IEc

Addressing Uncertainty in Regulatory Impact Analysis

Final White Paper | September 3, 2021

prepared for:

Office of the Assistant Secretary for Planning and Evaluation
U.S. Department of Health and Human Services

prepared by:

Jennifer Baxter & William Raich
Industrial Economics, Incorporated
2067 Massachusetts Avenue
Cambridge, MA 02140
617/354-0074
www.indecon.com

Dr. Robin Dillon-Merrill
ITA International, LLC
700 Tech Center Parkway
Suite 202
Newport News, VA 23606
www.ita-intl.com

INDUSTRIAL ECONOMICS, INCORPORATED

AKNOWLEDGEMENTS

This paper was prepared by Industrial Economics, Incorporated (IEc) and ITA International, LLC (ITA) for the Office of the Assistant Secretary for Planning and Evaluation (ASPE) in the U.S. Department of Health and Human Services (HHS) under Contract No. HHSP233201500035I, Task Order No. HHSP23337015T, Subcontract No. 50295S05856. Mathematica Incorporated is the prime contractor.

The ASPE project leads are Aaron Kearsley and Scott Douglas; we are grateful for their technical input and support. The paper was drafted by Jennifer Baxter (IEc) and Dr. Robin Dillon-Merrill (ITA and Georgetown University). William Raich (IEc) provided additional support preparing and drafting the Appendix B.

TABLE OF CONTENTS

ACRONYMS

ANPRM	-	Advance Notice of Proposed Rulemaking
APPS	-	Advanced Personal Protection System
ASPE	-	Assistant Secretary for Planning and Evaluation
CBP	-	U.S. Customs and Border Protection
DHS	-	U.S. Department of Homeland Security
DOT	-	U.S. Department of Transportation
EPA	-	U.S. Environmental Protection Agency
HHS	-	U.S. Department of Health and Human Services
IEc	-	Industrial Economics, Incorporated
IAM	-	Integrated Assessment Models
ITA	-	ITA International, LLC
NPS	-	National Park Service
NRC	-	Nuclear Regulatory Commission
OMB	-	U.S. Office of Management and Budget
PHMSA	-	U.S. Pipeline and Hazardous Materials Safety Administration
RIA	-	Regulatory Impact Analysis
SMEs	-	Subject Matter Experts
SSHAC	-	Senior Seismic Hazard Analysis Committee
USFWS	-	U.S. Fish and Wildlife Service
VSL	-	Value Per Statistical Life
WTP	-	Willingness to pay

1.0 INTRODUCTION

Estimates of the likely costs, benefits, and distributional effects of proposed regulations are inherently uncertain. To develop these estimates, analysts must collect data and construct models that attempt to predict the future. These models rely on information that may be subject to limitations related to: the quality of the methods used to collect the data; the extent to which the data address the same population, industries, or geographic area as the regulation; and the degree to which conditions may change between when the data were collected and when the regulation is implemented. The models also require many assumptions. For example, analysts must make assumptions about how regulated entities will respond to the regulation (e.g., likely compliance rates; the methods likely to be chosen to achieve compliance; etc.). They must also make assumptions about the future state of the world in the absence of the regulation (e.g., future output in regulated industries; exposure risks absent intervention, etc.).[1]

A critical challenge for analysts is to clearly describe the key sources of uncertainty associated with these estimates, in qualitative or quantitative terms, in the regulatory impact analysis (RIA). The goal is to ensure that decision-makers and other stakeholders understand the extent to which uncertainty – in the data, models, and assumptions – affects the main analytic conclusions. A well-developed presentation of uncertainty can aid decision-makers in understanding the confidence they should have in the results and the magnitude of any bias.

For example, if the agency's best estimates suggest that benefits exceed costs for a particular regulatory option, how likely is it that this conclusion would be reversed given uncertainty about the magnitudes of the quantified effects and the potential impact of non-quantified effects? Might these uncertainties affect the relative rankings of the policy options? Answering these questions requires quantifying impacts to the greatest extent possible, and identifying key sources of uncertainty and exploring them in both quantitative and qualitative terms. Over time, analysts can work to reduce uncertainty and minimize the types of effects that cannot be quantified by anticipating future analytic needs and investing in research that will be useful across a variety of regulatory analyses.

