analysts should preserve and refine this list of physical effects as the analysis proceeds. Maintaining the full list of potential effects — even though the quantitative analysis will (at least initially) focus on a sub-set of them — will allow easy revision of the analysis plan if new information warrants it.

EPA has developed extensive guidance on the assessment of human health and ecological risks, and analysts should refer to those documents and the offices responsible for their production and implementation for further information (U.S. EPA 2009a). No specific guidance exists for assessing changes in amenities or material damages. Analysts should consult relevant experts and existing literature to determine the "best practices" appropriate for these categories of benefits.

Step 2: Quantify significant endpoints

The second step is to quantify the physical endpoints related to each category, focusing on changes attributable to each policy option relative to the baseline. Data are usually needed on the extent, timing, and severity of the endpoints. For example, if the risk of lung cancer is an endpoint of concern, required information will usually include the change in risk associated with each option, the timing of the risk changes, the age distribution of affected populations, and fatality rates. If visibility is the attribute of concern, required information will usually include the geographical areas affected and the change in visibility resulting from each policy option.

Analysts should keep the following issues in mind while quantifying significant physical effects.

Work closely with analysts in other fields. Estimating physical effects is largely, but not completely, the domain of other experts, including human health and ecological risk assessors and other natural scientists. These experts generally are responsible for evaluating the likely transport of the pollutant through the environment and its potential effects on humans, ecological systems, and manufactured materials.

2 This third criterion relates to distributional considerations detailed in Chapter 10.

Text Box 7.2 - Integrating Economics and Risk Assessment

Historically, health and ecological risk assessments have been designed not to support benefits analyses per se but rather to support the setting of standards or to rank the severity of different hazards. Traditional measures of risk can be difficult or impossible to incorporate into benefits analyses. For example, traditional measures of risk are often based on endpoints not directly related to health outcomes or ecological services that can be valued using economic methods. These measures are often based on outcomes near the tails of the risk distribution for highly sensitive endpoints, which would lead to biased benefits estimates if extrapolated to the general population.

Because economists rely on risk assessment outcomes as key inputs into benefits analysis, it is important that risk assessments and economic valuation studies be undertaken together. Economists can contribute information and insights into how behavioral changes may affect realized risk changes. For example, if health outcomes in a particular risk assessment are such that early medical intervention could reduce the chances of illness, economists may be able to estimate changes in the probability that individuals will seek preventative care. Even in cases where the economists' contribution to the risk characterization is not direct, economists and risk assessors should communicate frequently to ensure that economic analyses are complete. Specifically risk assessors and economists should:

- Identify a set of human health and ecological endpoints that are economically meaningful. The endpoints should be linked to human well-being and monetized using economic valuation methods. This may require risk assessors to model more or different outcomes than they would if they were attempting to capture only the most sensitive endpoint. This also may require risk assessors and economists to convert specific human health or ecological endpoints measured in laboratory or epidemiological studies to other effects that can be valued in the economic analysis.
- Estimate changes in the probabilities of human health or ecological outcomes rather than "safety assessment" measures such as reference doses and reference concentrations.
- Work to produce expected or central estimates of risk, rather than bounding estimates as in safety assessments. At a minimum, any expected bias in the risk estimates should be clearly described.
- Attempt to estimate the "cessation lag" associated with reductions in exposure. That is, the analysis should characterize the time profile of changes in exposures and risks.
- Attempt to characterize the full uncertainty distribution associated with risk estimates. Not only does this contribute to a better understanding of potential regulatory outcomes, it also enables economists to incorporate risk assessment uncertainty into a broader analysis of uncertainty. Formal probabilistic assessment is required for some regulations by *Circular A-4* (OMB 2003). Also refer to EPA's guidance and reference documents on Monte Carlo methods and probabilistic risk assessment, including EPA's Policy for Use of Probabilistic Analysis in Risk Assessments (U.S. EPA 1997e), and the 1997 Guiding Principles for Monte Carlo Analysis (U.S. EPA 1997d).

The principal role of the economist at this stage is to ensure that the information provided is useful for the subsequent economic valuation models that may be used later in the benefits analysis. The analyst should give special care to ensuring that the endpoints evaluated are appropriate for use in benefits estimation. Effects that are described too broadly or that cannot be linked to human well-being limit the ability of the analysis to capture the full range of a policy's benefits. Text Box 7.2 provides examples and a more detailed discussion.

Another important role for economists at this stage is to provide insights, information, and analysis on behavioral changes that can affect the results of the risk assessment as needed. Changes in behavior due to changes in environmental quality (e.g., staying indoors to avoid detrimental effects of air pollution) can be significant and care should be taken to account for such responses in risk assessments and benefit estimations.

Step 3: Estimate the values of the effects

The next step is to estimate willingness to pay (WTP) of all affected individuals for the quantified benefits in each benefit category, and then to aggregate these to estimate the total social benefits of each policy option. Typically, a representative agent approach is used when deriving estimates of benefits. The analyst calculates WTP for an "average" individual in a sample of people from the relevant population and then multiply that average value by the number of individuals in the exposed population to derive an estimate of total benefits. As discussed earlier, markets do not exist for many of the types of benefits expected to result from environmental regulations. Details on the economic valuation methods suitable for this step and examples of how they can be applied can be found in Section 7.3. In applying these methods, analysts should:

Consider using multiple valuation methods when possible. Different methods often address different subsets of total benefits and the use of multiple methods allows for comparison of alternative measures of value when applied to the same category of benefits. Double counting is a significant concern when applying more than one method. Any potential overlap should be noted when presenting the results. The discussion of benefit transfer in Section 7.4 describes many of the issues involved in applying value estimates from previous studies to new policy cases, including various meta-analysis techniques for combining estimates from multiple studies.

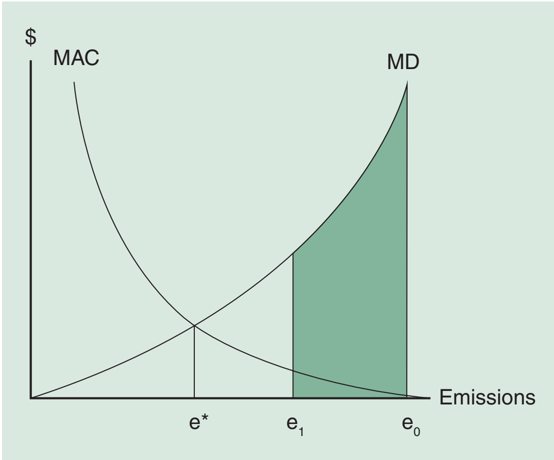
Describe the source of estimates and confidence in those sources. Valuation estimates always contain a degree of uncertainty. Using them in a context other than the one in which they were initially estimated can only increase that uncertainty. If many high-quality studies of the same effect have produced comparable values, analysts can have more confidence in using these estimates in their benefits calculations. In other cases, analysts may have only a single study, or even no directly comparable study, to draw from. In all cases, the benefits analysis should clearly describe the sources of the value estimates used and provide a qualitative discussion of the reliability of those sources. The analyst should include a quantitative uncertainty assessment when possible. Guiding principles for presenting uncertainty are addressed in Chapter 11.

7.2 Economic Value and Types of Benefits

Economic valuation is based on the traditional economic theory of human behavior and preferences, which centers on the concept of "utility" (or "satisfaction" or "welfare") that people realize from goods and services, both market and non-market. Different levels and combinations of goods and services afford different levels of utility for any one person. Because different people have different preferences, different sets of goods and services will appeal more or less to different people. Utility is inherently subjective and cannot be measured directly. Therefore, in order to give "value" an operational definition it must be expressed in a quantifiable metric. Money generally is used as the metric, but this choice for the unit of account has no special theoretical significance. One could use "apples," "bananas," or anything else that is widely valued and consumed by individuals. The crucial assumption is that a person could be compensated for the loss of some additional quantity of any good by some quantity of another good that is selected as the metric. Table 7.1 summarizes the types of benefits associated with environmental protection policies and provides examples of each of the benefits types, as well as valuation methods commonly used to monetize the benefits for each type.

When goods and services are bought and sold in competitive markets, the ratio of the marginal utility (the utility afforded by the last unit purchased) of any two goods that a person consumes must be equal to the ratio of the prices of those goods. If it were otherwise, that person could reallocate her budget to buy a little more

Figure 7.1 - Benefits of an Environmental Improvement



of one good and a little less of the other good to achieve a higher level of utility. Thus, market prices can be used to measure the value of market goods and services directly. A practical rationale for using money as the metric for non-market valuation is that money is the principal medium of exchange for the wide variety of market goods and services among which people choose on a daily basis.

The benefits of an environmental improvement are shown graphically in Figure 7.1. Reducing emissions from e_0 to e_1 produces benefits equal to the shaded area under the marginal damages (MD) curve. Many environmental goods and services, such as air quality and biological diversity, are not traded in markets. The challenge of valuing non-market goods that do not have prices is to relate them to one or more market goods that do. This can be done either by determining how the non-market good contributes to the production of one or more market goods (often in combination with other market good inputs), or by observing the trade-offs people make between non-market goods and market goods. One way or another, this is what each of the revealed and stated preference valuation methods discussed in Section 7.3 is designed to do. Of course, some methods will be more suitable than others in any particular case for a variety of reasons, and some will be better able to capture certain types of benefits than others. In principle, though, they are all different ways of measuring the same thing, which is the total amount of money required to make all individuals indifferent between the baseline and policy scenarios.

The economic valuation of an environmental improvement is the dollar value of the private goods and services that individuals would be willing to trade for the improvement at prevailing market prices. The willingness to trade compensation for goods or services can be measured either as *willingness to pay* (WTP) or *willingness to accept* (WTA). WTP is the maximum amount of money an individual would voluntarily pay to obtain an improvement. WTA is the least amount of money an individual would accept to forego the improvement.[3] The key theoretical distinction between WTP and WTA is their respective reference utility levels. For environmental improvements, WTP uses the level of utility *without* the improvement as the reference point while WTA uses the level of utility *with* the improvement as the reference point. Because of their different reference points, one relevant factor to consider when deciding whether WTP or WTA is the appropriate value measure to use in a BCA is the property rights for the environmental resource(s) in question. WTP is consistent with individuals or firms having rights to pollute and the affected parties needing to pay them to desist. WTA is consistent with individuals being entitled to a clean environment and needing to be compensated for any infringements of that right (Freeman 2003).

Economists generally expect that the difference between WTP and WTA will be small, provided the amounts in question are a relatively small proportion of income and the goods in question are not without substitutes, either market or non-market. However, there may be instances in which income and substitution effects are important.[4] To simplify the presentation, the term WTP is used throughout the remainder of this chapter to refer

3 For simplicity, the discussion in this section is restricted to the case of environmental improvements, but similar definitions hold for environmental damages. For a more detailed treatment of WTP and WTA and the closely related concepts of compensating variation, equivalent variation, and Hicksian and Marshallian consumer surplus, see Hanley and Spash (1993), Freeman (2003), Just et al. (2005), and Appendix A of these *Guidelines*.

4 For more information see Appendix A and Hanemann (1991).

to the underlying economic principles behind both WTA and WTP, but the analyst should keep the potential differences between the two measures in mind.

Based on the connection to individual welfare just described, estimates of WTP are needed for the Kaldor and Hicks potential compensation tests that form the basis of BCA (Boadway and Bruce 1984, Just et al. 1982, and Freeman 2003). To carry out these tests, sum the WTP for all affected individuals and compare the summed WTP value to the estimated costs of the proposed policy. Because environmental policy typically deals with improvements rather than deliberate degradation of the environment, WTP is generally the relevant measure.[5]

The types of benefits that may arise from environmental policies can be classified in multiple ways (Freeman 2003). As shown in Table 7.1, these *Guidelines* categorize benefits as human health improvements, ecological improvements, and other types of benefits, including aesthetic improvements and reduced materials damages, and list commonly used valuation methods for reference. The list is not meant to be exhaustive, but rather to provide examples and commonly used methods for estimating values.[6] The sections below provide a more detailed discussion of each of the benefit categories listed in Table 7.1.

7.2.1 Human Health Improvements

Human health improvements from environmental policies include effects such as reduced mortality rates, decreased incidence of non-fatal cancers, chronic conditions and other illnesses, and reduced adverse reproductive or developmental effects. While the most appropriate approach to valuation would consider mortality and morbidity together, in practice these effects are typically valued separately, and are therefore discussed separately in these *Guidelines*.

7.2.1.1 Mortality

Some EPA policies will lead to decreases in human mortality risks due to potentially fatal health conditions such as cancers. In considering the impact of environmental policy on mortality, it is important to remember that environmental policies do not assure that particular individuals will not die of environmental causes. Rather, they lead to small changes in the probability of death for many people.

EPA currently recommends a default central "value of statistical life" (VSL) of $7.9 million (in 2008 dollars) to value reduced mortality for all programs and policies.[7] This value is based on a distribution fitted to 26 published VSL estimates. The distribution itself can be used in uncertainty analysis. The underlying studies, the distribution parameters, and other useful information are available in Appendix B.

As a general matter, the impact of risk and population characteristics should be addressed qualitatively. In some cases, the analysis may include a quantitative sensitivity analysis. Analysts should account for latency and cessation lag when valuing reduced mortality risks, and should discount appropriately.

Valuing mortality risk changes in children is particularly challenging. EPA's *Handbook for Valuing Children's Health Risks* (2003b) provides some information on this topic, including key benefit transfer issues when using adult-based studies. *Circular A-4* also recognizes this subject, specifically advising: "For rules where health gains are expected among both children and adults and you decide to perform a BCA, the monetary values for children should be at least as large as the values for adults (for the same probabilities and outcomes) unless there is specific and compelling evidence to suggest otherwise" (OMB 2003). OMB guidance applies to risk of mortality and of morbidity.

5 See Section A.3 of Appendix A for further explanation of Kaldor-Hicks conditions.

6 In very rare cases with employment implications for the structurally unemployed, analysts may need to include job creation as a benefits category. See Appendix C for more detail.

7 This value is adjusted from the base value reported in U.S. EPA 2000d ($4.8 million in 1990 dollars) using the Consumer Price Index (CPI). The value is not adjusted for income growth over time.

Table 7.1 - Types of Benefits Associated With Environmental Policies: Categories, Examples, and Commonly Used Valuation Methods

Benefit Category	Examples	Commonly Used Valuation Methods
Human Health Improvements		
Mortality risk reductions	Reduced risk of: Cancer fatality Acute fatality	Averting behaviors Hedonics Stated preference
Morbidity risk reductions	Reduced risk of: Cancer Asthma Nausea	Averting behaviors Cost of illness Hedonics Stated preference
Ecological Improvements		
Market products	Harvests or extraction of: Food Fuel Fiber Timber Fur and Leather	Production function
Recreation activities and aesthetics	Wildlife viewing Fishing Boating Swimming Hiking Scenic views	Production function Averting behaviors Hedonics Recreation demand Stated preference
Valued ecosystem functions	Climate moderation Flood moderation Groundwater recharge Sediment trapping Soil retention Nutrient cycling Pollination by wild species Biodiversity, genetic library Water filtration Soil fertilization Pest control	Production function Averting behaviors Stated preference
Non-use values	Relevant species populations, communities, or ecosystems	Stated preference
Other Benefits		
Aesthetic improvements	Visibility Taste Odor	Averting behaviors Hedonics Stated preference
Reduced materials damages	Reduced soiling Reduced corrosion	Averting behaviors Production / cost functions

Note: "Stated preference" refers to all valuation studies based on hypothetical choices, as distinguished from "revealed preference," which refers to valuation studies based on observations of actual choices.

Methods for valuing mortality risk changes

Because individuals appear to make risk-income trade-offs in a variety of ways, the value of mortality risk changes are estimated using three primary methods. The most commonly used method is the hedonic wage, or wage-risk, method in which value is inferred from the income-risk trade-offs made by workers for on-the-job risks. Averting behavior studies value risk changes by examining purchases of goods that can affect mortality risk (e.g., bicycle helmets). Finally, stated preference studies are increasingly used to estimate WTP for reduced mortality risks. Key considerations in all of these studies include the extent to which individuals know and understand the risks involved, and the ability of the study to control for aspects of the actual or hypothetical transaction that are not risk-related. Because the value of risk reduction may depend on the risk context (e.g., work-related vs. environmental), results from any single study may not be directly applicable to a typical environmental policy case.

There are additional methods that can be used to derive information on risk trade-offs. Van Houtven et al. (2008) use a risk-risk trade-off model to examine preferences for avoiding fatal cancers. Carthy et al. (1999) examine trade-offs between fatal and non-fatal risks to indirectly estimate a WTP. This approach may make the task more manageable for the respondent, but the analyst should consider and evaluate the complexity of the additional steps and the indirect nature of the resulting estimates.

At one time, reduced mortality risk was valued under a human capital approach that equated the value of a statistical life with foregone earnings. This has largely been rejected as an inappropriate measure of the value of reducing mortality risks because it is not based on WTP for small risk reductions and as such does not capture the value associated with avoided pain and suffering, dread, and other risk factors that are thought to affect value (Viscusi 1993).

Previous studies

While there are many unresolved issues in valuing mortality risks, the field is relatively rich in empirical estimates and several substantial reviews of the literature are available. A general overview of common approaches and issues in mortality risk valuation can be found in Hammitt (2003). Viscusi (1993) and Viscusi and Aldy (2003) provide detailed reviews of the hedonic wage literature. Black, Galdo, and Liu (2003) provide a technical review of the statistical issues associated with hedonic wage studies. Blomquist (2004) provides a review of the averting behavior literature. Some key issues related to stated preference studies are included in Alberini (2004). Recently, some researchers have begun to use meta-analysis to combine study results and examine the impact of study design. Recent examples include Viscusi and Aldy (2003), Mrozek and Taylor (2002), and Kochi et al. (2006). EPA applications of VSL are numerous, and include the Clean Air Interstate Rule (CAIR), the Non-Road Diesel Rule, and the Stage 2 Disinfection By-Products Rule (DBP).[8]

Important considerations

The analyst should keep three important considerations in mind when estimating mortality benefits:

- Characterizing and measuring mortality effects;
- Heterogeneity in risk and population characteristics; and
- The timing of health risk changes.

Characterizing and measuring mortality effects

Reduced mortality risks are typically measured in terms of "statistical lives." This measure is the

8 The economic analyses for these three rules are available electronically as follows (accessed May 23, 2008):

CAIR (http://www.epa.gov/air/interstateairquality/pdfs/finaltech08.pdf);

Non-Road Diesel (http://www.epa.gov/nonroad-diesel/2004fr.htm#ria); and

Stage 2 DBP (http://www.epa.gov/safewater/disinfection/stage2/pdfs/anaylsis_stage2_ecconomic_main.pdf).

aggregation of many small risks over an exposed population. Suppose, for example, that a policy affects 100,000 people and reduces the risk of premature mortality by one in 10,000 for each individual. Summing these individual risk reductions across the entire affected population shows that the policy leads to 10 premature fatalities averted, or 10 statistical lives "saved."

Alternative measures attempt to capture the remaining life expectancy associated with the risk reductions. This is sometimes referred to as the "quantity of life" saved (Moore and Viscusi 1988) and is typically expressed as "statistical life years." Looking again at the policy described above, suppose the risks were spread over a population where each individual had 20 years of remaining life expectancy. The policy would then save 200 statistical life years (10 statistical lives times 20 life years each). In practice, estimating statistical life years saved requires risk information for specific subpopulations (e.g., age groups or health status). It is typical to use statistical life years saved in CEA, but valuing a statistical life year remains a subject of debate in the economics literature. Theoretical models show that the relationship between WTP and factors such as age, baseline risk, and the presence of co-morbidities is ambiguous and empirical findings are generally mixed (U.S. EPA 2006e).

Heterogeneity in risk and population characteristics

The value of mortality risks can vary both by risk characteristics and by the characteristics of the affected population. Key risk characteristics include voluntariness (i.e., whether risks are voluntarily assumed), timing (immediate or delayed), risk source (e.g., natural vs. man-made), and the causative event (e.g., cancer vs. accidents). Population characteristics include those generally expected to influence WTP for any good (e.g., income and education) as well as those more closely related to mortality risks such as baseline risk or remaining lifespan, health status, risk aversion, and familiarity with the type of risk. The empirical and theoretical literature on many of these characteristics is incomplete or ambiguous. For example, some studies suggest that older populations are willing to pay less for risk reductions (Jones-Lee et al. 1993), but others find this effect to be small if it exists at all (Alberini et al. 2004). Still others suggest older populations have higher WTP (Kniesner, Viscusi, and Ziliak 2006). Smith et al. (2004) and Viscusi and Aldy (2007a) discuss the relationship between age and VSL in the context of hedonic wage studies. Appendix B contains a more complete discussion of risk and population characteristics and how they may affect WTP.

Timing of health risk changes

Environmental contamination can cause immediate or delayed health effects. If individuals typically prefer health improvements earlier in time rather than later, all else equal, then the WTP for reductions in exposure to environmental pollutants will depend on when the resulting health risk changes will occur. The description here focuses on mortality risk, but the same principles apply to non-fatal health risks.

The effects of timing on the present or annualized value of reduced mortality risk can be considered in the context of a lifecycle consumption model with uncertain lifetime (Cropper and Sussman 1990, Cropper and Portney 1990, and U.S. EPA 2007g). In this framework reductions in mortality risk are represented as a shift in the survival curve — the probability an individual will survive to all future ages — which leads to a corresponding change in life expectancy and future utility.

If the basis for benefit transfer is a marginal WTP for contemporaneous risk reductions, then calculating the benefits of a policy with delayed risk reductions requires three steps: (1) estimating the time path of future mortality risk reductions; (2) estimating the annual WTP in all future years; and (3) calculating the present value of these annual WTP amounts. The first step should account for all the factors that ultimately relate changes in exposure to changes in mortality risk as defined by shifts in the survival curve.

7.2.1.2 Morbidity

Morbidity benefits consist of reductions in the risk of non-fatal health effects ranging from mild illnesses, such as headaches and nausea, to very serious illnesses such as cancer (see Table 7.1). Non-fatal health effects also include conditions such as birth defects or low birth weight. Non-fatal health effects differ with respect to the availability of existing value estimates. Values for reducing the risks of some of these health effects have been estimated multiple times using a variety of different methods, while others have been the subject of only a few or no valuation studies.

WTP to reduce the risk of experiencing an illness is the preferred measure of value for morbidity effects. As described in Freeman (2003), this measure consists of four components:

- "Averting costs" to reduce the risk of illness;
- "Mitigating costs" for treatments such as medical care and medication;
- Indirect costs such as lost time from paid work, maintaining a home, and pursuing leisure activities; and
- Less easily measured but equally real costs of discomfort, anxiety, pain, and suffering.

Methods used to estimate WTP vary in the extent to which they capture these components. For example, cost-of-illness (COI) estimates generally only capture mitigating and indirect costs, and omit averting expenditures and lost utility associated with pain and suffering. [9]

Methods for valuing morbidity

Researchers have developed a variety of methods to value changes in morbidity risks. Some methods measure the theoretically preferred value of individual WTP to avoid a health effect. Others can provide useful data, but that data must be interpreted carefully if it is to inform economically meaningful measures. Methods also differ in the perspective from which values are measured (e.g., before or after the incidence of morbidity), whether they control for the opportunity to mitigate the illness (e.g., before or after taking medication) and the degree to which they account for all of the components of total WTP. The three primary methods most often used to value morbidity in an environmental context are stated preference (Section 7.3.2), averting behavior (Section 7.3.1.4), and COI (Section 7.3.1.5). Hedonic methods (Section 7.3.1.3) are used less frequently to value morbidity from environmental causes.

Many other approaches do not estimate WTP and their ability to inform benefits analyses consequently varies. Risk-risk trade-offs, for example, do not directly estimate dollar values for risk reductions, but rather provide rankings of relative risks based on consumer preferences. Risk-risk trade-offs can be linked to WTP estimates for related risks.[10]

Other methods suffer from certain methodological limitations and are therefore generally less useful for policy analysis. For example, health-state indices, composite metrics that combine information on quality and quantity of life lived under various scenarios, are often used for cost-effectiveness or cost-utility analyses. These methods cannot be directly related to WTP estimates as the indices were developed using very different paradigms than those for WTP values. As such, they should not be used for deriving monetary estimates for use in BCA [Hammitt 2003, and Institute of Medicine (IOM) 2006], although there is evidence that components of these indices may still be useful in a benefit-transfer context (Van Houtven et al. 2006). Another commonly suggested alternative is jury awards, but these generally should *not* be used in benefits analysis, for reasons explained in Text Box 7.3.

9 This is why COI estimates generally understate WTP to reduce the same risk or avoid a given health effect. Some studies have estimated that total WTP can be two to four times as large as COI even for minor acute respiratory illnesses (Alberini and Krupnick 2000). Still, there is not any broadly applicable "scaling factor" that relates COI to WTP generally.

10 EPA analyses have used risk-risk trade-offs for non-fatal cancers in conjunction with VSL estimates as one method to assess the benefits of reduced carcinogens in drinking water (U.S. EPA 2005a).

Text Box 7.3 - Non-Willingness to Pay Measures

Economic measures of value calculate willingness to pay (WTP) for environmental changes. WTP is defined as that amount of money that, if taken away from income, would make an individual exactly indifferent between experiencing an environmental improvement and not experiencing either the improvement or any change in income. (An analogous measure can also be constructed for "not experiencing degradation" rather than "experiencing an improvement"). WTP is a valid measure of "economic value" because it is directly useful for applying the potential compensation tests of Kaldor and Hicks.

Some measures of economic value are not valid, as they do not measure WTP, and cannot be related to changes in utility. Others should be used only in a limited set of circumstances. Some examples are provided below.

Replacement cost. One of the common consequences of environmental deterioration is damage to assets. Some analysts have suggested that the economic value of the damage is the cost of replacing the asset. This will only be true if: (1) damage to the asset is the only cost of the environmental deterioration; and (2) the least expensive way to achieve the level of satisfaction realized before the deterioration would be to replace the asset. If the first condition is not met, consideration of replacement costs alone might underestimate the economic consequences of environmental degradation. If the second condition is not met, replacement costs might overestimate the consequences. Suppose that water pollution kills fish in a pond. Replacing those fish with healthy, edible ones might prove extremely expensive: the pond might need to be dredged and restocked. However, people who are no longer able to catch fish in the pond might be compensated simply by giving them enough money to purchase substitutes at their local supermarket.

Proxy costs. A closely related concept to replacement cost is the cost of a substitute for the damaged asset. In widely cited work, ecologist H.T. Odum (1996) calculated the number of barrels of petroleum that would be required to provide the energy to replace the services of wetland ecosystems. However, this number is economically irrelevant. There is no reason to suppose that people would choose to replace services of damaged wetlands with those of purchased oil. A similar argument can be made against the interpretation of "ecological footprints" as an estimate of economic consequences (Wackernagel and Rees 1996). Dasgupta (2002) interprets these approaches as single-factor theories of value (Karl Marx's labor theory of value is the best known example), fallacies that were disproved in general by Samuelson's (1951) "non-substitution theorem."

Cost-of-illness (COI). Health effects are often proxied by the "cost of illness," which are the total costs of treatment and time lost due to illness. Although COI is discussed in greater detail in Section 7.3.1.5, note here that: (1) COI does not record other expenses incurred in efforts to avoid illness; (2) health insurance may drive a wedge between the costs incurred to treat illness and WTP to avoid it; and (3) COI ignores factors such as discomfort and dread that patients would also be willing to pay to avoid.

Jury awards. Another approach sometimes taken to measure environmental damages is derived from the awards made by juries. Using such awards may also prove problematic for at least two reasons. First, cases only go to trial if both sides prefer the risk of an adverse outcome to the certainty of a pre-trial settlement. Cases that go to juries are "atypical" by definition. Second, since adjudication does not always occur and can never be infallible, jury awards often do, and arguably should (Shavell 1979), embody "punitive" as well as "compensatory" elements. Juries make examples of guilty defendants in an attempt to deter others from committing similar offenses. For this reason, jury awards may overstate typical damages. Finally, jury awards reflect a certain outcome and not the probability of experiencing an adverse event and therefore include the influence of characteristics typically not included in statistical analysis, such as pain, suffering, and likeability. These estimates are not appropriate for application to ex ante evaluation of the value associated with a statistical probability.

Previous studies

A comprehensive summary of existing studies of morbidity values is beyond the scope of these *Guidelines*. Below is a short list of references that can serve as a starting point for reviewing available morbidity value estimates for benefit transfer or for designing a new study. Some recent estimates for particular health effects include Hammitt and Haninger (2007), who examine food-related illnesses, and Chestnut et al. (2006), who examine respiratory and cardiovascular effects. Tolley et al. (1994) and Johanneson (1995) are useful general references for valuing non-fatal health effects. EPA's *Handbook for Non-Cancer Valuation* (U.S. EPA 2000c) provides published estimates for many illnesses and reproductive and developmental effects. Desvousges et al. (1998) assess a number of existing studies in the context of performing a benefit transfer for a benefits analysis of improved air quality. EPA's *Cost of Illness Handbook* (U.S. EPA 2007c) includes estimates for many cancers, developmental illnesses, disabilities, and other conditions. EPA analyses of regulations and policies, including EPA's two comprehensive studies of the benefits and costs of the Clean Air Act (U.S. EPA 1997a and U.S. EPA 1999) draw upon a number of existing studies to obtain values for reductions of a variety of health effects. These sources describe how the central estimates were derived, and attempt to quantify the uncertainty associated with using the estimates.

At least two meta-analyses have attempted to examine how the value of non-fatal risk reductions varies with characteristics of the condition, the affected population, and the approach to valuation. Vassanadumrongdee et al. (2004) focus on air pollution-related morbidity risks and posit a meta-regression based benefit transfer function. Van Houtven et al. (2006) evaluate more than 230 WTP estimates from 17 stated preference studies, finding evidence that illness severity, measured systematically, is a significant factor explaining variation in WTP. The authors also illustrate how a meta-regression-based function can facilitate benefit transfer based on duration and severity of acute illnesses, along with population characteristics. While the specific benefit-transfer functions in these articles might not be suitable for application in any particular context, the estimates contained in them can be helpful. Other studies are available through the Environmental Valuation Reference Inventory (EVRI). EVRI is maintained by Environment Canada and contains more than 1,100 studies that can be referenced according to medium, resource, stressor, method, and country.[11]

Important considerations

The analyst should keep two important considerations in mind when estimating morbidity benefits:

- Characterizing and measuring morbidity effects; and
- Incomplete estimates of WTP.

Characterizing and measuring morbidity effects

The key characteristics that will influence the values of morbidity effects are their severity, frequency, duration, and symptoms. Severity defines the degree of impairment associated with the illness. Examples of how researchers have measured severity include "restricted activity days," "bed disability days," and "lost work days."[12] Severity can also be described in terms of health state indices that combine multiple health dimensions into a single measure.[13] For duration, the primary distinction is between acute effects and chronic effects. Acute effects are discrete episodes usually lasting only a few days, while chronic effects last much longer and are generally associated with long-term illnesses. The

11 See www.evri.ca for more information.

12 As Cropper and Freeman (1991) note, these descriptions are essentially characterizations of a behavioral response to the illness. Lost workdays, for example, in some cases require a decision on an individual's part not to go to work due to illness. Such a response may depend upon various socioeconomic factors as well as the physical effect of the illness.

13 The difference in the indices is intended to reflect the relative difference in disutility associated with symptoms or illnesses. There are serious questions about the theoretic and empirical consistency between these "health-related quality of life" index values and WTP measures for improved health outcomes (Hammitt 2002). Still the inclusion of some aspects of these indices may prove useful in valuation studies (Van Houtven et al. 2006). Examples of economic analyses that have employed some form of health state index include Desvousges et al. (1998) and Magat et al. (1996).

frequency of effects also can vary widely across illnesses. Some effects are one-time events that are unlikely to recur, such as a gastrointestinal illness. Other effects, such as asthma, do recur or can be aggravated regularly, causing disruptions in work, school, or recreational activities.

For chronic conditions or more serious outcomes, morbidity effects are usually measured in terms of the number of expected cases of a particular illness. Given the risks faced by each individual and the number of people exposed to this risk, an estimate of "statistical cases" can be defined analogously to "statistical lives." In contrast, morbidity effects that are considered acute or mild in nature can be estimated as the expected number of times a particular symptom associated with an illness occurs. These estimates of "symptom days" may be used in benefits analysis when appropriate estimates of economic value are available, although a richer characterization of combinations of symptoms, severity, duration, and episode frequency would be an improvement over much of the existing literature. Some studies have attempted to deal with these complexities in a more systematic manner, but the results have not yet been widely applied and interpreted for policy analysis (Cameron and DeShazo 2008). (Refer to Section 7.3.1.5 and Text Box 7.3 on the use of COI versus WTP measures of value.)

Incomplete estimates of WTP

The widespread availability of health insurance and paid sick leave shift some of the costs of illness from individuals to others. While this cost-shifting can be addressed explicitly in COI studies, it may lead to problems in estimating total WTP. If the researcher does not adequately address these concerns, individuals may understate their WTP, assuming that some related costs would be borne by others. However, to the extent that these costs represent diversions from other uses in the economy, they represent real costs to society and should be accounted for in the analysis.

More information on these and other issues to consider when conducting or evaluating morbidity value studies is provided in EPA's *Handbook for Non-Cancer Health Effects Valuation* (U.S. EPA 2000c).

7.2.2 Ecological Benefits

In addition to human health benefits, many EPA policies will produce ecological benefits by increasing the delivery of "ecosystem services," which are the end products of ecological functions that are valued by people (Daily 1997, National Research Council 2005, and Millennium Ecosystem Assessment 2005). There is a large and growing literature on the valuation of ecosystem services. Fisher et al. (2009) document an exponentially increasing number of published articles on ecosystem services, growing from essentially none in the early 1980s to around 250 in 2007. Much of this literature focuses on the impacts of habitat loss and other land use changes on ecosystem service flows. Because EPA has only limited authority over private and public land use decisions, analysts may find that only a subset of the results in these studies will be directly transferable to traditional EPA regulations. Nevertheless, this growing literature can provide a useful conceptual framework and potentially transferable methods for analyzing a wide range of EPA policies that may affect ecological services.

In principle, once the pollutants (or other environmental stressors) whose emissions will be altered by the regulation have been identified, the same general approach used to estimate human health benefits can be used to estimate ecological benefits: identify the endpoints that are affected by those pollutants and that are valuable to society; estimate dose-response relationships between stressors and endpoints; and estimate people's WTP for changes in the endpoints using revealed or stated preference valuation methods. In the case of ecological benefits estimation, the relevant endpoints will include measures of ecosystem health rather than human health, and the methods and data required to estimate the dose-response functions and WTP will differ accordingly. As in the human health case, the estimation of dose-response relationships between pollutants and endpoints will fall mainly to natural scientists, although collaboration between scientists and economists often is needed to help focus the analysis on the most important endpoints. [The Agency's *Ecological Benefits Assessment Strategic Plan* describes an interdisciplinary approach for conducting ecological benefits assessments, as well as research priorities for improving such assessments (U.S. EPA 2006a)]. Even though the basic approach for valuing ecological benefits is similar to that used to value human health benefits, an entirely different set of complications may arise when estimating ecological benefits (Freeman 2003 pp. 457-460). Some of these complications are explored below.

A hypothetical policy

To illustrate some of the complications that can arise when assessing ecological benefits, consider a hypothetical policy that would control the emissions of an industrial chemical that are believed to decrease survival and reproductive rates in one or more fish species. First, compared to the commonly accepted individual-level mortality and morbidity endpoints used in human health benefit assessments, it may be more difficult to identify or define the relevant endpoints in an ecological benefits assessment (de Groot et al. 2002, Boyd and Banzhaf 2007, Wallace 2007, and Fisher and Turner 2008). Identifying endpoints for estimating use values may be relatively straightforward. For example, endpoints for this hypothetical policy would include the abundances and distributions of species that are directly or indirectly affected by the chemical and are harvested or targeted for wildlife viewing or other non-consumptive outdoor activities. Identifying relevant endpoints for non-use values, on the other hand, can be more complicated. Even for this simplified hypothetical policy, it may not be clear which among the wide variety of measureable ecosystem attributes — beyond those previously identified as relevant for use values — would provide an adequate basis for eliciting non-use values in a stated preference survey. Evans et al. (2008) discuss some of the challenges they faced in defining endpoints for a stated preference survey to value reductions in acid rain in the Adirondacks. Boyd and Krupnick (2009) discuss problems of identifying ecological endpoints more generally.