In 2016, the U.S. Department of Health and Human Services (HHS) finalized its *Guidelines for Regulatory Impact Analysis* (hereafter *Guidelines*) under the leadership of its Assistant Secretary for Planning Evaluation (ASPE) and Analytics Team. In Chapter 6, "Address Uncertainty and Nonquantifiable Effects," the *Guidelines* discuss strategies for characterizing the uncertainty in quantified effects as well as the potential impacts of

[1] In this white paper, the discussion references prospective analysis of future regulations. The same challenges and tools apply to retrospective analysis of existing regulations. In retrospective analysis, analysts estimate the incremental effects of the regulation by comparing two scenarios: the world with the regulation (the "incremental scenario") and the world without the regulation (the "counterfactual scenario"). The counterfactual scenario cannot be observed; it must be modeled, introducing uncertainty into estimates of incremental costs and benefits.

non-quantified effects. It provides an overview of basic concepts, along with a summary discussion of different approaches and their complexity.

This white paper expands on the discussion provided in the *Guidelines*. It provides a more detailed discussion of terminology, tools, and methods that may be used in uncertainty and sensitivity analyses conducted as part of the development of an RIA. It also provides a discussion of methods that may be used for communicating uncertainty to different stakeholder audiences and summarizes best practices for conducting these analyses. This white paper does not constitute new *Guidelines*.

2.0 TERMINOLOGY

In this section, we introduce terminology commonly employed by analysts engaged in uncertainty analysis. We begin with a discussion of different types of uncertainty. Then, we provide definition for common terms used throughout the remainder of this white paper.

2.1 TYPES OF UNCERTAINTY

Uncertainty comes in two general forms: one that can more easily be modeled and quantified and one that cannot. These two forms are often referred to as “variability” and “uncertainty.”

- **Variability:** refers to the inherent heterogeneity or diversity of data in an assessment (EPA 2021). Variability cannot be reduced, but it can be better characterized. For example, the height of individuals in a population is variable. Sampling a larger population of individuals provides a better picture of distribution of individual height in the population; however, variation will always exist. Variability is also referred to as “randomness” (Morgan and Henrion 1990).
- **Uncertainty:** refers to a lack of data or an incomplete understanding of the context of the assessment (EPA 2021). For example, predictions about future trends in regulated industries are inherently uncertain, because one is attempting to predict the future. Similarly, characterizations of the likely change in behavior of affected entities in response to a new regulation may be subject to significant uncertainty. In both cases, uncertainty can be reduced or eliminated over time with new or better data.

In fact, both are sources of uncertainty. The more academic terms for the distinction in types of uncertainty are “aleatory and epistemic uncertainties.”

- **Aleatory uncertainty.** Uncertainty in assessments because events are inherently stochastic (Tannenbaum et al. 2017, Paté-Cornell, 1996).
- **Epistemic uncertainty.** Uncertainty in assessments because of a lack of knowledge about what is true or what may happen. Epistemic uncertainties, while

not known now, are potentially knowable (Tannenbaum et al. 2017, Paté-Cornell, 1996).[2]

While technical in nature, it is clearer to describe the two distinct types of uncertainties using these terms.[3,4]

Importantly, in their seminal text on addressing uncertainty in quantitative policy analysis, Morgan and Henrion (1990) note,

> *A common mistake is failure to distinguish between variability due to sampling from a frequency distribution and empirical uncertainty that arises from incomplete scientific or technical knowledge…The uncertainty due to the variability can be reduced by disaggregation…whereas the scientific uncertainty can be reduced by further research…Examination of the first kind of uncertainty can tell you how much disaggregation is worthwhile in performing an assessment. Examination of the second kind of uncertainty may tell you about the relative importance of carrying out more…research. But this sort of uncertainty analysis is impossible unless the two kinds of uncertainty are carefully distinguished.*

Thus, understanding the difference between these types of uncertainty has important implications for interpreting information about uncertainty provided in an RIA and for developing and evaluating options for reducing that uncertainty in future analyses. In cases where aleatory uncertainty is more dominant, analysts might consider disaggregating the data and modeling subpopulations. In contrast, if epistemic uncertainty is more important, additional research may be necessary.