After relevant endpoints are identified, there may be additional complications in modeling the effects of the chemical on those endpoints. For example, the emissions-transport-exposure pathway(s) — i.e., the "ecological production function" (U.S. EPA 2009b) — may involve complex food web linkages that are less direct or have more convoluted feedbacks than in the human health context. Furthermore, some of the important feedbacks may involve human responses to the changed ecological conditions. For example, if some of the fish species in our hypothetical policy scenario are harvested by recreational or commercial fishers, then the nature of the management regime in the fisheries may influence the response of the fish stocks to the policy. In an extreme case, if the commercial fisheries are completely unregulated and subject to open access conditions, then any increases in the stock sizes from the policy may be completely offset in the long run by new entrants to the fishery (Freeman 1991, Barbier et al. 2002, Smith 2007, and Newbold and Iovanna 2007). Therefore, an integrated bio-economic modeling approach may be needed to accurately project the bio-physical effects of the policy. Some examples of such an approach include Smith and Crowder (2006), Massey et al. (2006), and Finnoff and Tschirhart (2008).

After the ecological effects of the policy are characterized, there may be further complications in valuing those effects. For this hypothetical policy, the main requirement for revealed preference valuation methods might be data on commercial and recreational fishing activities in the affected water bodies. Other recreational activities also might be affected, and water-related amenities might influence property values. As with human health benefits, care must be taken to avoid double counting when using multiple datasets and methods that could include overlapping values (McConnell 1990, and Phaneuf et al. 2008). Furthermore, if a significant portion of the benefits for ecological changes are thought to consist of non-use values rather than use values, analysts may need to rely more heavily on stated preference methods when estimating ecological benefits. Considering the challenges in conducting reliable stated preference valuation studies even for well-defined and familiar commodities (described in detail in Section 7.3.2), this compounds the extra complications already discussed. This also points to a larger potential role for non-monetized and non-quantified benefits in the overall analysis (U.S. EPA 2006a, and U.S. EPA 2009b).

Application of economic valuation methods to ecological changes

Extensive treatments of the valuation of ecosystem services can be found in recent reports from the National Academy of Science (NAS) (2005) and EPA's SAB Committee on the Valuation of Ecological Systems and Services (U.S. EPA 2009b). Analysts are referred to these reports for more detailed discussions on the application of economic valuation methods to ecological benefits than are provided in these *Guidelines*. In this section are examples of studies that apply traditional valuation methods (discussed more generally in the following sections of this chapter) to ecosystem goods and services. Some of the special complications that can arise in such studies are highlighted.

Production functions

A number of recent contributions to the literature on valuing of ecosystem services emphasize the importance of understanding the production functions relating natural systems to the provision of products that are valuable to people (Polasky et al. 2008a, 2008b; Boyd and Banzhaf 2007; and U.S. EPA 2009b). Some simple examples have long been known: commercially valuable species "produce" themselves. Early work such as Faustmann's 1848 analysis of optimal rotations in forestry (see also Samuelson 1976), Clark's (1990) work in fisheries, and Hammack and Brown's (1974) work on wetlands and waterfowl have provided templates for later studies. It may be possible to value the effects of pollution on the exploitation of renewable resources when biological production possibilities are affected by environmental conditions — for example, when fish stocks are affected by water quality, or when waterfowl populations are affected by the extent and configuration of wetlands (Bell 1997, Ellis and Fischer 1987, and Massey et al. 2006). As discussed above, analysts should keep in mind that institutional features such as open access to renewable resources may dissipate values that might otherwise be realized from environmental improvement.

Ecological resources also can contribute to the production of other useful goods and services, such as crop yields, groundwater quality, and surface water flow characteristics. Hence the degradation of supporting ecological resources should be reflected in diminished outputs of these commodities. Direct application of production function approaches often is hampered by data and methodological limitations. Specifically, it can be difficult to measure the flow of non-market ecosystem services that a particular production process receives, as well as to statistically control for the effects of unobserved characteristics of climate and topography. One approach is to design observational studies to mimic controlled experiments as closely as possible. Ricketts et al. (2004) use this approach in a study of the value of pollination services to coffee crops. In some cases production functions might plausibly be derived from first principles. For example, Weitzman (1992), Simpson et al. (1996), Rausser and Small (2000), and Costello and Ward (2006) use simple probability models to examine the role of biodiversity in the development of new pharmaceutical products. Further examples of studies relating ecological conditions to economic outputs through production processes include Acharya and Barbier (2002), who examine ground water recharge as a function of surrounding land cover, and Pattanayak and Kramer (2001), who examine stream flow as a function of land cover.

Hedonic models

Econometricians generally have favored estimating cost or profit functions to estimating production functions. This is because the prices that are the arguments of the former will be uncorrelated with unobserved factors, whereas input choices will not (see Varian 1992). While a cost or profit function approach could be adopted in the estimation of ecosystem service values, a more common, and theoretically equivalent, approach is to estimate a hedonic price function. In theory, the rental price of land is equal to the earnings that could be derived from its use, while the purchase price is equal to the net discounted value of the stream of such earnings. A number of authors have estimated hedonic models relating the value of residential properties to the proximity and attributes of nearby forests (Anderson and Cordell 1988, Tyrväinen and Miettinen 2000, and Willis and Garrod 1991), wetlands (Lupi et al. 1991, Mahan et al. 2000, Woodward and Wui 2001, Bin and Polasky 2005, and Costanza et al. 2008), or other varieties of "open space" (Geoghegan et al. 1997, Benson et al. 1998, Irwin and Bockstael 2002, Irwin 2002, and Thorsnes 2002).

Travel cost models

A large number of studies use travel cost models to value ecological endpoints. The predominant activity in the recreational use value literature has been fishing; where the ecological endpoint is expected fish catch (or one or more proxy measures thereof) at one or more recreation sites. For example, 122 of 325 studies in the recreational use value database assembled by Rosenberger and Stanley (2007) focused on either freshwater or saltwater recreational fishing. The remaining studies in the database focus on one of 25 other categories of activities, including bird watching (Hay and McConnell 1979), wildlife hunting (Creel and Loomis 1990, Coyne and Adamowicz 1992, Boxall 1995, Peters et al. 1995, and Adamowicz et al. 1997), beach use (Bockstael et al. 1987a, and Parsons and Massey 2003), backcountry recreation (Boxall et al. 1996), rock-climbing (Shaw and Jakus 1996), and kayaking (Phaneuf and Siderelis 2003).

Stated preference methods

Revealed preference methods cannot capture non-use values, such as those associated with the existence of biological diversity. This is because it is not possible to use data on market transactions or any other observed choices to estimate the value of goods that leave no "behavioral trail" (Larson 1993) in their enjoyment. In such cases only stated preference methods can provide estimates of WTP or WTA (Freeman 2003). More generally, stated preference methods may be employed when researchers want to identify the widest possible spectrum of values, both use and non-use (Loomis et al. 2000).

Stated preference studies have been used to value a number of ecosystem services. Examples include the protection of endangered species (Brown and Shogren 1998), the ecological consequences of water quality improvements in Europe (Hanley et al. 2006), improved ecological conditions resulting from reduced air pollution in the United States (Banzhaf et al. 2006), and restoration of the Florida Everglades (Milon and Scrogin 2006). In some instances researchers may want to combine results of stated preference valuation studies of particular ecological endpoints with other data on the effects of pollution, land use, or other factors on the production of ecosystem services. See Boyd and Krupnick (2009) for an extended discussion.

Complications that may apply to all methods

When using these valuation methods or when transferring the results of previous valuation studies to assess ecological benefits for new policy cases, analysts should be prepared to confront several complications. For example:

For new studies, it may be difficult to identify and/or measure the ecological endpoints that are most relevant for the policy case. Without a set of observable measures of ecological conditions (or measures that can be linked to ecological conditions through supplemental bio-physical modeling) thought to be relevant for outdoor recreation behavior, housing decisions, etc., it will not be possible to use revealed preference methods to value ecological effects. For example, users may care mainly about water clarity for a certain type of recreational activity, while the most readily available data might measure nutrient loading in the water bodies that would be affected by a policy change. Under such circumstances it may be difficult to relate revealed preferences regarding housing decisions, recreational behavior, etc., to the available nutrient loading data, as those data are imperfect proxies for water clarity. There are well-known statistical pitfalls associated both with specifying the wrong "right-hand side" variables in an econometric relationship, as well as with "data mining" by including right-hand side variables in the absence of theoretical justification. The best, if not always practicable, advice that can be given is to think as carefully as possible about which variables should motivate choices before running any regressions.

For benefit transfers, it may be difficult to find existing studies that value ecological endpoints that are the same as, or sufficiently similar to, those of interest in the policy case. This problem is likely to be more common for ecological benefits than for human health benefits because the latter has a larger set of studies to draw from and a smaller set of common endpoints that have been used in multiple studies. The less similar are the commodities valued in the existing ecological benefit studies, the more difficult it will be to synthesize those studies in a meta-analysis or preference calibration exercise, and the less valid will be the transfer of the resulting value estimate or function.

Estimation difficulties are likely to arise in many cases of interest. In particular, explanatory variables may not meet the exogeneity requirement for estimating their associated coefficients. For example, in performing hedonic regressions of property prices on, among other things, the development status of nearby properties, it is likely that both the price of the property in question and the use made of nearby properties would be determined by factors that cannot be observed by the econometrician (Irwin and Bockstael 2002, and Irwin 2002). Similarly, in estimating recreation demand models in which a recreationist's decision to visit a particular site depends on, among other things, congestion (i.e., how many others decide to visit the site at the same time), it is likely that *all* recreationists' site visit choices will be influenced by the same unobserved factors (Timmins and Murdock 2007). Similar difficulties arise in other areas of economics; for example Durlauf's (2004) survey of empirical approaches to "neighborhood effects" in urban economics. The solution in each instance is to identify appropriate instrumental variables, but this can be difficult in many cases. One way around such problems may be to identify "natural experiments." Thorsnes (2002), for example, identifies instances in which historical accidents influenced land use patterns independently of the later realization of adjacent land value in order to conduct a hedonic study of the effects of open space.

For resources subject to consumptive use, such as harvested fish or wildlife species, expected harvest levels are endogenous variables, which can lead to biases similar to that introduced by congestion effects. If the policy of interest leads to spatially heterogeneous environmental quality improvements, then it may lead to a re-sorting not only of recreators but also of the target species among the recreation sites. Ignoring this spatial re-sorting effect can give biased welfare estimates (Newbold and Massey 2010). This can complicate both the estimation of preference parameters and the transfer of the estimated preference function to the policy case.

A basic goal of any benefits assessment is to count all categories of benefits, but to count each only once. This may be particularly important for ecological benefits assessments since stated preference studies employed to estimate intangible values, such as existence values of biodiversity, might also capture use values that are already covered by revealed preference studies such as recreation demand or hedonic studies. When combining values estimated using multiple methods, the analyst should take care to avoid double counting.

It is important to identify and discuss any omitted benefit categories that are thought to be important but that cannot be monetized, or possibly even quantified. There may be circumstances in which provision of some additional information may be helpful, even if does not rise to the level of presenting an explicit comparison of benefits with costs. For example, analysts may be able to identify the most cost-effective approach among different alternatives, or to present natural science information that can convey the biophysical impact of a policy even if it does not quantify the WTP or WTA for such a policy. It is better to acknowledge gaps in information by discussing them qualitatively or by reporting physical measures (if available) than to employ conceptually flawed methods of monetization. In particular, analysts should avoid the use of replacement cost, embodied energy-based evaluation methods, or ecological footprint analysis to derive estimates of WTP or WTA.

7.2.3 Other Benefits

Other types of potential benefits from environmental policies include aesthetic improvements and reduced material damages.

Aesthetic improvements include effects such as improved taste and odor of tap water resulting from water treatment requirements and enhanced visibility resulting from reduced air pollution. EPA typically considers two types of benefits from increased visibility due to improvements in air quality: residential visibility benefits and recreational visibility benefits. Improvements in residential visibility are typically assumed to only benefit residents living in the areas in which the improvements are occurring, while all households in the United States are usually assumed to derive some benefit from improvements in visibility in areas such as National Parks. The benefits received, however, are assumed to decrease with the distance from the recreational area in which the improvements occur.

Reduced materials damages include welfare impacts that arise from changes in the provision of service flows from human-made capital assets such as buildings, roads, and bridges. Materials damages can include changes in both the quantity and quality of such assets. Benefits from reduced material damages typically involve cost savings from reduced maintenance or restoration of soiled or corroded buildings, machinery, or monuments.

Methods and previous studies

Changes in the stock and quality of human-made capital assets are assessed in a manner similar to their "natural capital" counterparts. Analytically, the valuation of reduced materials damages parallels the methods for valuing the tangible end products from managed ecosystems such as agriculture or forestry. Effects from changes in air quality on the provision of the service flows from physical resources are handled in a similar fashion to the effects from changes in air quality on crops or commercial timber stocks. The most common empirical applications involve air pollution damages and the soiling of structures and other property.

Linking changes in environmental quality with the provision of service flows from materials can be difficult because of the limited scientific understanding of the physical effects, the timing of the effects, and the behavioral responses of producers and consumers. An analysis of reduced materials damages typically begins with an environmental fate and transport model to determine the direct effects of the policy on the stocks and flows of pollutants in the environment. Then stressor-response functions are used to relate local concentrations of pollutants to corrosion, soiling, or other physical damages that affect the production (inputs) or consumption (output) of the material service flows. The market response to these impacts serves as the basis for the final stage of the assessment, in which some type of structural or reduced-form economic model that relates averting or mitigating expenditures to pollution levels is used to value the physical impacts. The degree to which behavioral adjustments are considered when measuring the market response is important, and models that incorporate behavioral responses are preferred to those that do not. Adams and Crocker (1991) provide a detailed discussion of this and other features of materials damages benefits assessment. Also see EPA's benefits analysis of household soiling for an example that employs a reduced-form economic model relating defensive expenditures to ambient pollution (U.S. EPA 1997f).

7.3 Economic Valuation Methods for Benefits Analysis

For goods bought and sold in undistorted markets, the market price indicates the marginal social value of an extra unit of the good. There are virtually no markets for environmental goods. While some natural products are sold in private markets, such as trees and fish, these are "products of the environment" and not the types of "environmental goods and services" analysts typically need to value. The analyst's concern is typically with *non-market* inputs, which are, by definition, not traded in markets.[14] To overcome this lack of market data, economists have developed a number of methods to value environmental quality changes. Most of these methods can be broadly categorized as either revealed preference or stated preference methods.

In cases where markets for environmental goods do not exist, WTP can often be inferred from choices people make in related markets. Specifically, because environmental quality is often a characteristic or component of a private good or service, it is sometimes possible to disentangle the value a consumer places on environmental quality from the overall value of a good. Methods that employ this general approach are referred to as *revealed preference methods* because values are estimated using data gathered from observed choices that reveal the preferences of individuals. Revealed preference methods include production or cost functions, travel cost models, hedonic pricing models, and averting behavior models. This section also discusses COI methods, which are sometimes used to value human health effects when estimates of WTP are unavailable.

In situations where no markets for environmental or related goods exist to infer WTP, economists sometimes rely on survey techniques to gather choice data from hypothetical markets. The methods that use this type of data are referred to as *stated preference methods* because they rely on choice data that are stated in response to hypothetical situations, rather than on choice behavior observed in actual markets. Stated preference methods include contingent valuation, conjoint analysis, and contingent ranking.

Each of these revealed and stated preference methods is discussed in detail below. Included are an overview of each method, a description of its general application to environmental benefits analysis, and a discussion of issues involved in interpreting and understanding valuation studies. The discussion concludes with a separate overview of benefit-transfer methods. It is important to keep in mind that research on all of these methods is ongoing. The limitations and qualifications described here are meant to characterize the state of the science at the time these *Guidelines* were written. Analysts should consult additional resources as they become available.

7.3.1 Revealed Preference Methods

A variety of revealed preference methods for valuing environmental changes have been developed and are widely used by economists. The following common types of revealed preference methods are discussed in this section:

- Production or cost functions;
- Travel cost models;
- Hedonic models;
- Averting behavior models; and
- Cost of Illness (COI).[15]

7.3.1.1 Production and Cost Functions

Discrete changes in environmental circumstances generally cause both consumer and producer effects, and it is common practice to separate the welfare effects brought about by changes in environmental circumstances into consumer surplus and producer surplus.[16] Marginal changes can be evaluated by considering the production side of the market alone.

14 There are examples in which environmental goods have been traded in markets. The Clean Air Act Amendments of 1990, for example, initiated a market in sulfur dioxide (SO_2). However prices in such markets are determined by regulation-induced scarcity, and not by considerations of marginal utilities or marginal products.

15 Although not a revealed preference method (as it does not measure WTP) COI methods are discussed in this section since estimates are based on observable data.

16 See Appendix A for more detail.

Economic foundations of production and cost functions

Inputs to production contribute to welfare indirectly. The marginal contribution of a productive input is calculated by multiplying the marginal product of the input by the marginal utility obtained from the consumption good, in whose production the input is employed. The marginal utility of a consumption good is recorded in its price. While marginal products are rarely observed, the need to observe them is obviated when both inputs and outputs are sold in private markets because *prices* can be observed. Environmental goods and services are typically not traded in private markets, and therefore the values of environmental inputs must be estimated indirectly.

Production possibilities can be represented in three equivalent ways:

- As a production function relating output to inputs;
- As a cost function relating production expenses to output and to input prices; and
- As a profit function relating earnings to the prices of both output and inputs (see Varian 1992, for an explication of the relationships among these functions).

The value of a marginal change in some environmental condition can be represented as a marginal change in the value of production, as a marginal change in the cost of production, or as a marginal change in the profitability of production.[17] It should be noted, however, that problems of data availability and reliability often arise. Such problems may motivate the choice among these conceptually equivalent approaches, or in favor of another approach.

Note that derivation of values *on the margin* does not require any detailed understanding of consumer demand conditions. To evaluate marginal effects via the production function approach, the analyst needs to know the price of output and the marginal product of the environmental input. To derive the equivalent measure using a cost function approach, the analyst needs to know the derivative of the cost function with respect to the environmental input. In the profit function approach, the analyst needs to know the derivative of the profit function with respect to the environmental input.[18]

In the statements note the emphasis that *marginal* effects are being estimated. Estimating the net benefits of larger, non-marginal, changes represent a greater challenge to the analyst. In general this requires consideration of changes in both producer and consumer surplus. The latter necessitates application of techniques such as travel cost, hedonics, and stated preference, which are discussed elsewhere in this chapter.