2.2 GENERAL DEFINITIONS

In this section, we provide definitions for additional key terminology.[5]

- **Correlation:** A mathematical relationship between two sets of data indicating the probability that a dependency exists between the parameters underlying the two data sets and the strength of that (linear) dependency. This means that knowing the value of one parameter gives you some information about the value of the other. For example, a person's height and weight are strongly correlated. Thinking

---

[2] Makridakis, Hogarth, and Gaba (2009) provides an illustrative story of a person facing these two types of uncertainty, which they have nicknamed "subway" vs. "coconut" uncertainty. Their fictitious character Klaus faces regular "subway" uncertainty every day as his commute time varies around an average travel time of 43 minutes with 95 percent of the trips between 37 and 49 minutes. Their fictitious character Klaus also faces the possibility of a totally unexpected event with significant consequences such as a death by a coconut falling from a tree.

[3] For example, under EPA's more limited definition of uncertainty, "randomness" would not qualify as uncertainty. Such a distinction would be inconsistent with the academic literature.

[4] In *Circular A-4*, OMB (2003) writes, "In some contexts, the word 'variability' is used as a synonym for statistical variation that can be described by a theoretically valid distribution function, whereas 'uncertainty' refers to a more fundamental lack of knowledge. Throughout this discussion [of the treatment of uncertainty in RIAs], we use the term 'uncertainty' to refer to both concepts.

[5] Definitions are taken from Leach (2006), with the exceptions of "uncertainty", "value of information," and "dependence."

about correlation is particularly important in Monte Carlo analysis, where analysts must account for correlation between distributions. For example, it could be problematic to independently draw height and weight from two different distributions during a simulation; the results might be nonsensical (e.g., a 6 foot tall individual is unlikely to weigh 100 pounds).

- **Cumulative Distribution:** For a cumulative distribution, also called the “S-curve,” values are plotted along the x-axis and the probability of occurrence of a value less than or equal to the given x-value is plotted on the y-axis. By definition, the y-axis goes from 0 to 1. It would be useful to display the output of a Monte Carlo simulation as a cumulative distribution if, for example, it is important to describe the value of concern at a 70 percent confidence level.
- **Decision Tree:** An analysis tool in which several discrete values are used to represent uncertainty. A true decision tree would contain decision nodes representing different alternative choices. The term decision tree is often used to refer to tree diagrams without decisions, but tree diagrams without decision trees are more correctly referred to as event trees or probability trees. The “tree” is built from left to right with each uncertainty and/or decision appearing in chronological order. Probabilities are assigned for each branch of each uncertainty. Decision trees are used to both illustrate the relationships between key inputs in a process and as a modeling tool (see software tools such as the Palisade Company’s Precision Tree).
- **Dependence:** Correlation (defined above) measures the linear dependence of two random variables, but variables can have non-linear dependency. Variables are considered dependent when knowing the value of one tells us something about the value of the other. Alternatively, two random variables are considered independent (and not correlated) when knowing the value that one random variable takes on tells us nothing about the distribution of the other.
- **Monte Carlo Simulation:** A process in which a probability distribution is given for each of the inputs (e.g., number of affected entities, hours per inspection, wage rates per inspector) to some model (e.g., a model estimating total compliance costs) and then a computer calculates the output of the model many times (usually hundreds or thousands of times). On each trial, the computer draws a different value for each input parameter based on the defined input probability distributions. At the end of the simulation, statistics are calculated on the output variables of interest. Monte Carlo simulation is specifically referenced by OMB in *Circular A-4* (OMB 2003) as a useful tool for formal, probabilistic analysis of proposed regulations. Descriptions of commonly used probability distributions are provided in Appendix A.
- **Tornado Chart:** A chart used in sensitivity analyses to indicate the impact each uncertain variable has on the model output values. The tornado chart is useful for determining which sources of uncertainty have the potential to have the most significant impacts and which will have minimal impacts. For example, if the

output variable is the total costs of a rule, a tornado chart can be used to understand which uncertain variables have the biggest impact on the total cost estimate.

- **Uncertainty:** A measure of a value of which we do not know and cannot know until sometime in the future. See the discussion above in Section 2.1.
- **Value of Information:** The increase in the output of the analysis that comes from acquiring a certain piece of information. Analysts use the value of information to make decisions about how to allocate scarce research resources and to decide whether delaying action while collecting more information is preferable.

3.0 SELECTING AN APPROACH

As described in Chapter 1 of the *Guidelines*, Executive Orders 12866 and 13563 (Clinton 1993, Obama 2011) direct Federal agencies to assess the costs and benefits of proposed regulations, as permitted by existing statutes. The Office of Information and Regulatory Affairs within the U.S. Office of Management and Budget (OMB) provides specific guidance in *Circular A-4* (OMB 2003) on the estimation of regulatory costs and benefits and the preparation of RIAs. OMB encourages analysts to present "probability distributions of benefits and costs and include the upper and lower bound estimates as complements to central tendency and other estimates" (OMB 2003). However, it also recognizes that presenting probabilistic estimates may not always be practical or feasible. In such cases, OMB directs analysts to "balance thoroughness with the practical limits on your analytical capabilities."

Noting that an analysis does not need to include an exhaustive exploration of every source of uncertainty, OMB suggests first identifying the assumptions or data that have the largest potential effect on decision-making. For example, as described in Chapter 2 of the *Guidelines*, analysts might undertake screening analysis using readily available information and simple assumptions to provide preliminary information on potential costs and benefits and to identify key assumptions or data needs. The screening analysis may highlight the assumptions that make the biggest difference in the results and that could be the focus of additional research or consideration. Sensitivity analysis, described below and in Section 4.0 of this document, is another good tool for this purpose.

If key sources of uncertainty are likely to have a significant effect on the conclusions about net benefits (e.g., the ranking of regulatory alternatives or whether net benefits are positive), OMB (2003) suggests that the agency consider additional research prior to rulemaking. Uncertainties may be significant enough to warrant delaying a decision until more information can be collected and assessed. When considering whether to recommend a delay, analysts should take into account both costs (e.g., of further data gathering efforts) and benefits (e.g., of the knowledge likely to be obtained from the new data). Delay may also have consequences for social welfare (e.g., if it allows dangerous practices to continue), which should also be considered along with the impacts of any interim protective measures.[6] In Section 6.0 of this document, we discuss approaches for estimating the value of collecting additional information in order to reduce uncertainty.

The discussion above suggests that analysts have discretion when deciding which sources of uncertainty to explore, and what methods to use. Below, we provide additional guidance on how to make these choices. First, we discuss the guidance provided by OMB in *Circular A-4*. Then, we highlight additional considerations and summarize key questions for analysts to consider as they make choices about how to proceed with uncertainty analysis.

3.1 OMB GUIDANCE

In *Circular A-4*, OMB outlines three approaches for addressing uncertainty and provides general guidance on when each should be used. Each of these approaches serves to improve decision-makers' and the public's understanding of the range of impacts possible under a regulation and is superior to ignoring uncertainty entirely. Ranked from less rigorous to more rigorous, these approaches include:

- **Qualitative Discussion.** Qualitative discussion of key sources of uncertainty is the least rigorous approach, but is of significant importance. It should always be included in the RIA. This approach involves disclosing key assumptions and uncertainties and including information about the implications. To the greatest extent possible, the qualitative discussion should include both the likely direction of the potential bias (i.e., whether the assumption may lead to an under- or over-estimate of the impacts) and the likely magnitude of the effect (e.g., whether it is major or minor). Such information will help decision-makers and others better understand the implications of the analysis.

- **Numerical Sensitivity Analysis.** Numerical sensitivity analysis allows the analyst to explore the effects of varying the values of key parameters and is often used to determine whether uncertainty about particular components or assumptions may substantially affect the analytic result, as well as when data limitations or constrained resources prevent full probabilistic analysis. Sensitivity analysis may be particularly useful in situations where the qualitative discussion raises questions about the robustness of the results or where the consequences of the rule are large. Methods for communicating the results of a sensitivity analysis vary and may include the presentation of alternative scenarios reflecting different, plausible assumptions for key parameters.

- **Probabilistic Analysis.** OMB (2003) recommends using "probabilistic analysis" to explore key sources of uncertainty for "complex rules where there are large, multiple uncertainties whose analysis raises technical challenges, or where the effects cascade." Additionally, probabilistic analysis is required for rules with annual costs or benefits exceeding $1 billion. Such analysis often involves the use of Monte Carlo simulation (discussed in Section 4.0) to quantify the probability

[6] If the timing of the regulation is determined by statute or court order, delay may not be possible.

distributions of anticipated costs or benefits. It provides decision-makers with information about the variance, or spread, of the statistical distribution of estimated impacts. This information may be particularly useful when the expected value of net benefits is close to zero or similar across multiple policy alternatives. In such cases, decision-makers may feel more confident about the results if they have a smaller variance, because the realized results are more likely to be near the expected value. OMB also suggests the use of expert elicitation (discussed in Section 5.0) to characterize uncertainty in terms of a distribution for use in simulations.