Before moving on to those topics, note a fourth equivalent way to estimate environmental effects on production possibilities. Such effects are reflected in the profitability of enterprises engaged in production. That profitability also can be related to the return on fixed assets such as land. The value of a parcel of land is related to the stream of earnings that can be achieved by employing it in its "highest and best use." Its rental value is equal to the profits that can be earned from it over the period of rental (the terms "rent" and "profit" are often used synonymously in economics). The purchase price of the land parcel is equal to the expected discounted present value of the stream of earnings that can be realized from its use over time. Therefore, the production, cost, and profit function approaches described above are also equivalent to inferences drawn from the effects of environmental conditions on asset values. This fourth approach is known as "hedonic pricing," and will be discussed in detail in Section 7.3.1.3.

17 For a good review of statistical procedures used for estimating production, cost, and profit functions see Berndt (1991).

18 Derivation of marginal values often involves an application of the "envelope theorem" that states that effects from variables that are already optimized are negligible. In determining the effect of an improvement in a particular environmental input on welfare arising from the consumption of a particular product using the cost function approach, the analyst would determine how $\int_0^Q p(q)dq - C(Q, e)$ varies with e, the environmental variable. The integral is consumer surplus, i.e., the area under the demand curve, and the second term is the cost of producing quantity Q given environmental conditions, e. Differentiating with respect to e yields $[p(Q) - \partial C/\partial Q]\, dQ/de - \partial C/\partial e = - \partial C/\partial e$, where the last equality results because competitive firms set price equal to marginal cost, i.e., $p(Q) = \partial C/\partial Q$. This is the basis for the general proposition that marginal values can be estimated by looking solely at the production side of the market.

It is introduced now to show that production, cost, or profit function approaches are generally equivalent to hedonic approaches.

"Production" as a term is broad in meaning and application, especially with regard to hedonic pricing. While businesses produce goods and services in their industrial facilities, one might also say that developers "produce" housing services when they build residences. Therefore, hedonic pricing approaches can measure the value of the environment in "production," whether they are focusing on commercial or residential properties. Similarly, households may "produce" their health status by combining inputs such as air and water filtration systems and medical services along with whatever environmental circumstances they face. Or they "produce" recreational opportunities by combining "travel services" from private vehicles, their own time, recreational equipment purchases, and the attributes of their destination. Much of what is discussed elsewhere in this section is associated with this "production" analysis. This is not to say that estimation of production, cost, or profit functions is necessarily the best way to approach such problems, but rather, that all of these approaches are conceptually consistent.

General application of production and cost functions

Empirical applications of production and cost function approaches are diverse. Among other topics, the empirical literature has addressed the effects of air quality changes on agriculture and commercial timber industries. It also has assessed the effects of water quality changes on water supply treatment costs and on the production costs of industry processors, irrigation operations, and commercial fisheries.[19] Production, cost, or profit functions have found interesting applications to the estimation of some ecological benefits.[20] Probabilistic models of new product discovery from among diverse collections of natural organisms can also be regarded as a type of "production."[21] Finally, work in ecology points to "productive" relationships among natural systems that may yield insights to economists as well.[22]

Considerations in evaluating and understanding production and cost functions

The analyst should consider the following factors when estimating the values of environmental inputs into production:

Data requirements and implications. Estimating production, cost, or profit functions requires data on *all* inputs and/or their prices. Omitted variable bias is likely to arise absent such information, and may motivate the choice of one form over another. Econometricians have typically preferred to estimate cost or, better yet, profit functions. Data on prices are often more complete than are data on quantities and prices are typically uncorrelated to unobserved conditions of production, whereas input quantities are not.

The model for estimation. Standard practice involves the estimation of "flexible functional forms," i.e., functions that can be regarded as second-order approximations to any production technology. The translog and generalized Leontief specifications are examples. Estimation often will be more efficient if a system of equations is estimated (e.g., simultaneous estimation of a cost function and its associated factor demand equations), although data limitations may impose constraints.

Market imperfections. Most analyses assume perfectly competitive behavior on the part of producers and input suppliers, and assume an absence of other distortions. When these assumptions do not hold, the interpretation of welfare results becomes more problematic. While there is an extensive literature on the regulation of externalities under imperfect competition, originating with Buchanan (1969), analysts should exercise caution and restraint in attempting to correct for departures from competitive behavior.

19 Refer to Adams et al. (1986), Kopp and Krupnick (1987), Ellis and Fisher (1987), Taylor (1993), and U.S. EPA (1997a) for examples.

20 See, for example, Acharya and Barbier (2002) on groundwater recharge, and Pattanayak and Kramer (2001) on water supply.

21 For example, see Weitzman (1992), Simpson et al. (1996), and Rausser and Small (2000).

22 For example, see Tilman, Lehman, and Polasky 2005.

The issues can become quite complex and, as is the case with environmental externalities, there is typically no direct evidence of the magnitude of departures from perfectly competitive behavior. Moreover, in many circumstances it might reasonably be argued that departures from perfect competition are not of much practical concern (Oates and Strassman 1984). Perhaps a more pressing concern in many instances will be the wedge between private and social welfare consequences that arise with taxation. An increase in the value of production occasioned by environmental improvement typically will be split between private producers and the general public through tax collection. The issues here also can become quite complex (see Parry et al. 1997), with interactions among taxes leading to sometimes surprising implications. While it is difficult to give general advice, analysts may wish to alert policy makers to the possibility that the benefits of environmental improvements in production may accrue to different constituencies.

7.3.1.2 Travel Costs

Recreational values constitute a potentially large class of environmental use benefits. However, measuring these values is complicated by the fact that the full benefits of access to recreation activities are rarely reflected in admission prices. Travel cost models address this problem by inferring the value of changes in environmental quality through observing the trade-offs recreators make between environmental quality and travel costs. A common situation recreators may face is choosing between visiting a nearby lake with low water quality and a more distant lake with high water quality. The outcome of the decision of whether to incur the additional travel cost to visit the lake with higher water quality reveals information about the recreator's value for water quality. Travel cost models are often referred to as recreation demand models because they are most often used to value the availability or quality of recreational opportunities.

Economic foundation of travel cost models

Travel cost models of recreation demand focus on the choice of the number of trips to a given site or set of sites that a traveler makes for recreational purposes. Because there is no explicit market or price for recreation trips, travel cost models are frequently based on the assumption that the "price" of a recreational trip is equal to the cost of traveling to and from the site. These costs include both participants' monetary cost and opportunity cost of time. Monetary costs include all travel expenses. For example, when modeling day trips taken primarily in private automobiles, travel expenses would include roundtrip travel distance in miles multiplied by an estimate of the average cost per mile of operating a vehicle, plus any tolls, parking, and admission fees.

A participant's opportunity cost of time for a recreational day trip is the value of the participant's time spent traveling to and from the recreation site plus the time spent recreating.[23] A variety of approaches have been used in the literature to define the opportunity cost of time. Most commonly, researchers have used a fixed fraction ranging from one third to one whole of a person's hourly wage as an estimate of participants' hourly opportunity cost of time. In most cases, the fraction used depends on how freely individuals are assumed to be able to substitute labor and leisure. If a person can freely choose their work hours then their opportunity cost of time will be equal to their full wage rate. However, if a person cannot freely substitute labor for leisure (for example if they have a set 40 hour work week), then the opportunity cost of the time they have available for recreation is unobservable and may be less or more than the full wage rate. Many other factors can influence recreators' opportunity cost of time, including the utility received from traveling, non-wage income, and other non-work time constraints. A number of researchers have developed methods for estimating recreators' endogenous opportunity cost of time although no one method has yet been fully embraced in the literature. For examples, see McConnell and Strand (1981); Smith,

23 If the amount of time spent recreating or doing something else (not including the time spent traveling to and from the sites) is assumed to be the same across all alternatives then it will not be identifiable in estimation and therefore it is not necessary to include it in the estimation of the participant's opportunity cost of time. See Smith, Desvousges, and McGivney (1983); and McConnell (1992) for discussions of the implication of and the methods for allowing time onsite to vary across trip and alternatives.

Desvousges, and McGivney (1983); Bockstael et al. (1987b); McConnell (1992); and Feather and Shaw (1999). Hourly opportunity costs are multiplied by round trip travel time and time on-site to calculate a person's full opportunity cost of time. Total travel costs are the sum of monetary travel costs and full opportunity costs. Following the law of demand, as the cost of a trip increases the quantity of trips demanded generally falls, all else equal. This means that participants are more likely to visit a closer site than a site farther away.

While travel costs are the driving force of the model, they do not completely determine a participant's choice of sites to visit. Site characteristics, such as parking, restrooms, or boat ramps; participant characteristics, such as age, income, experience, and work status; and environmental quality also can affect demand for sites. The identification and specification of the appropriate site and participant characteristics are generally determined by a combination of data availability, statistical tests, and the researcher's best judgment. Ultimately, every recreation demand study strikes a compromise in defining sites and choice sets, balancing data needs and availability, costs, and time.[24]

General application by type of travel cost model

Travel cost models can logically be divided into two groups: single-site models and multiple-site models. Apart from the number of sites they address, the two types of models differ in several ways. The basic features of both model types are discussed below.

Single-site models. Single-site travel cost models examine recreators' choice of *how many trips to make to a specific site over a fixed period of time* (generally a season or year). It is expected that the number of trips taken will increase as the cost of visiting the site decreases and/or as the benefits realized from visiting increase. Site, participant, and environmental attributes, as well as the prices of substitute sites, act as demand curve shifters. For example, sites with good water quality are likely to be visited more often than sites with poor water quality, all else equal. Most current single-site travel cost models are estimated using count data models because the dependent variable (number of trips taken to a site) is a non-negative integer. See Haab and McConnell (2003) and Parsons (2003a) for detailed discussions and examples of recreation demand count data models.

Single-site models are most commonly used to estimate the value of a change in access to a site, particularly site closures (e.g., the closure of a lake due to unhealthy water quality). The lost access value due to a site closure is the difference between the participant's WTP for the option of visiting the site, which is given by the area between the site's estimated demand curve and the implicit "price" paid to visit it. Estimating the value of a change in the cost of a site visit, for example the addition or increase of an admission fee, is another common application of the model.

A weakness of the single-site model is its inability to deal with large numbers of substitute sites. If, as is often the case, a policy affects several recreation sites in a region, then traditional single-site models are required for each site. In cases with large numbers of sites, defining the appropriate substitute sites for each participant and estimating individual models for each site can impose overwhelming data collection and computational costs. Because of these difficulties, most researchers have opted to refrain from using single-site models when examining situations with large numbers of substitute sites.[25]

Multiple-site models. Multiple-site models examine a recreator's choice of *which site to visit from a set of available site (known as the choice set) on a given choice occasion* and in some cases can also examine *how many trips to make to each specific site*

24 For a comprehensive treatment of the theoretical and econometric properties of recreation demand models see Phaneuf and Smith (2005).

25 Researchers have developed methods to extend the single-site travel cost model to multiple sites. These variations usually involve estimating a system of demand equations. One example is the Kuhn-Tucker (KT) model discussed in the following multiple-site model section. See Bockstael, McConnell, and Strand (1991) and Shonkwiler (1999) for more discussion and other examples of extensions of the single-site model.

over a fixed period of time. Compared to the single-site model, the strength of multiple-site models lies in their ability to account for the availability and characteristics of substitute sites. By examining how recreators trade the differing levels of each site characteristic and travel costs when choosing among sites it is possible to place a per trip (or choice occasion) dollar value on site attributes or on site availability for single sites or multiple sites simultaneously.

The two most common multiple-site models are the random utility maximization (RUM) travel cost model and Kuhn-Tucker (KT) system of demand models. Both models may be described by a similar utility theoretic foundation, but they differ in important ways. In particular, the RUM model is a choice occasion model while the KT model is a model of seasonal demand.

Random utility maximization models. In a RUM model each alternative in the recreator's choice set is assumed to provide the recreator with a given level of utility, and on any given choice occasion the recreator is assumed to choose the alternative that provides the highest level of utility on that choice occasion.[26] The attributes of each of the available alternatives, such as the amenities available, environmental quality, and the travel costs, are assumed to affect the utility of choosing each alternative. Because people generally do not choose to recreate at every opportunity, a non-participation option is often included as a potential alternative.[27] From the researcher's perspective, the observable components of utility enter the recreator's assumed utility function. The unobservable portions of utility are captured by an error term whose assumed distribution gives rise to different model structures. Assuming that error terms have type 1 extreme values distribution leads to the closed form logit probability expression and allows for maximum likelihood estimation of utility function parameters. Using these estimated parameters it is then possible to estimate WTP for a given change in sites quality or availability.

However, because the RUM model examines recreation decisions on a choice occasion level, it is less suited for predicting the number of trips over a time period and measuring seasonal welfare changes. A number of approaches have been used to link the RUM model's estimates of values per choice occasion to estimates of seasonal participation rates. See Parsons, Jakus, and Tomasi (1999) for a detailed discussion of methods of incorporating seasonal participation estimates into the RUM framework.

The nested logit and mixed logit models are extensions of the basic logit. The nested logit model groups similar alternatives into nests where alternatives within a nest are more similar with each other than they are with alternatives outside of the nest. In very general terms, recreators are first assumed to choose a nest and then, conditional on the choice of nest, they then choose an alternative within that nest. Nesting similar alternatives allows for more realistic substitution patterns among sites than is possible with a basic logit. The mixed logit is a random parameter logit model that allows for even more flexible substitution patterns by estimating the variation in preferences (or correlation in errors) across the sample. If preferences do not vary across the sample then the mixed logit collapses to a basic logit.[28]

The Kuhn-Tucker (KT) model. The KT model is a seasonal demand model that estimates recreators' *choice of which sites to visit (like a multiple-site model) and how often to visit them over a season (like a single-site model).* The model is built on the theory that people maximize their seasonal utility subject to their budget constraint by purchasing

26 While the standard logit recreation demand model treats each choice occasion as an independent event, the model can also be generalized to account for repeated choices by an individual.

27 In a standard nested logit RUM model, recreators are commonly assumed to first decide whether or not to take a trip, and then conditional on taking a trip, to next choose which site to visit. By not including a non-participation option, the researcher in effect assumes that the recreator has already decided to take a trip, or in other words, that the utility of taking a trip is higher than the utility of doing something else for that choice occasion. Another way to think of it is that models lacking a participation decision only estimate the recreation values of the segment of the population that participates in recreation activities (i.e., recreators), while models that allow for non-participation incorporate the recreation values of the whole population (i.e., recreators and non-recreators combined). Because of this, recreation demand models without participation decisions tend to predict larger per person welfare changes than models allowing non-participation.

28 See Train (1998) and Train (2003) for detailed descriptions of the nested and mixed logit models.

the quantities of recreation and other goods that give them the greatest overall utility. Similar to the RUM model, the researcher begins by specifying the recreator's utility function. Taking the derivative of this utility function with respect to the number of trips taken, subject to a budget and non-negative trip constraint, yields the "Kuhn-Tucker" conditions. The KT conditions show that trips will be purchased up to the point that the marginal rate of substitution between trips and other spending is equal to the ratio of their prices. In cases where the price of a good exceeds its marginal value none will be purchased. Given assumptions on the form of the utility function and the distribution of the error term, probability expressions can be derived and parameter estimates may then be recovered. While recent applications have shown that the KT model is capable of accommodating a large number of substitute sites (von Haefen, Phaneuf, and Parsons 2004) the model is computationally intensive compared to traditional models. For a basic application of the KT model see Phaneuf and Siderelis (2003). For more advanced treatments of the models see Phaneuf, Kling, and Herriges (2000), and von Haefen and Phaneuf (2005).

Considerations in evaluating and understanding recreation demand studies

Definition of a site and the choice set. The definition of what constitutes a unique site has been shown to have a significant effect on estimation results. Ideally, one could estimate a recreation demand model in which sites are defined as specific points such as exact fishing location, campsites, etc. The more exact the site definition, the more exact the measure of travel costs and site attributes, and therefore WTP, that can be calculated. However, in situations with a large number of potential alternatives, the large data requirements may be cost and time prohibitive, estimation may be problematic, and aggregation may be required. The method of aggregation has been shown to have a significant effect on estimated values. The direction of the effect will depend on the situation being evaluated and the method of aggregation chosen (Parsons and Needleman 1992; Feather 1994; Kaoru, Smith, and Liu 1995; and Parsons, Plantinga, and Boyle 2000).

In addition to the definition of what constitutes a site, the number of sites included in a recreator's choice set can have a significant effect on estimated values. When defining choice sets, the most common practice in the literature has been to include all possible alternatives available to the recreator. In many cases availability has been defined by location with a given distance or travel time.[29] This strategy has been criticized on the grounds that people may not know about all possible sites, or even if they do know they exist they may not seriously consider them as alternatives. In response to this, a number of researchers have suggested methods that either restrict choice sets to include only those sites that the recreators seriously consider visiting (Peters et al. 1995, and Haab and Hicks 1997) or that weight seriously-considered alternatives more heavily than less-seriously-considered alternatives (Parsons, Massey, and Tomasi 2000).

Multiple-site or multipurpose trips. Recreation demand models assume that the particular recreation activity being studied is the sole purpose for a given trip. If a trip has more than one purpose, it almost certainly violates the travel cost model's central assumption that the "price" of a visit is equal to the travel cost. The common strategy for dealing with multipurpose trips is simply to exclude them from the data used in estimation.[30] See Mendelsohn et al. (1992) and Parsons (2003b) for further discussion.

Day trips versus multi-day trips. The recreation demand literature has focused almost exclusively on single-day trip recreation choices. One main reason researchers have focused mostly on day trips is that adding the option to stay longer than one day adds another choice variable in estimation,

29 Parsons and Hauber (1998) explore the implication of this strategy by expanding the choice set geographically and find that beyond some threshold the effect of additional sites is negligible.