Finally, OMB (2003) also identifies break-even analysis as a useful tool in situations where it is not possible to quantify all of the important benefits and costs likely to result from a proposed regulation. Break-even analysis assists decision-makers in thinking about whether it is plausible that the regulation will result in positive net benefits (e.g., benefits that exceed costs), given existing data gaps (e.g., uncertainties). Break-even analysis is discussed below in Section 4.0.

3.2 OTHER CONSIDERATIONS

In addition to the guidance provided in *Circular A-4*, analysts should also consider what is at stake in the rulemaking effort when selecting an approach for addressing uncertainty. A high stakes rule may be one that is likely to result in large costs or benefits (described above in Section 3.1), have significant distributional impacts, raise major concerns about equity, or otherwise be highly controversial.[7] A low stakes rule has the opposite characteristics. Ultimately, the decision about how to categorize a particular rule for the purpose of making analytic choices is a judgment made by the agency. Generally, higher stakes regulations justify allocating more time and resources and employing more complex approaches for investigating or resolving uncertainty. For lower stakes rules, simpler approaches and qualitative discussion of uncertainty may be sufficient.

Analysts should be careful to use quantitative tools judiciously. The simple availability of a tool does not justify its use. For example, in some cases, data limitations prevent the quantification of the benefits of a proposed rule, and analysts discuss benefits qualitatively. In these cases, complex analysis of costs may be important for many of the reasons described above (e.g., the rule is likely to be subject to considerable scrutiny, probabilistic analysis is required because costs are likely to exceed $1 billion in a given year). However, if costs are small, and stakes are low, and quantitative analysis of uncertainty will not help improve decision-making, then a qualitative discussion of uncertainty may be sufficient.[8]

---

[7] For example, even when a rule is unlikely to result in costs or benefits exceeding $100 million in a given year, the agency may be aware that the regulation is likely to be the subject of intense public scrutiny due to a high degree of stakeholder interest. The agency may also be concerned about the likelihood of litigation after the regulation in promulgated. In these situations, more robust uncertainty analysis will likely provide additional support to the agency in demonstrating that it followed a reasoned decision-making process (i.e., its decision was not arbitrary or capricious).

[8] For example, imagine a low-stakes rule where the analyst is able to quantify costs but not benefits. For the cost analysis, she has probability distributions for parameters with small variance, while distributions for key parameters with large

3.3 KEY QUESTIONS TO ASK WHEN SELECTING AN APPROACH

In Exhibit 1, we summarize key questions for analysts to consider as they make decisions about whether to address uncertainty in an RIA qualitatively or quantitively. At a minimum, all RIAs should include a qualitative discussion. Many combine qualitative and quantitative approaches.

EXHIBIT 1. SELECTING AN APPROACH FOR ADDRESSING UNCERTAINTY

QUESTION	RECOMMENDATION
Will the regulation result in costs or benefits exceeding $1 billion in a given year?	If the answer is "yes," then probabilistic analysis is required by OMB.
Will the regulation result in costs or benefits exceeding $100 million but less than $1 billion in a given year?	Probabilistic analysis is not required however OMB recommends using more complex approaches. These approaches may include, for example, Monte Carlo simulation, numerical sensitivity analysis, or scenario analysis.
Will the regulation result in costs and benefits each less than $100 million in a given year?	The recommended approach depends on additional factors, discussed below.
Do simpler methods provide robust results?	If "yes," then simpler methods are sufficient.
Is the uncertainty in the modeling results sufficiently large that net benefits could be positive or negative?	If "yes" then more complex methods are justified.
Is the uncertainty in the modeling results sufficiently large that the ranking of regulatory alternatives could change?	If "yes," then more complex methods are justified.
Does available information lend itself to a quantitative tool that will improve decision-making?	If "yes," then more complex methods are justified.
Is this a "high stakes" rule?	If "yes," then more complex methods are justified.
Do data limitations prevent the quantification of certain impacts, and would the benefits of additional information outweigh the costs of collecting it and any delay in regulatory action?	If the answer is "yes," then additional effort to quantify impacts is justified. Otherwise, qualitative analysis is likely to be sufficient.