30 Excluding any type or class of trip (like multiple-site or multipurpose) will produce an underestimate of the population's total use value of a site. The amount by which benefits will be underestimated will depend on the number and type of trips excluded.

thereby greatly increasing estimation difficulty. A second reason is that as trip length increases multipurpose trips become increasingly more likely, again casting doubt on the assumption that trip's travel costs represent the "price" of one single activity (see previous paragraph). A few researchers have estimated models that allow for varying trip length. The most common strategy has been to estimate a nested logit model in which each choice nest represents a different trip length option. See Kaoru (1995) and Shaw and Ozog (1999) for examples. The few multi-day trip models in the literature find that the *per-day* value of multi-day trips is generally less than the value of a single-day trip, which suggests that estimating the value of multi-day trips by multiplying a value estimated for single-day trips value by the number of days of will overestimate the multi-day trip value.

7.3.1.3 Hedonics

Hedonic pricing models use statistical methods to measure the contribution of a good's characteristics to its price. Cars differ in size, shape, power, passenger capacity, and other features. Houses differ in size, layout, and location. Even labor hours can be thought of as "goods" differing in attributes like risk levels, and supervisory nature, that should be reflected in wages. Hedonic pricing models use variations in property prices or wages and are commonly used to value the characteristics of properties or jobs. The models are based on the assumption that heterogeneous goods and services (e.g., houses or labor) consist of "bundles" of attributes (e.g., size, location, environmental quality, or risk) that are differentiated from each other by the quantity and quality of these attributes. Environmental conditions are among the many attributes that differ across neighborhoods and job locations.

Economic foundations of hedonic models

Hedonic pricing studies estimate economic benefits by weighing the advantages against the costs of different choices. A standard assumption underlying hedonic pricing models is that markets are in equilibrium, which means that no individual can improve her welfare by choosing a different home or job. For example, if an individual changed location she might move to a larger house, or one in the midst of a cleaner environment. However, to receive such amenities, the individual must pay for a more expensive house and incur transaction costs to move. The more the individual spends on her house, the less she has to spend on food, clothing, transportation, and all the other things she wants or needs. Thus, individuals are assumed to choose a better available option such that the benefits derived from it are exactly offset by the increased cost. So, if the difference in prices paid to live in a cleaner neighborhood is observable, then that price difference can be interpreted as the WTP for a better environment.

One key requirement in conducting a hedonic pricing study is that the available options differ in measurable ways. To see why, suppose that all locations in a city's housing market were polluted to the same degree, or all jobs in a particular labor market expose workers to the same risks. Homeowners and workers would, of course, be worse off due to their exposure to pollution and job risks, but their losses could not be measured unless a comparison could be made to purchasers of more expensive houses in less polluted neighborhoods, or wages in lower-paying but safer jobs. However, there is also a practical limit on the heterogeneity of the sample. Workers in different countries earn very different wages and face very different job risks, but this does not mean it is possible to value the difference in job risks by reference to international differences in wages. This is because: (1) there are many other factors that differ between widely separated markets; and (2) people simply are not mobile between very disparate sites. For these reasons it is important to exercise care in defining the market in which choices are made.[31]

Another aspect of the heterogeneity in locations required to make hedonic pricing studies work is that people must *be able to perceive* the differences among their options. If homeowners are unable to recognize differences in health outcomes, visibility, and other consequences of differences

31 Michaels and Smith (1990) offer guidance for defining the extent of the market.

in air quality at different locations, or if workers are unaware of differences in risks at different jobs, then a hedonic pricing study would not be suitable for estimating the values for those attributes.

Hedonic pricing studies can be used in different ways in environmental economics. Some are intended to provide direct evidence of the value of environmental improvements. Hedonic housing price studies are good examples. House prices are related to environmental conditions. The most frequent example is probably air quality (see Smith and Huang 1995 for a meta-analysis of many studies), although water quality (Leggett and Bockstael 2000), natural amenities (Thorsnes 2002), land contamination (Messer et al. 2006) and other examples have been studied. Other hedonic studies evaluate endpoints other than environmental conditions. A good example would be hedonic wage studies that are used in the computation of the VSL. (See Viscusi 2004 for a recent example.)

General application by type of hedonic pricing study

Hedonic wage studies, also known as wage-risk or compensating wage studies, are based on the premise that individuals make trade-offs between wages and occupational risks of death or injury. Most analysts assume that workers understand on-the-job risks, but others argue that workers generally underestimate them (Viscusi 1993). Some studies attempt to account for workers' perceived risks, but the results of these studies are not markedly different from those that do not (Gerking, de Haan, and Schulze 1988). Two of the most frequently used data sources for hedonic wage studies are the National Institute of Occupational Safety and Health (NIOSH) and Bureau of Labor Statistics (BLS) Survey on Working Conditions (SWC) data. The NIOSH data are state-level data of fatalities by occupation or industry, while the SWC data provide a finer resolution of occupation or industry fatalities, but do not vary by location. Black and Kneiser (2003), however, question the ability of hedonic wage studies using these data sources to measure job risks accurately due to severe measurement error. They find that the measurement error in the fatality rates reported from these sources is correlated with covariates commonly used in the wage equations, making the consistent estimation of the coefficient on risk in the standard hedonic wage equation a challenge. More recent hedonic wage studies have used the BLS Census of Fatal Occupational Injuries (CFOI) as the source for workplace risk information (Viscusi 2004; Viscusi and Aldy 2007b; Aldy and Viscusi 2008; Kniesner, Viscusi, and Ziliak 2006; Leeth and Ruser 2003; Viscusi 2003; and Scotton and Taylor 2009). These data are considered the most comprehensive data on workplace fatalities available (Viscusi 2004), compiling detailed information since 1992 from all states and the District of Columbia. Not only are the counts of fatal events reported by 3-digit occupation and 4-digit industry classifications, but the circumstances of the fatal events, as well as worker characteristics like age, gender and race, are also captured.[32] To ensure the veracity and completeness of the reported data, multiple sources, including death certificates, workers' compensation reports and federal and state administration reports are consulted and cross-referenced.

Although questions still persist about the applicability of hedonic wage study results to environmental benefits assessment, hedonic wage studies have been used most frequently in benefits assessments to estimate the value of fatal risk reductions.[33] When a benefits assessment requires a VSL estimate, hedonic wage estimates are a good source of information. Historically, EPA has used a VSL estimate primarily derived from hedonic wage studies. For more information on the Agency's VSL estimate, see Section 7.1.1 and Appendix C.[34] The VSL determined by a hedonic wage study, for example, typically relates WTA higher wages in exchange for the increased likelihood of accidental death during a person's working years. However,

32 More information on the CFOI data is available at: http://www.bls.gov/iif/oshfat1.htm.

33 For example, EPA's SAB has recognized the limitations of these estimates for use in estimating the benefits of reduced cancer incidence from environmental exposure. Despite these limitations, however, the SAB concluded that these estimates were the best available at the time (U.S. EPA 2000d).

34 As part of the revision of this document, EPA is revisiting the VSL estimate used in policy analysis; further guidance will be forthcoming.

analysts should take care when applying results from one hedonic study to a new policy case, for example, if there are differences in the age groups facing mortality risks from longer-term conditions.

Hedonic property value studies measure the different contributions of various characteristics to the value of property. These studies are typically conducted using residential housing data, but they have also been applied to commercial and industrial property, agricultural land, and vacant land.[35] Bartik (1988) and Palmquist (1988, 1991) provide detailed discussions of benefits assessment using hedonic methods. Property value studies require large amounts of disaggregated data. To avoid aggregation problems, market transaction prices on individual parcels or housing units are preferred to aggregate data such as census tract information on average housing units. Problems can arise from errors in measuring prices (aggregated data) and errors in measuring product characteristics (particularly those related to the neighborhood and the environment). There are numerous statistical issues associated with applying hedonic methods to property value studies. These include the choice of functional form, the definition of the extent of the market, identification, endogeneity, and spatial correlation. Refer to Palmquist (1991) for a thorough treatment of the main econometric issues. Recently, advances have been made in modeling spatial correlation in hedonic models (see Text Box 7.4 on spatial correlation for more information).

Other hedonic studies. Applicability of the hedonic pricing method is not limited to the property and labor markets. For example, hedonic pricing methods can be combined with travel cost methods to examine the implicit price of recreation site characteristics (Brown and Mendelsohn 1984). Results from other studies can be used to infer the value of reductions in mortality, cancer, or injury risks. For example, Dreyfus and Viscusi (1995) use a hedonic analysis to determine the trade-offs between automobile price and safety features to infer the VSL.

Considerations in evaluating and understanding hedonic pricing studies

Unobservable factors. A concern common to hedonic pricing studies is that it is impossible to observe all factors that go into a decision. People will choose among different jobs or houses not only because they can trade off differences in amenities and risks against differences in prices or wages, but also because they have different preferences for risks. Idiosyncratic personal tastes that cannot be observed may be responsible for a substantial portion of differences in observed choices. For example, mountain climbers have been known to pay tens of thousands of dollars to undertake expeditions that substantially increase their likelihood of early death.

Source of risks. Similarly, analysts need to be careful in distinguishing the source of the risks used to estimate risk premia. Consider an individual who both works a dangerous job and lives in unhealthy circumstances. Such a person may be at greater risk of premature death than someone who works a different job or lives elsewhere. Analysts risk underestimating the wage premium demanded on the job if they fail to distinguish between causes of death — for example between on-the-job accidents and environmentally induced conditions acquired at home — when relating the wage premium paid on dangerous jobs to the statistics on premature mortality. Conversely, if the same job poses multiple risks — say the risk of both accidental death and serious, but nonfatal injury were higher on a particular job — the wage premium the job offers would overstate WTP for reductions in mortality risks if the injury risks were not properly controlled for in the analysis. See Eeckhoudt and Hammitt (2001), and Evans and Smith (2006) for more discussion of competing versus specific risks.

Marginal changes. As with many results in economics, hedonic pricing models are best suited to the valuation of small, or marginal, changes in attributes. Under such circumstances, the slope

35 See Xu, Mittlehammer, and Barkley (1993), and Palmquist and Danielson (1989) for hedonic values of agricultural land; Ihlanfeldt and Taylor (2004) for commercial property; Dale, Murdoch, Thayer, and Waddell (1999), and McCluskey and Rausser (2003) for residential property; and Clapp (1990), and Thorsnes (2002) for vacant land.

Text Box 7.4 - Spatial Correlation

Real property, such as buildings and land, and their associated characteristics are spatially distributed over the landscape. As such, the characteristics of some of the properties may be spatially correlated. If some of these characteristics are unobserved or for any other reason are not incorporated into the econometric model, there may be dependence across the error terms of the model. Spatial econometrics is a subfield of econometrics that has gained more attention as the capability for assessing such locational relationships within hedonic property data has improved. Such improvements are primarily due to the increasing use of geographic information systems (GIS) technology and geographically referenced data sets.

The nature of the correlation in the data can manifest itself so that there is either spatial heterogeneity across observations, or more importantly, so that the characteristic values (e.g., price of homes) are correlated with those of nearby observations. Standard econometric techniques can readily deal with the former, but are not well equipped to handle the latter case. The econometric techniques allow for testing for the presence of spatial correlation, and specifically modeling and correcting the correlation between observations and correcting for the biasing effect it can have on parameter estimates. In practice, a relationship is defined between every variable at a given location and the same variable at other, usually nearby, locations in the data set. In most cases this relationship is based on common boundaries or is some specified function based on the distances between observations. This relationship between observations is then accounted for in the econometric model in order to correct the error terms and obtain unbiased model estimates. For more details on the fundamentals of spatial statistics see Anselin (1988).

of the hedonic price function can be interpreted as WTP for a small change in the attribute. Public policy, however, is sometimes geared to larger, discrete changes in attributes. When this is the case, calculation of benefits can become significantly more complicated. Hedonic price functions typically reflect equilibria between consumer demands and producer supplies for fixed levels of the attributes being evaluated. The demand and supply functions are tangent to the hedonic price function only in the immediate neighborhood of an equilibrium point. Palmquist (1991) describes conditions under which exact welfare measures can be calculated for discrete changes. See Freeman (2003) and Ekeland, Heckman, and Nesheim (2004) for recent treatments.

7.3.1.4 Averting Behaviors

The averting behavior method infers values for environmental quality from observations of actions people take to avoid or mitigate the increased health risks or other undesirable consequences of reductions in ambient environmental quality conditions. Examples of such defensive actions can include the purchase and use of air filters, boiling water prior to drinking it, and the purchase of preventative medical care or treatment. By analyzing the expenditures associated with these averting behaviors economists can attempt to estimate the value individuals place on small changes in risk (Shogren and Crocker 1991, and Quiggin 1992).

Economic foundations of averting behavior methods

Averting behavior methods can be best understood from the perspective of a household production framework. Households can be thought of as producing health outcomes by combining an exogenous level of environmental quality with inputs such as purchases of goods that involve protection against health and safety risks (Freeman 2003). To the extent that averting behaviors are available, the model assumes that a person will continue to take protective action as long as the expected benefit exceeds the cost of doing so. If there is a continuous relationship between defensive actions and reductions in health risks, then the individual will continue to avert until the marginal cost just equals her marginal WTP for these reductions. Thus, the value of a small change in health risks can be estimated from two primary pieces of information:

- The cost of the averting behavior or good; and
- Its effectiveness, as perceived by the individual, in offsetting the loss in environmental quality.

Blomquist (2004) provides a detailed description of the basic household production model of averting behavior. More detail on the difficulties inherent in applying the averting behavior model can be found in Cropper and Freeman (1991).

One approach to estimation is to use observable expenditures on averting and mitigating activities to generate values that may be interpreted as a lower bound on WTP. Harrington and Portney (1987) demonstrate this by showing that WTP for small changes in environmental quality can be expressed as the sum of the values of four components: changes in averting expenditures, changes in mitigating expenditures, lost time, and the loss of utility from pain and suffering. The first three terms of this expression are observable, in principle, and can be approximated by calculating changes in these costs after a change in environmental quality. The resulting estimate can be interpreted as a lower bound on WTP that may be used in benefits analysis (Shogren and Crocker 1991, and Quiggin 1992).

General application of averting behavior method

Although the first applications of the method were directed toward values for benefits of reduced soiling of materials from environmental quality changes (Harford 1984), recent research has primarily focused on health risk changes. Conceptually, the averting behavior method can provide WTP estimates for a variety of other environmental benefits such as damages to ecological systems and materials.

Some averting behavior studies focus on behaviors that prevent or mitigate the impact of particular symptoms (e.g., shortness of breath or headaches), while others have examined averting expenditures in response to specific episodes of contamination (e.g., groundwater contamination). The difference in these endpoints is important. Because many contaminants can produce similar symptoms, studies that estimate values for symptoms may be more amenable to benefit transfer than those that are episode-specific. The latter could potentially be more useful, however, for assessing the benefits of a regulation expected to reduce the probability of similar contamination episodes.

Considerations in evaluating and understanding averting behavior studies

Perceived versus actual risks. Analysts should remember that consumers base their actions on perceived benefits from defensive behaviors. Many averting behavior studies explicitly acknowledge that their estimates rest on consistency between the consumer's perception of risk reduction and actual risk reduction. While there is some evidence that consumers are rational with regard to risk — for example, consumer expenditures to reduce risk vary positively with risk increases — there is also evidence that there are predictable differences between consumers' perceptions and actual risks. Thus, averting behavior studies can produce biased WTP estimates for a given change in objective risk. Surveys may be necessary to determine the benefits individuals perceive they are receiving when engaging in defensive activities. These perceived benefits can then be used as the object of the valuation estimates. For example, if surveys reveal that perceived risks are lower than expert risk estimates, then WTP can be estimated with the lower, perceived risk (Blomquist 2004).

Data requirements and implications. Data needed for averting behavior studies include information detailing the severity, frequency, and duration of symptoms; exposure to environmental contaminants; actions taken to avert or mitigate damages; the costs of those behaviors and activities; and other variables that affect health outcomes, like age, health status, or chronic conditions.

Separability of joint benefits. Analysts should exercise caution in interpreting the results of studies that focus on goods in which there may be significant joint benefits (or costs). Many defensive behaviors not only avert or mitigate environmental damages, but also provide other benefits. For example, air conditioners obviously provide cooling in addition to air filtering, and bottled water may not only reduce health risks, but may also taste better. Conversely, it also is possible that the averting behavior may have negative effects on utility. For example, wearing helmets when riding bicycles or motorcycles may be uncomfortable. Failure to account for these “joint” benefits and costs associated with averting behaviors will result in biased estimates of WTP.

Modeling assumptions. Restrictive assumptions are sometimes needed to make averting behavior models tractable. Analysts drawing upon averting behavior studies will need to review and assess the implications of these assumptions for the valuation estimates.

7.3.1.5 Cost of Illness

A frequently encountered alternative to WTP estimates is the avoided cost of illness (COI). The COI method estimates the financial burden of an illness based on the combined value of direct and indirect costs associated with the illness. Direct costs represent the expenditures associated with diagnosis, treatment, rehabilitation, and accommodation. Indirect costs represent the value of illness-related lost income, productivity, and leisure time. COI is better suited as a WTP proxy when the missing components (e.g., pain and suffering) are relatively small as they usually are in cases of in minor, acute illnesses. However, there are usually better medical treatment and lost productivity estimates for more severe illnesses.

The COI method is straightforward to implement and explain to policy makers, and has a number of other advantages. The method has been used for many years and is well developed. Collecting data to implement it often is less expensive than for other methods, improving the feasibility of developing original COI estimates in support of a specific policy.

Economic foundations of COI studies

Two conditions must be met for the COI method to approximate a market value of reduced health risk. First, the direct costs of morbidity must reflect the economic value of goods and services used to treat illness. Second, a person’s earnings must reflect the economic value of lost work time, productivity, and leisure time. Because of distortions in medical and labor markets, these assumptions do not routinely hold. Further, COI estimates are not necessarily equal to WTP. The method generally does not attempt to measure the loss in utility due to pain and suffering, and does not account for the costs of any averting behaviors that individuals have taken to avoid an illness. When estimates of WTP are not available, the potential bias inherent in relying on COI estimates should be acknowledged and discussed. A second shortcoming of the COI method is that by focusing on ex post costs, it does not capture the risk attitudes associated with ex ante measures of reduced health risk.

Although COI estimates do not adequately capture several components of WTP, COI does not necessarily serve as a lower bound estimate of WTP. This is because, for some illnesses, the cost of behaviors that allow one to avoid an illness might be far lower than the cost of the illness itself. Depending on the design of the research question, WTP could reflect the lower avoidance costs while COI would reflect the higher costs of treating the illness once it has been contracted. In addition, COI estimates capture medical expenses passed on to third parties such as health insurance companies and hospitals, whereas WTP estimates generally do not. Finally, COI estimates capture the value of lost productivity (see Text Box 7.4 above), whereas these costs may be overlooked in WTP estimates — especially when derived from consumers or employees covered by sick leave.

Available comparisons of COI and total WTP estimates suggest that the difference can be large (Rowe et al. 1995). This difference varies greatly across health effects and across individuals.

General application by type of COI study

Prevalence-based estimates. Prevalence-based COI estimates are derived from the costs faced by all individuals who have a sickness in a specified time period. For example, an estimate of the total number of individuals who currently have asthma, as diagnosed by a physician, reflects the current prevalence of physician-diagnosed asthma. Prevalence-based COI estimates for asthma include all direct and indirect costs associated with asthma within a given time period, such as a year. Prevalence-based COI estimates are a measure of the full financial burden of a disease, but generally will be lower bound estimates of the total WTP for avoiding the disease altogether. They are useful for evaluating the financial burden of policies aimed at improving the effectiveness of treatment or at reducing the morbidity and mortality associated with a disease.

Incidence-based estimates. By contrast, incidence-based COI estimates reflect expected costs for *new* individuals who develop a disease in a given time period. For example, the number of individuals who receive a *new* diagnosis of asthma from a physician in a year reflects the annual incidence of physician-diagnosed asthma. Incidence-based COI estimates reflect the expected value of direct medical expenditures and lost income and productivity associated with a disease from the time of diagnosis until recovery or death. Because these expenses can occur over an extended time period, incidence-based estimates are usually discounted to the year the illness is diagnosed and expressed in present value terms. Incidence-based COI estimates are useful for evaluating the financial burden of policies that are aimed at reducing the incidence of new cases of disease.

Most existing COI studies estimate indirect costs based on the typical hours lost from a work schedule or home production, evaluated at an average hourly wage. The direct medical costs of illness are generally derived in one of two ways. The empirical approach estimates the total medical costs of the disease by using a database of actual costs incurred for patients with the illness. The “expert elicitation” approach uses a panel of physicians to develop a generic treatment profile for the illness. Illness costs are estimated by multiplying the probability of a patient receiving a treatment by the cost of the treatment. For any particular application, the preferred approach will depend on availability of reliable actual cost data as well as characteristics of the illness under study.

COI estimates for many illnesses are readily available from existing studies and span a wide range of health effects. EPA’s *Cost of Illness Handbook* (U.S. EPA 2007c) provides estimates for many cancers, developmental illnesses and disabilities, and other illnesses.

Considerations in evaluating and understanding COI studies

Technological change. Medical treatment technologies and methods are constantly changing, and this could push the true cost estimate for a given illness either higher or lower. When using previous COI studies, the analyst should be sure to research whether and how the generally accepted treatment has changed from the time of the study.

Measuring the value of lost productivity. Simply valuing the actual lost work time due to an illness may not capture the full loss of an individual’s productivity in the case of a long-term chronic illness. Chronic illness may force an individual to work less than a full-time schedule, take a job at a lower pay rate than she would otherwise qualify for as a healthy person, or drop out of the labor force altogether. A second issue is the choice of wage rate. Even if the direct medical costs are estimated using individual actual cost data, it is highly unlikely that the individual data will include wages. Therefore, the wage rate chosen should reflect the demographic distribution of the illness under study. Furthermore, the value of lost time should include the productivity of those persons not involved in paid jobs. Homemakers’ household upkeep and childcare services, retired persons’ volunteering efforts, and students’ time in school all directly or indirectly contribute to the productivity of society. Finally, the value of lost leisure time to an individual and her family is not

Text Box 7.5 - Value of Time

Estimating the cost of an illness by examining only medical costs clearly understates the true costs experienced by an individual with ill health. Not only does the individual incur medical expenditures, they also miss production and consumption opportunities. In particular they miss opportunities to work for wages, produce household goods and services (e.g., laundry, home-cooked meals), and enjoy leisure activities. These latter two categories are jointly referred to as non-work time. The value of these lost opportunities has typically been estimated by examining the value of time.

EPA has developed an approach for valuing time losses based on the opportunity cost of time. For paid work, the approach is relatively straightforward. It rests on the assumption that total compensation (wages and employment benefits) is equal to the employers' valuation of the worker's output. Therefore, if a worker is absent due to illness, society loses the value of the foregone output, which can be estimated by examining the worker's wages and employment benefit values. To value time spent on non-market work and leisure activities, the assumption is made that an individual will engage in such unpaid activities only if, at the margin, the value of these activities is greater than the wages that could be earned in paid employment. Hence after-tax wages provide a lower bound estimate of the value of non-work time.

The loss of work time and leisure activities due to illness need not be complete. When an illness reduces but does not eliminate productivity at work or enjoyment of leisure time, estimates of the value of the diminishments in these opportunities are legitimate components of the cost of the illness.

Valuing time lost due to illness experienced by children and other subpopulations that do not earn wages is more difficult. Examples of such subpopulations include the elderly, unemployed, or individuals who are out of the work force. Analysts could surmise the post-tax wage if such individuals were employed; however, the situation involves less certainty. For example, the time loss of children who suffer illness is sometimes estimated by considering the effect of the illness, if any, on future earnings. For this case, however, *Circular A-4* (OMB 2003) currently suggests that, in the absence of better data, monetary values for children should be at least be as large as the values for adults (for the same risk probabilities and health outcomes).

Accounting for time losses in COI estimates comes closer to a full accounting of the losses borne by individuals suffering illness than simply assessing medical costs. However, a third cost category remains neglected — the value of pain and suffering. When an individual is sick, she not only misses opportunities to produce or relax, she also would be willing to pay some amount to avoid the pain or discomfort of the illness. In most economic models, these costs are represented as declines in utility and as such are inherently difficult to estimate. To date, there are no good estimates, or methods for obtaining good estimates, of the value of avoiding pain.

included in most COI studies. (See Text Box 7.5 for a discussion of the value of time.)

7.3.2 Stated Preference

The distinguishing feature of stated preference methods compared to revealed preference methods is that stated preference methods rely on data drawn from people's responses to hypothetical questions while revealed preference methods rely on observations of actual choices. Stated preference methods use surveys that ask respondents to consider one or a series of hypothetical scenarios that describe a potential change in a non-market good. The advantages of stated preference methods include their ability to estimate non-use values and to incorporate hypothetical scenarios that closely correspond to a policy case. The main disadvantage of stated preference methods is that they may be subject to systematic biases that are difficult to test for and correct.

National Oceanic and Atmospheric Administration's (NOAA) *The Report of the NOAA Panel on Contingent Valuation* is often cited as a primary source of information on stated preference techniques. Often referred to as the "NOAA Blue Ribbon Panel," this panel, comprised of five distinguished economists including two Nobel Laureates, deliberated on the usefulness of stated preference studies for policy analysis (Arrow et al. 1993). While their findings generally mirror the recommendations offered below, since the release of their report a number of changes in the survey administration "landscape" have occurred including the advent of internet surveys, the decline in representativeness of telephone surveys, and the growth in popularity of stated choice experiments.

7.3.2.1 Economic Foundation of Stated Preference Methods

The responses elicited from stated preference surveys, if truthful, are either direct expressions of WTP or can be used to estimate WTP for the good in question. However, the "if truthful" caveat is paramount. While many environmental economists believe that respondents can provide truthful answers to hypothetical questions and therefore view stated preference methods as useful and reliable if conducted properly, a non-trivial fraction of economists are more skeptical of the results elicited from stated preference surveys. Due to this skepticism, it is important to employ validity and reliability tests of stated preference results when applying them to policy decisions.

If the analyst decides to conduct a stated preference survey or use stated preference results in a benefit transfer exercise, then a number of survey design issues should be considered. Stated preference researchers have attempted to develop methods to make individuals' choices in stated preference studies as consistent as possible with market transactions. Reasonable consistency with the framework of market transactions is a guiding criterion for ensuring the validity of stated preference value estimates. Three components of market transactions need to be constructed in stated preference surveys: the commodity, the payment, and the scenario (Fischoff and Furby 1988).

Stated preference studies need to carefully define the commodity to be valued, including characteristics of the commodity such as the timing of provision, certainty of provision, and availability of substitutes and complements. The definition of the commodity generally involves identifying and characterizing attributes of the commodity that are relevant to respondents. Commodity definition also includes defining or explaining baseline or current conditions, property rights in the baseline, and the policy scenarios, as well as the source of the change in the environmental commodity.[36]

Respondents also must be informed about the transaction context, including the method, timing, and duration of payment. The transaction must not be coerced and the individual should be aware of her budget constraint. The payment vehicle should be described as a credible and binding commitment should the respondent decide to purchase the good. The timing and duration of a payment involves individuals implicitly discounting payments and calculating expected utility for future events. The transaction context and the commodity definition should describe and account for these temporal issues.

The hypothetical scenario(s) should be described so as to minimize potential strategic behavior such as "free riding" or "overpledging." In the case of free riding, respondents will underbid their true WTP for a good if they feel they will actually be made to pay for it but believe the good will be provided nevertheless. In the case of overpledging, respondents pledge amounts greater than their true WTP with the expectation that they will not be made to pay for the good, but believing that their response could influence whether or not the good will be provided. Incentive-compatible choice scenarios and attribute-based response formats have been shown to mitigate strategic responses. Both are discussed below.

It is recognized in both the experimental economics literature and the survey methodology

36 Depending on the scenario, the description of the commodity may produce strong reactions in respondents and could introduce bias. In these cases, the detail with which the commodity of the change is specified needs to be balanced against the ultimate goals of the survey. Regardless, the commodity needs to be specified with enough detail to make the scenario credible.

literature that different survey formats can elicit different responses. Changing the wording or order of questions also can influence the responses. Therefore, the researcher should provide a justification for her choice of survey format and include a discussion of the ramifications of that choice.

7.3.2.2 General Application by Type of Stated Preference Study

Two main types of stated preference survey format are currently used: direct WTP questions and stated choice questions. Stated choice questions can be either dichotomous choice questions or multi-attribute choice questions. Following a general discussion of survey format, each of the stated preference survey formats is described in detail below.

Goals that should guide selection of the survey format include the minimization of survey costs, of non-responsiveness, of unexplained variance, and of complications associated with WTP estimation. For example, open-ended questions require smaller sample sizes and are simpler to analyze than other methods of asking the valuation question. These advantages could lead to significant cost reductions. However, these advantages may be mitigated by higher non-response rates and large unexplained variance in the responses. Moreover, there remains a great deal of uncertainty over the effect of the choice mechanism (i.e., open-ended, dichotomous choice, etc.) on the ability and willingness of respondents to provide accurate and well-considered responses.

Because survey formats are still evolving and many different approaches have been used in the literature, no definitive recommendations are offered here regarding selection of the survey format. Rather, the following sections describe some of the most commonly used formats and discuss some of their known and suspected strengths and weaknesses. Researchers should select a format that suits their topic, and should strive to use focus groups, pretests, and statistical validity tests to address known and suspected weaknesses in the selected approach.

Direct/open-ended WTP questions

Direct/open-ended WTP questions ask respondents to indicate their maximum WTP for the specific quantity or quality changes of a good or service that has been described to them. An important advantage of open-ended stated preference questions is that the answers provide direct, individual-specific estimates of WTP. Although this is the measure that economists want to estimate, early stated preference studies found that some respondents had difficulty answering open-ended WTP questions and non-response rates to such questions were high. Such problems are more common when the respondent is not familiar with the good or with the idea of exchanging a direct dollar payment for the good. An example of a stated preference study using open-ended questions is Brown et al. (1996).

Various modifications of the direct/open-ended WTP question format have been developed in an effort to help respondents arrive at their maximum WTP estimate. In *iterative bidding* respondents are asked if they would pay some initial amount, and then the amount is changed up or down depending on whether the respondent says "yes" or "no" to the first amount. This continues until a maximum WTP is determined for that respondent. Iterative bidding has been shown to suffer from "starting point bias," wherein respondents' maximum WTP estimates are systematically related to the dollar starting point in the iterative bidding process (Rowe and Chestnut 1983, Boyle et al. 1988, and Whitehead 2002). A *payment card* is a list of dollar amounts from which respondents can choose, allowing respondents an opportunity to look over a range of dollar amounts while they consider their maximum WTP. Mitchell and Carson (1989) and Rowe et al. (1996) discuss concerns that the range and intervals of the dollar amounts used in payment card methods may influence respondents' WTP answers.

Stated choice questions

While direct/open-ended WTP questions are efficient in principle, researchers have generally turned to other stated preference techniques in recent years. This is largely due to the difficulties respondents face in answering direct WTP questions and the lack of easily implemented procedures to mitigate these difficulties. Researchers also have noted that direct WTP questions with various forms of follow-up bidding may not be "incentive compatible." That is, the respondents' best strategy in answering these questions is not necessarily to be truthful (Freeman 2003).

In contrast to direct/open-ended WTP questions, stated choice questions ask respondents to choose a single preferred option or to rank options from two or more choices. When analyzing the data the dependent variable will be continuous for open-ended WTP formats and discrete for stated choice formats.[37] In principle, stated choice questions can be distinguished along three dimensions:

- *The number of alternatives each respondent can choose from in each choice scenario* — surveys may offer only two alternatives (e.g., yes/no, or "live in area A or area B); two alternatives with an additional option to choose "don't know" or "don't care;" or multiple alternatives (e.g., "choose option A, B, or C").
- *The number of attributes varied across alternatives in each choice question (other than price)* — alternatives may be distinguished by variation in only a single attribute (e.g., mortality risk) or by variation in multiple attributes (e.g., price, water quality, air quality, etc.).
- *The number of choice scenarios an individual is asked to evaluate through the survey.*

Any particular stated choice survey design could combine these dimensions in any given way. For example, a survey may offer two options to choose from in each choice scenario, vary several attributes across the two options, and present each respondent with multiple choice scenarios through the course of the survey. Using the taxonomy presented in these *Guidelines*, a complete (though cumbersome) description of this format would be a dichotomous choice/multi-attribute/multi-scenario survey. The statistical strategy for estimating WTP is largely determined by the survey format adopted, as described below.

The earliest stated choice questions were simple yes/no questions. These were often called *referendum* questions because they were often posed as, "Would you vote for . . ., if the cost to you were $X?" However, these questions are not always posed as a vote decision and are now commonly called *dichotomous choice* questions.

In recent years, stated preference researchers have been adapting a choice question approach used in the marketing literature called *conjoint analysis*. These are more complex choice questions in which the respondent is asked repeatedly to pick her preferred option from a list of two or more options. Each option represents a package of product attributes. By incorporating a dollar price or cost in each option, stated preference researchers are able to extract WTP estimates for incremental changes in the attributes of the good, based on the preferences expressed by the respondents. Holmes and Adamowicz (2003) refer to this as *attribute-based stated choice*.

Dichotomous choice WTP questions. Dichotomous choice questions present respondents with a specified environmental change costing a specific dollar amount and then ask whether or not they would be willing to pay that amount for the change. The primary advantage of dichotomous choice WTP questions is that they are easier to answer than direct WTP questions, because the respondent is not required to determine her exact WTP, only whether it is above or below the stated amount. Sample mean and median WTP values can be derived from analysis of the frequencies of the yes/no responses to each dollar amount. Bishop and Heberlein (1979), Hanemann (1984), and Cameron and James (1987) describe the necessary statistical procedures for analyzing dichotomous choice responses using logit or probit models. Dichotomous choice responses will reveal an interval containing WTP and in the case of a 'yes' response this interval will be unbounded from above. As a result, significantly larger sample sizes are needed for

37 Some researchers use the term "contingent valuation" to refer to direct WTP and dichotomous choice/referendum formats and "stated preference" to refer to other stated choice formats. In these *Guidelines* the term "stated preference" is used to refer to all valuation studies based on hypothetical choices (including open-ended WTP and stated choice formats), as distinguished from "revealed preference."

dichotomous choice questions to obtain the same degree of statistical efficiency in the sample means as direct/open-ended responses that reveal point-values for WTP (Cameron and James 1987).

To increase the estimation efficiency of dichotomous choice questions, recent applications have commonly used what is called a double-bounded approach. In double-bounded questions the respondent is asked whether she would be willing to pay a second amount, higher if she said yes to the first amount, and lower if she said no to the first amount.[38] Sometimes multiple follow-up questions are used to try to narrow the interval around WTP even further. These begin to resemble iterative bidding style questions if many follow-up questions are asked. Similar to starting point bias in iterative bidding questions, the analyses of double-bounded dichotomous choice question results suggest that the second responses may not be independent of the first responses (Cameron and Quiggin 1994, 1998; and Kanninen 1995).

Multi-attribute choice questions. In multi-attribute choice questions, respondents are presented with alternative choices that are characterized by different combinations of goods and services attributes and prices. Multi-attribute choice questions ask respondents to choose the most preferred alternative (a partial ranking) from multiple alternative goods (i.e., a choice set), in which the alternatives within a choice set are differentiated by their attributes including price (Johnson et al. 1995 and Roe et al. 1996). The analysis takes advantage of the differences in the attribute levels across the choice options to determine how respondents value marginal changes in each of the attributes. To measure WTP, a price (often a tax or a measure of travel costs), is included in multi-attribute choice questions as one of the attributes of each alternative. This price and the mechanism by which the price would be paid need to be explained clearly and plausibly, as with any payment mechanism in a stated preference study. Boyle and Özdemir (2009) examine the impact of question design choices, such as the ordering of attributes and the number of alternatives in a single question, on the mean WTP estimate.

There are many desirable aspects of multi-attribute choice questions, including the nature of the choice being made. To choose the most preferred alternative from some set of alternatives is a common decision experience in posted-price markets, especially when one of the attributes of the alternatives is a price. One can argue that such a decision encourages respondents to concentrate on the trade-offs between attributes rather than taking a position for or against an initiative or policy. This type of repeated decision process may also diffuse the strong emotions often associated with environmental goods, thereby reducing the likelihood of yea-saying or of rejecting the premise of having to pay for an environmental improvement.[39] Presenting repeated choices also gives the respondent some practice with the question format, which may improve the overall accuracy of her responses, and gives her repeated opportunities to express support for a program without always selecting the highest price option.

Some applications of multi-attribute survey formats include Opaluch et al. (1993), Adamowicz et al. (1994), Viscusi et al. (1991), Adamowicz et al. (1997), Adamowicz et al. (1998a), Layton and Brown (2000), Johnson and Desvousges (1997), Boyle et al. (2001), and Morey et al. (2002). Studies that investigate the effects of multi-attribute choice question design parameters include Johnson et al. (2000) and Adamowicz et al. (1997).

7.3.2.3 Considerations in Evaluating Stated Preference Results

Survey mode. The mode used to administer a survey is an important component of survey research design because it is the mechanism by

38 Alberini (1995) illustrated an analysis approach for deriving WTP estimates from such responses and demonstrates the increased efficiency of double-bounded questions. The same study showed that the most efficient range of dollar amounts in a dichotomous choice study design was one that covered the mid-range of the distribution and did not extend very far into the tails at either end.

39 Yea-saying refers to the behavior of respondents when they overstate their true WTP in order to show support for a situation described in survey questions.

which information is conveyed to respondents, and likewise determines the way in which individuals can provide responses for analysis. Until recently there were three primary survey modes: telephone, in-person, and mail. Telephone surveys are primarily conducted with a trained interviewer using random digit dialing (RDD) to contact households. In-person surveys are conducted in a variety of ways, including door-to-door, intercepts at public locations, and via telephone recruiting to a central facility. Mail surveys are conducted by providing written survey materials for respondents to self-administer. As technology and society has changed, so has the preference for one mode over the other. With the influx of market research and telemarketing, the telephone has become a less convenient way to administer surveys. Many people refuse to answer the phone, or to answer questions over the phone. The same can be said of mail surveys. People are quick to ignore unsolicited mail. In recent years the Internet has emerged as a possible mode for conducting surveys. Internet access and email accounts are more prevalent and computer literacy is high in the United States and other developed countries. As with all of the survey modes mentioned, there are inherent biases. These biases are generally classified as social desirability bias, sample frame bias, avidity bias, and non-response bias. See Maguire (2009), Loomis and King (1994), Mannesto and Loomis (1991), Lindberg et al. (1997), and Ethier et al. (2000) for a discussion of different biases in survey mode.

Framing issues. An important issue regarding survey formats is whether information provided in the questions influences the respondents' answers in one way or another. For example, Cameron and Huppert (1991) and Cooper and Loomis (1992) find that mean WTP estimates based on dichotomous choice questions may be sensitive to the ranges and intervals of dollar amounts included in the WTP questions. Kanninen and Kriström (1993) show that the sensitivity of mean WTP to bid values can be caused by model misspecification, failure to include bid values that cover the middle of the distribution, or inclusion of bids from the extreme tails of the distribution.

Selection of payment vehicle. The payment vehicle in a stated preference study refers to the method by which individuals or households would pay for the good described in a particular survey instrument. Examples include increases in electricity prices, changes in cost of living, a one-time tax, or a donation to a special fund. It is imperative that the payment vehicle is incentive compatible and does not introduce any strategic or other bias. Incentive compatibility means that the individual is motivated to respond truthfully and does not use their responses to try to influence a particular outcome (e.g., state a WTP value that is higher than their true WTP to try to make sure a particular outcome succeeds).

Strategic behavior. Adamowicz et al. (1998a) also suggests that respondents may be less likely to behave strategically when responding to multi-attribute choice experiments. Repeatedly choosing from several options gives the respondent some practice with the question format that may improve the overall accuracy of her responses, and gives her repeated opportunities to express support for a program without always selecting the highest price option.

Yea-saying. As mentioned above, yea-saying refers to the behavior of respondents when they overstate their true WTP in order to show support for situation described in survey questions. For example, Kanninen (1995) finds some evidence of yea-saying in dichotomous choice responses through testing in follow-up questions. The extent of this potential problem is not well established, but it may provide an explanation for the fact that mean WTP values based on dichotomous choice responses tend to be equal to or higher than values from direct WTP questions for the same good (Cummings et al. 1986, Boyle et al. 1993, Brown et al. 1996, Ready et al. 1996, and Balistreri et al. 2001). It has not been determined whether yea-saying can be reduced by double-bounded dichotomous choice because in this case the respondent has more than one opportunity to say yes.

Treatment of "don't know" or neutral responses. Based on recommendations from the NOAA Blue Ribbon panel (Arrow et al. 1993), many surveys now include "don't know" or "no preference" options for respondents to choose from. There have been questions about how such responses should enter the empirical analysis. Examining referendum-style dichotomous choice questions, Carson et al. (1998) found that when those who chose not to vote were coded as "no" responses, the mean WTP values were the same as when the "would not vote" option was not offered. Offering the "would not vote" option did not change the percentage of respondents saying "yes". Thus, they recommend that if a "would not vote" option is included, it should be coded as a "no" vote, a practice that has become widespread. Stated preference studies should always be explicit about how they treat "don't know," "would not vote," or other neutral responses.

Reliability, in general terms, means consistency or repeatability. If a method is used numerous times to measure the same commodity, then the method is considered more reliable the lower the variability in the results.

- **Test-retest approach.** Possibly the most widely applied approach for assessing reliability in stated preference studies has been the test-retest approach. Test-retest assesses the variability of a measure between different time periods. Loomis (1989), Teisl et al. (1995), McConnell et al. (1998), and Hoban and Whitehead (1999) all provide examples of the test-retest method for reliability.
- **Meta-analysis of stated preference survey results** for the same good also may provide evidence of reliability. Meta-analysis evaluates multiple studies as though each was constructed to measure the same phenomenon. Meta-analysis attempts to sort out the effects of differences in the valuation approach used in different surveys, along with other factors influencing the elicited value. For example Boyle et al. (1994) use meta-analysis to evaluate eight studies conducted to measure values for groundwater protection. (Also see Section 7.4.)

Validity tests seek to assess whether WTP estimates from stated preference methods behave as a theoretically correct WTP should. Three types of validity discussed below are: content validity, criterion validity, and convergent validity.

- **Content validity.** Content validity refers to the extent to which the estimate captures the concept being evaluated. Content validity is largely a subjective evaluation of whether a study has been designed and executed in a way that incorporates the essential characteristics of the WTP concept. In a sense, it is akin to asking, "On the face of it, does the estimate capture the concept of WTP?" (This approach is sometimes referred to as "face validity.")

  To evaluate a survey instrument, analysts look for features that researchers should have incorporated into the survey scenario. First, the environmental change being valued should be clearly defined. A careful exposition of the conditions in the baseline case and how these would be expected to change over time if no action were taken should be included. Next, the action or policy change should be described, including an illustration of how and when it would affect aspects of the environment that people might care about. Boyd and Banzahf (2007), and Boyd and Krupnick (2009) put a finer point on this concept and advocate developing the valuation scenario based on "ecological endpoints" rather than intermediate goods that are less clearly associated with outcomes of interest. For example, if respondents ultimately care about the survival of a certain species, it is more sensible to structure questions to ask about WTP for the species' survival than to ask about degradation of habitat, as respondents are unlikely to know the relationship between habitat attributes and species survival. Respondent attitudes about the provider and the implied property rights of the survey scenario can be used to evaluate the appropriateness of features related to the payment mechanism (Fischhoff and Furby 1988). Survey questions that probe for respondent comprehension and acceptance of the commodity scenario can offer important indications about the validity of the results (Bishop et al. 1997).

- **Criterion validity.** Criterion validity assesses whether stated preference results relate to other measures that are considered to be closer to the concept being assessed (WTP). Ideally, one would compare results from a stated preference study (the measure) with those from actual market data (the criterion). This is because market data can be used to estimate WTP more reliably than a stated preference survey. Another approach would be to estimate a sample of individuals' WTP for a commodity using a stated preference survey and then later give the same sample of individuals or a different random sample of individuals drawn from the same population a real opportunity to buy the good. (See Mitchell and Carson 1989, Carson et al. 1987a, Kealy et al. 1990, Brown et al. 1996, and Champ et al. 1997 for examples.)

  When unable to conduct such comparisons, sensitivity to scope and income has been used to assess criterion validity. "Scope tests" are concerned with how WTP responds to changes in the amount of the referenced good provided in the valuation scenario (Smith and Osborne 1996, Rollins and Lyke 1998, and Heberlien et al. 2005). If the referenced good is indeed a "normal good" utility theory implies that WTP should increase with the provision of the good. For the same reason one would expect WTP to exhibit positive income elasticity (McFadden 1994, and Schlapfer 2006). Neither test is necessary or sufficient to establish criterion validity (Heberlein et al. 2005) but can serve as useful proxies when an alternate measure of WTP for the same good is unavailable. Diamond (1996) suggests that stronger scope tests can be conducted by comparing departures from strict "adding up" of WTP for partial changes and relating them to the income elasticity of WTP. Other researchers, however, argue that the Diamond test may not be practicable or even necessarily correct (Carson et al. 2001).

- **Convergent validity.** Convergent validity examines the relationship between different measures of a concept.[40] This differs from criterion validity in that one of the measures is not taken as a criterion upon which to judge the other measure. The measure of interest and the other measure are judged together to assess consistency with one another. If they differ in a systematic way (e.g., one is usually larger than another for the same good), it is not clear which one is more correct. However, if stated preference results are found to be larger than revealed preference results for the same good, it is often presumed that the difference is the result of hypothetical bias because revealed preference results are based on actual behavior. There can be many other sources of bias and error in both stated preference and revealed preference results that cause them to differ from one another and from "true" WTP.

  Empirical convergent validity tests use comparisons of stated preference results with revealed preference or experimental results that are thought to be free of hypothetical bias.[41] In some circumstances, convergent validity tests may be incorporated as part of the study design. Such a test might compare results of an actual market exercise with the results of a hypothetical market exercise in which the exercises are otherwise identical. In this case there might be evidence of an upward or downward bias in the hypothetical results as compared to the simulated market results. See Section 7.3.3 for a discussion on combining revealed preference and stated preference data.

Hypothetical bias occurs when the responses to hypothetical stated preference questions are

40 Mitchell and Carson (1989) define convergent validity and theoretical validity as two types of construct validity. Construct validity examines the degree to which the measure is related to other measures as predicted by theory.

41 Some analysts include the comparisons of stated preference results to actual markets under convergent validity rather than criterion validity, as discussed in the previous section, because there is no actual observable measure of the theoretical construct WTP. Here, a distinction is made between simulated markets, as in a laboratory experiment in which values may be "induced" by giving subject cash at the end based on their choices, and actual markets in which subjects must pay with their own money.

systematically different than what individuals would pay if the transactions were to actually occur. Widely cited as one of the most common problems with the stated preference method (List and Gallet 2001, and Murphy and Allen 2005), and researchers have made advances in techniques to minimize such bias. These techniques include the use of "cheap talk" methods to directly tell respondents about the potential for hypothetical bias (Cummings and Taylor 1999, and List 2001); calibrating hypothetical values (List and Shogren 1998, and Blomquist et al. 2009); and allowing respondents to express uncertainty in their responses and restricting the set of positive responses to those about which the respondent was most certain (Vossler et al. 2003). Several studies have shown that attribute-based choice experiments reduce hypothetical bias in the bid amounts and the marginal value of attributes relative to other elicitation methods (Carlsson and Martinsson 2001, Murphy and Allen 2005, and List et al. 2006).

Tests for hypothetical bias often involve a comparison of actual payments and responses to hypothetical scenarios that use the same solicitation approach. The actual payments typically occur in one of three scenarios. Market transactions are the most common (Cummings et al. 1995, and List and Shogren 1998) but generally involve payments for private goods while most stated preference applications are concerned with public or quasi-public goods. Simulated markets can be used to solicit actual donations for public good provision (Champ et al. 1997). However, donation solicitations are subject to free riding, so while it may be possible to test for hypothetical bias using this approach, both the actual and hypothetical payment scenarios lack incentive compatibility and may not represent total WTP. In rare instances comparisons have been made between actual referenda for public good provision and hypothetical responses to the same scenario but the conditions for a valid comparison of this sort are exceedingly difficult to satisfy (Johnston 2006).

Non-response bias is introduced when non-respondents would have answered questions systematically differently than those who did answer. Non-response bias can take two forms: item non-response and survey non-response.

- **Item non-response bias** occurs when respondents who agreed to take the survey do not answer all of the choice questions in the survey. Information available about respondents from other questions they answered can support an assessment of potential item non-response bias for the WTP questions that were unanswered. The key issue is whether there were systematic differences in potential WTP-related characteristics of those who answered the WTP questions and those who did not. Characteristics of interest include income, gender, age, expressed attitudes and opinions about the good or service, and information reported on current use or familiarity with the good or service. Statistically significant differences may indicate the potential for item non-response bias, while finding no such differences suggests that the chance of significant non-response bias is lower. However, the results of this comparison are only suggestive because respondents and non-respondents may only differ in their preference for the good in question (McClelland et al. 1991).

- **Survey non-response bias** is created when those who refuse to take the survey have preferences that are systematically different from the preferences of those who do respond. Although it is generally thought that surveys with high response rates are less likely to suffer from survey non-response bias, it is not a guarantee.[42] For survey non-respondents, there may be no available data to determine how they might systematically differ from those who responded to the survey. The

42 Note that OMB's *Guidance on Agency Survey and Statistical Collections* (OMB 2006) has fairly strict requirements for response rates and their calculation for Agency-sponsored surveys, recommending that "ICRs for surveys with expected response rates of 80 percent or higher need complete descriptions of the basis of the estimated response rate...ICRs for surveys with expected response rates lower than 80 percent need complete descriptions of how the expected response rate was determined, a detailed description of steps that will be taken to maximize the response rate...and a description of plans to evaluate non-response bias" (pp. 60-70).

most common approach is to examine the relevant measurable characteristics of the respondent group, such as income, resource use, gender, age, etc., and to compare them to the characteristics of the study population. Similarity in mean characteristics across the two groups suggests that the respondents are representative of the study population and that non-response bias is expected to be minimal.

A second way to evaluate potential survey non-response bias is to conduct a short follow-up survey with non-respondents. This can sometimes be accomplished through interviews conducted during the recruiting phase. Such follow-ups typically ask a few questions about attitudes and opinions on the topic of the study as well as collecting basic socioeconomic information. Questions need to match those in the full survey closely enough to compare non-respondents to respondents. The follow-up must be very brief or response rates will be low (OMB 2006).

7.3.3 Combining Revealed and Stated Preference Data

Instead of looking at revealed preference and stated preference data as two separate methods for estimating environmental benefits, an increasing number of researchers are using them in combination. The practice has been in use much longer in the marketing and transportation literature and many of the lessons learned by those researchers are now being employed in environmental economics. In theory, the strengths of each data type should help overcome some of the weaknesses of the other. As described by Whitehead et al. (2008) in a recent assessment of the state of the science, the advantages of combining revealed preference and stated preference data include:

- Helping to ground the hypothetical stated preference data with real world behavior potentially decreasing any hypothetical bias;
- Providing the ability to test the validity of both data sources;[43]
- Increasing the range of historical stated preference data to include conditions not observed in the past and thereby reducing the need to make predictions outside of the sample;
- Increasing the sample size;
- Extending the size of the market or population to include larger segments than captured by either method alone; and
- Exploiting the flexibility of stated preference experimental design to overcome revealed preference data's potential multicollinearity and endogeneity problems (von Haefen and Phaneuf 2008).

The different strategies for combining revealed preference and stated preference data can be roughly grouped into three main methods. The first two methods rely on joint estimation. If the revealed preference and stated preference data have similar dependent and independent variables and the same assumed error structures, then they can simply be pooled together and treated as additional observations (Adamowicz et al. 1994; Boxall, Englin, and Adamowicz 2003; and Morgan, Massey, and Huth 2009). If the revealed preference and stated preference data sources cannot be pooled, it is sometimes possible to use them in a jointly estimated mixed model that relies on a utility theoretic specification of the underlying WTP function (Huang, Haab, and Whitehead 1997; Kling 1997; and Eom and Larson 2006). If the data cannot be combined in estimation, it can still be useful to estimate results separately and then use them to test for convergent validity between the two data sources (Carson et al. 1996, and Schlapfer et al. 2004).

7.4 Benefit Transfer

Benefit transfer refers to the use of estimated non-market values of environmental quality changes from one study in the evaluation of a different policy that is of interest to the analyst (Freeman 2003, p. 453). The case under consideration for a

43 Herriges, Kling, and Phaneuf (2004) point out that revealed preference may not always be valid for estimating WTP for quality changes when weak complementarity cannot be assured.

new policy is referred to as the "policy case." Cases from which estimates are obtained are referred to as "study cases." A benefit transfer study identifies stated preference or revealed preference study cases that sufficiently relate to the policy context and "transfers" their results to the policy case.

Benefit transfer is necessary when it is infeasible to conduct an original study focused directly on the policy case. Original studies are time consuming and expensive; benefit transfer can reduce both the time and financial resources required to develop estimates of a proposed policy's benefits. While benefit transfer should only be used as a last resort and a clear justification for using this approach over conducting original valuation studies should be provided (OMB 2003), the reality is that benefit transfer is one of the most common approaches for completing a BCA at EPA. However, the advantages of benefit transfer in terms of time and cost savings must be weighed against the disadvantages in terms of potential reduced reliability of the final benefit estimates. The transfer of benefits estimates from any single study case is unlikely to be as accurate as a primary study tailored specifically to the policy case, although it is difficult to characterize the uncertainty associated with transferred benefits estimates.

The number and quality of relevant studies available for application to the policy case can limit the use of benefit-transfer methods.[44] Even when a study case is qualitatively similar to the policy case, the environmental change associated with the policy case may be of a different scope or nature than the changes considered in the study cases. In addition, methodological advances and changes in demographic, economic, and environmental conditions over time may make otherwise suitable studies obsolete.[45]

Steps for conducting benefit transfer

While there is no universally accepted single approach for conducting benefit transfer there are some generalized steps involved in the process. These steps are described below.

1. Describe the policy case. The first step in a benefit-transfer study is to clearly describe the policy case so that its characteristics and consequences are well understood. Are human health risks reduced by the policy intervention? Are ecological benefits expected (e.g., increases in populations of species of concern)? It is also important to identify to the extent possible the beneficiaries of the proposed policy and to describe their demographic and socioeconomic characteristics (e.g., users of a particular set of recreation sites, children living in urban areas, or older adults across the United States). Information on the affected population is generally required to translate per person (or per household) values to an aggregate benefits estimate.

2. Select study cases. A benefit-transfer study is only as good as the study cases from which it is derived, and it is therefore crucial that studies be carefully selected. First, the analyst should identify potentially relevant studies by conducting a comprehensive literature search. Because peer-reviewed academic journals may be more likely to publish work using novel approaches compared to established techniques, some studies of interest may be found in government reports, working papers, dissertations, unpublished research, and other "gray literature."[46] Including studies from the gray literature may also help mitigate "publication bias" that results from researchers being more likely to present and/or editors being more likely to publish studies that demonstrate statistically significant results, or results that are of an expected sign or magnitude.[47] Online searchable databases

44 One possible reason that a relatively limited number of value estimates exist in peer-reviewed literature is that researchers and editors of scholarly journals may be more interested in new theoretical or methodological advances than in studies that apply established valuation methods to confirm earlier findings.

45 A 2006 special issue of *Ecological Economics* (volume 60) focused exclusively on benefit transfer for environmental policy, covering diverse topics such as publication bias, theoretical motivation and emerging issues. Florax et al. (2002), and Navrud and Ready (2007) are two general references for benefit transfer studies.

46 Peer review of benefit-transfer studies using gray literature is highly advisable.

47 There is some evidence of publication bias towards studies showing statistically significant results. For example, in a meta-analysis of studies in labor economics, Card and Krueger (1995) argue that just-significant results are reported more frequently than would be predicted by chance. Similar practices may prevail in other areas of economic research. Combining results from a group of studies that suffer from publication bias may lead to inaccurate conclusions. See Stanley (2005, 2008) for a discussion of methods to correct for and identify publication bias.

summarizing valuation research may be especially helpful at this stage.[48]

Next, the analyst should develop an explicit set of selection criteria to evaluate each of the potentially relevant studies for quality and applicability to the policy case. The quality of the value estimates in the study cases will in large part determine the quality of the benefit transfer. As a first step, the analyst should review studies according to the criteria listed for each methodology in the previous sections in this chapter. Results from study cases must be valid as well as relevant. Concerns about the quality of the studies, as opposed to their relevance, will generally hinge on the methods used. Valuation approaches commonly used in the past may now be regarded as unacceptable for use in benefits analysis. Studies based on inappropriate methods or reporting obsolete results should be removed from consideration.

It is unlikely that any single study will match perfectly with the policy case; however each potential study case should inform at least some aspect of the policy decision. Study cases potentially suitable for use in benefit transfer should be similar to the policy case in their: (1) definition of the environmental commodity being valued (include scale and presence of substitutes); (2) baseline and extent of environmental changes; and (3) characteristics of affected populations. Analysts should avoid using benefit transfer in cases where the policy or study case is focused on a "good" with unique attributes or where the magnitude of the change or improvement across the two cases differs substantially (OMB 2003).[49]

48 For example, the EVRI is maintained by Environment Canada and managed by a working group that includes the U.S. EPA and members of the European Union. EVRI contains over 1,100 studies that can be referenced according to medium, resource, stressor, method, and country. EVRI also provides a bibliography on benefit transfer. See www.evri.ca for more information. Envalue, developed by the New South Wales EPA in 1995, is similar: Studies can be identified according to medium, stressor, method, country, and author.

49 In some cases the transfer method itself may inform the choice of study cases to include. For example, meta-analysis approaches (discussed below) can facilitate some forms of statistical validity testing (Hunter and Schmidt 1990, and Stanley 2001), so some otherwise suitable studies may be rejected as "outliers."

The analyst should determine whether adjustments should and can be made for important differences between each study and policy case. For example, some case studies will report Marshallian demand while others may report Hicksian demand.[50] The ability of the analyst to make these adjustments will depend, in part, on both the number of value estimates for suitably similar study sites and the method used to combine these estimates. These methods are now discussed in turn.

3. Transfer values. There are several approaches for transferring values from study cases to the policy case. These include unit value transfers, value function transfers, and non-structural or structural meta-analysis. Each of these approaches is typically used to develop per person or per household value estimates that are then aggregated over the affected population to compute a total benefits estimate. As a general rule, the more related case study estimates involved in a benefit transfer, the more reliable the estimate.

Unit value transfers are the simplest of the benefit-transfer approaches. They take a point estimate of WTP for a unit change in the environmental resource from a study case or cases and apply it directly to the policy case. The point estimate is commonly a single estimated value from a single case study, but it can also be the (otherwise unadjusted) average of a small number of estimates from a few case studies. For example, a study may have found a WTP of $20 per household for a one-unit increase on some water quality scale. A unit value transfer would estimate total benefits for the policy case by multiplying $20 by the number of units by which the policy is expected to increase water quality and by the number of households who will benefit from the change. This approach can be useful for developing preliminary, order-of-magnitude estimates of benefits, but it should be possible to base final benefit estimates on more

50 See Desvousges et al. (1992), Brouwer (2000), Florax et al. (2002), Bergstrom and Taylor (2006), and Navrud and Ready (2007) for additional information on criteria used to determine quality and applicability. For more information on applicability as related to specific benefit categories, see Desvousges et al. (1998), the draft *Handbook for Non-Cancer Valuation* (U.S. EPA 2000c), and the *Children's Health Valuation Handbook* (U.S. EPA 2003b). It may also be useful for the analyst to discuss her interpretation and intended use of the study case with the original authors.

Text Box 7.6 - The Benefits and Costs of the Clean Air Act 1990 to 2010: Reduced Acidification in Freshwater Adirondack Lakes

One component of the total benefits of the Clean Air Act (CAA) was determined to be improved recreational fishing due to reduced acidification in freshwater Adirondack lakes. To value this benefit, EPA relied on the results of Montgomery and Needleman's (1997) New York State Adirondack region recreational fishing study. EPA first developed estimates of the percentage Adirondack of lakes affected by acidification pre and post CAA. Then, using a probit model, the likelihood that each individual lake would become acidified was estimated (the model relates acidity to lake characteristics such as elevation, surface area, watershed, and others) and the lakes were ranked from highest to lowest probability of being acidified. The acidification status of individual lakes in the choice set was then assigned, starting with the highest probability lake and proceeding down until the appropriate number of lakes affected under each scenario (i.e., the estimated percentage of lakes affected) was achieved. Using these lake designations and the Montgomery and Needleman model's estimated coefficients, welfare was calculated for the pre and post CAA levels of lake acidification. The difference between the two welfare estimates was assumed to be the value of improved Adirondack freshwater recreational fishing under the CAA.

information than a single point estimate from a single study. Point estimates reported in study cases are typically functions of several variables, and simply transferring a summary estimate without controlling for differences among these variables can yield inaccurate results. It is important to recognize that unit value transfer assumes that the original good, as well as the characteristics and tastes of the population of beneficiaries, are the same as the policy good. Unit values transfers should only be used if the case and policy studies are evaluating the same environmental good, the same change in environmental levels, and same affected populations.

Function transfers also rely on a single study, but they use information on other factors that influence WTP to adjust the unit value for quantifiable differences between the study case and the policy case. This is accomplished by transferring the estimated function upon which the value estimate in the study case is based to the policy case. This approach implicitly assumes that the population of beneficiaries to which the values are being transferred has potentially different characteristics, but similar tastes, as the original one and allows the analyst to adjust for these different characteristics. Generally, benefit function transfers are preferable to unit value transfers as they incorporate information relevant to the policy scenario (OMB 2003). For example, suppose that in the hypothetical example above the $20 unit value was the result of averaging the results of an estimated WTP function over all individuals in the study case sample, where the WTP function included income, the baseline water quality level, and the change in the water quality level for each household. A function transfer would estimate total benefits for the policy case by:

1. Applying the WTP function to a random sample of households affected in the policy case using each household's observed levels of income, baseline water quality, and water quality change;
2. Averaging the resulting WTP estimates; and
3. Multiplying this average WTP by the total number of households affected in the policy case.

See Text Boxes 7.6 and 7.7 for examples of value and function transfers.

If the WTP function is nonlinear and statistics on average income, baseline water quality, and water quality changes are used in the transfer instead of household level values, then bias would result. Feather and Hellerstein (1997) provide an example of a function transfer that attempts to correct for such bias. Although unit transfers can adjust and compensate for small differences between the case and policy study populations, they are subject to the same basic usage rules governing unit value transfers. Function transfers should only be used if the case and policy studies are evaluating very similar environmental goods, change in environmental levels, and affected populations.

Text Box 7.7 - Benefits Transfer: Water Quality Benefits in the Combined Animal Feeding Operations Rule

There are two prominent water quality benefit-transfer applications in the 2002 Combined Animal Feeding Operations (CAFO) rule. The first looks at the recreational value of water quality improvements in fresh water lakes and streams (see Section 4 of U.S. EPA 2002c). Field pollutant loadings were modeled by the National Water Pollution Control Assessment Model (NWPCAM) to produce pre and post regulation water quality estimates. Predicted changes in water quality were then valued using the results of Carson and Mitchell's (1993) national water quality contingent valuation survey. First, benefits were calculated based on estimates of willingness to pay (WTP) for water quality improvements resulting in discrete movements to higher "rungs" of the water quality ladder (boatable, fishable, swimmable, drinkable). Very simply described, Carson and Mitchell's "in-state" WTP estimates for discrete movements up the water quality ladder were multiplied by the number of affected residents in every state and "out-of-state," non-use values were multiplied times the remaining population. State totals were then summed up to a national total (see Appendix A-4 of U.S. EPA 2002c for more details). Benefits were also estimated a second way based on a continuous (1 to 100) water quality index constructed from six water quality parameters measured in the NWPCAM model. The minimum thresholds between rungs on the water quality ladder were then translated into points along the continuous water quality index (i.e., boatable = 25, fishable = 50, swimmable = 70). Carson and Mitchell's WTP function was then used to value changes in water quality as measured by the water quality index (see Appendix B-4 of U.S. EPA 2002c for more details). Benefits estimated by the water quality index method are larger by roughly a factor of two (Exhibits 4-12 and 4-13 of U.S. EPA 2002c).

The second major benefit-transfer application in the CAFO rule involves the valuation of reduced eutrophication in estuaries (Section 9 of U.S. EPA 2002c). EPA used a case study of Albemarle and Pamlico sounds to demonstrate the potential importance and value of reduced eutrophication on recreational fishing in affected estuaries. Again, NWPCAM was used to estimate pre and post regulation water quality levels. In this case, the benefit transfer made use of three studies (Kaoru 1995; Kaoru, Smith, and Liu 1995; and Smith and Palmquist 1988), all of which were based in part on the same dataset. All "reasonable" estimates of WTP for reduced phosphorus or nitrogen from the studies were retained and translated into their corresponding dollar per trip per ton reduction in pollutant per year value. A range of total benefits was then calculated by multiplying each $/trip/ton/year estimate by the number of trips taken and the change in loadings (in tons) for each pollutant (see Exhibit 9-3 of U.S. EPA 2002c).

Meta-analysis uses results from multiple valuation studies to estimate a new transfer function. Meta-analysis is an umbrella term for a suite of techniques that synthesize the summary results of empirical research. This could include a simple ranking of results to a complex regression. The advantage of these methods is that they are generally easier to estimate while controlling for a relatively large number of confounding variables. This approach has been widely used in environmental economics (Poe et al. 2001, Shrestha and Loomis 2003a and 2003b, Rosenberger and Loomis 2000, and Bateman and Jones 2003).

There are a number of guidelines for meta-analyses that outline protocols that should be followed in conducting or evaluating a study. See Begg et al. (1996), Moher et al. (1999), and U.S. EPA (2006e) for more information.[51] More recently Bergstrom and Taylor (2006) discuss the theory and practice underlying meta-analysis for benefit transfer, discussing three major necessary steps: theory, data collection, and analysis. In general, when reporting meta-analysis results, researchers should provide information on the background of the problem, the strategy for selecting studies, analytic methods, results, discussion, and conclusions. See U.S. EPA (2006e) for a detailed discussion of meta-analysis as applied to VSL estimates. U.S. EPA (2006e) specifically recommends carefully specifying the search process, selection criteria, and analytical methods.

Structural benefit transfer is a relatively new approach to benefit transfer. The advantages of

51 The last reference contains a detailed discussion of the protocols for conduction a meta-analsysis.

Text Box 7.8 - Structural Benefit Transfer with an Application to Visibility

U.S. EPA (2006b) employs a structural benefit transfer to derive values for visibility improvements associated with the Particulate Matter (PM) National Ambient Air Quality Standards (NAAQS). It specified a constant elasticity of substitution utility function for visibility in residential and Class I (national park and similar) areas. This function assumes that the value for Class I visibility differs in and out of region but that residential visibility is valued the same everywhere. EPA also assumed that in-region visibility was valued more highly than out-of-region visibility. The function further specified utility as a function of: (1) consumption of all goods; (2) visibility in a person's residential area; (3) recreational visibility in a person's residential region; and (4) recreational visibility outside of a person's residential region. Given the utility function and a budget constraint, it was then possible to define households' WTP for changes in visibility as a function of income and visibility measures. The regional preference parameters of the function were calibrated using existing WTP estimates for visibility in Class I areas (Chestnut and Rowe 1990, and Chestnut 1997) if estimates existed for a given region. If not, estimates were adjusted by visitation rate. The preference parameter for residential visibility was assumed to be the same in all counties and was solved for based on a WTP estimate presented in McClelland et al. (1991). With estimates of visibility (pre and post regulation), county-level income, and the required preferences parameters, nationwide estimates of the value of increased visibility were then computed for each of the six regions of the country.

structural transfer functions are that they can accommodate different types of economic value measures (e.g., WTP, WTA, or consumer surplus) and can be constructed in such a way that certain theoretical consistency conditions (e.g., WTP bounded by income) can be satisfied. This could be applied to value transfer, function transfer, or meta-analysis; although applications to function transfer are the most common. Structural transfer functions that have been estimated have specified a theoretically consistent preference model that is calibrated according to existing benefit estimates from the literature (see Smith and Pattanayak 2002; and Smith, Pattanayak, and van Houtven 2006 for descriptions on the method). See Text Box 7.8 for an application to of structural benefit transfer to visibility benefits.

4. Report the results. In addition to reporting the final benefit estimates from the transfer exercise, the analyst should clearly describe all key judgments and assumptions, including the criteria used to select study cases and the choice of the transfer approach. The uncertainty in the final benefit estimate should be quantified and reported when possible. (See Chapter 11 on Presentation of Analysis and Results.)

7.5 Accommodating Non-Monetized Benefits

It often will not be possible to quantify all of the significant physical impacts for all policy options. For example, animal studies may suggest that a contaminant causes severe illnesses in humans, but the available data may not be adequate to determine the number of expected cases associated with different human exposure levels. Likewise, it often is not possible to quantify the various ecosystem changes that may result from an environmental policy. While Chapter 11 discusses how to present these benefits so as to provide a fuller accounting of all effects, this section discusses what analysts can do to incorporate these endpoints more fully into the analysis.

7.5.1 Qualitative Discussions

When there are potentially important effects that cannot be quantified, the analyst should include a qualitative discussion of benefits results. The discussion should explain why a quantitative analysis was not possible and the reasons for believing that these non-quantified effects may be important for decision making. Chapter 11 discusses how to describe benefit categories that are quantified in physical terms but not monetized.

7.5.2 Alternative Analyses

Alternative analyses exist that can support benefits valuation when robust value estimates and/or risk estimates are lacking. These analyses, including break-even analysis and bounding analysis, can provide decision makers with some useful information. However analysts should remember that because these alternatives do not estimate the net benefits of a policy or regulation, they fall short of BCA in their ability to identify an economically efficient policy. This and other shortcomings should be discussed when presenting results from these analyses to decision makers.

7.5.2.1 Break-Even Analysis

Break-even analysis is one alternative that can be used when either risk data or valuation data are lacking.[52] Analysts who have per unit estimates of economic value but lack risk estimates cannot quantify net benefits. They can, however, estimate the number of cases (each valued at the per unit value estimate) at which overall net benefits become positive, or where the policy action will break even.[53] Consider a proposed policy that is expected to reduce the number of cases of endpoint X with an associated cost estimate of $1 million. Further, suppose that the analyst estimates that WTP to avoid a case of endpoint X is $200, but that because of limitations in risk data, it is not possible to generate an estimate of the number of cases of this endpoint reduced by the policy. In this case, the proposed policy would need to reduce the number of cases by 5,000 in order to "break even." This estimate then can be assessed for plausibility either quantitatively or qualitatively. Policy makers will need to determine if the break-even value is acceptable or reasonable.

The same sort of analysis can be performed when analysts lack valuation estimates, producing a break-even value that should again be assessed for credibility and plausibility. Continuing with the example above, suppose the analyst estimates that the proposed policy would reduce the number of cases of endpoint X by 5,000 but does not have an estimate of WTP to avoid a case of this endpoint. In this case, the policy can be considered to break even if WTP is at least $200.

One way to assess the credibility of economic break-even values is to compare them to risk values for effects that are more or less severe than the endpoint being evaluated. For the break-even value to be plausible, it should fall between the estimates for these more and less severe effects. For the example above, if the estimate of WTP to avoid a case of a more serious effect was only $100, the above break-even point may not be considered plausible.

Break-even analysis is most effective when there is only one missing value in the analysis. For example, if an analyst is missing risk estimates for two different endpoints (but has valuation estimates for both), then they will need to consider a "break-even frontier" that allows the number of both effects to vary. It is possible to construct such a frontier, but it is difficult to determine which points on the frontier are relevant for policy analysis.

7.5.2.2 Bounding Analysis

Bounding analysis can help when analysts lack value estimates for a particular endpoint. As suggested above, reducing the risk of health effects that are more severe and of longer duration should be valued more highly than those that are less severe and of shorter duration, all else equal. If robust valuation estimates are available for effects that are unambiguously "worse" and others that are unambiguously "not as bad," then one can use these estimates as the upper and lower bounds on the value of the effect of concern. Presenting alternative benefit estimates based on each of these bounds can provide valuable information to policy makers. If the sign of the net benefit estimate is positive across this range then analysts can have some confidence that the program is welfare enhancing. Analysts should carefully describe judgments or assumptions made in selecting appropriate bounding values.

52 Boardman et al. (1996) describes determining break-even points under the general subject of sensitivity analysis and includes empirical examples.

53 *Circular A-4* (OMB 2003) refers to these values as "switch points" in its discussion of sensitivity analysis.