Note:
Examples of simpler methods include qualitative discussion of uncertainty and break-even analysis. More complex methods attempt to characterize uncertainty quantitatively using sensitivity and scenario analysis, or probabilistically using tools such Monte Carlo analysis.

variance are unknown. In this situation, presenting cost estimates resulting from a Monte Carlo simulation may provide a false sense of precision, because the range of possible costs will not reflect the key source of uncertainty in the analysis. In such a situation, qualitative discussion of the key sources of uncertainty in the cost estimate, paired with the qualitative discussion of possible benefits, may be more easily interpreted by the decision-maker.

4.0 TOOLS

In this section, we describe several tools that may be helpful in addressing uncertainty in regulatory analyses. These tools are not mutually exclusive; a strong analysis may use multiple tools in combination. At a general level, we present the tools in the order of increasing level of analytic sophistication.

We begin with a discussion of break-even analysis, which does not explicitly treat uncertainty but judges what reasonable values of a key unknown parameter would support the selection of a preferred regulatory alternative. Scenario analysis can be used to address different types of uncertainty quantitively, such as uncertainty regarding key parameters or model selection. Analysts present results for a defined number of defensible scenarios, not the entire range of possible combinations. In contrast, decision trees and Monte Carlo simulation are useful tools for including the entire range of possible probabilistic outcomes. Finally, we discuss sensitivity analysis last, not because it is the most sophisticated tool, but because it is valuable for all the previously discussed tools.

4.1 BREAK-EVEN ANALYSIS

A break-even analysis can be used when key uncertainties are too difficult to quantify. A break-even analysis answers the question, “How small could the value of the non-quantified benefits be (or how large would the value of the non-quantified costs need to be) before the rule would yield zero net benefits?” (OMB 2003). For example, imagine that an agency is considering a regulation intended to reduce cases of food poisoning, but information quantifying the likely effectiveness of the proposed intervention, in terms of avoided cases, is unavailable. Without this information, an analyst cannot estimate the net benefits of the proposed rule. However, if the analyst can estimate the cost of the policy and the willingness to pay (WTP) to avoid a case of illness, then a break-even analysis will determine the number of cases that would need to be avoided for benefits to equal costs (i.e., for net benefits to equal zero). A break-even analysis is most effective when there is only one missing value in the analysis.

For example, Thrift and von Winterfeldt (2021) document a U.S. Department of Homeland Security (DHS) study that examines the benefits and costs of investments in an advanced personal protection system (APPS) for wildland firefighters. In this analysis two key uncertain parameters are the reduction of injury and fatality risks due to use of APPS and the market penetration rate of the subject technology. Using break-even analysis, they identify that the break-even point (net benefits = $0) is reached when the reduction of fatality and injury risks is about 3.42 percent or when the market penetration rate is 0.18 percent, which corresponds to the sale of 985 APPD garments. The authors believe that these break-even values are at the very low-end of plausible estimates and thus, the break-even analysis suggests that there is a high likelihood of a positive net benefits for the APPS.

Advantages

Analysis tends to be simple to calculate and easy to interpret.

Limitations

Break-even analysis only examines one variable or factor at a time and may require significant assumptions regarding other variables or factors that could be explored with a more complex analysis tool. Additionally, it does not provide a quantitative estimate of the benefits or costs of a proposed regulation.

4.2 SCENARIO ANALYSIS

Scenario analysis, also referred to in the literature as "scenario planning," is a disciplined method for imagining possible futures (Schoemaker 1995). The goal is to simplify situations with large amounts of data and uncertainty into a limited number of possible states, where each scenario describes how elements interact under certain conditions. Scenario analysis organizes the range of uncertainties into narratives that are easier to understand than large volumes of data. Examples of uncertain aspects of the future are rates of innovation (high versus low) or sales growth (positive versus negative).

Scenario analysis will not try to depict all possible outcomes of each uncertainty. The purpose is not to cover all possibilities but to circumscribe them (high, medium, and low). The focus is not on forecasting the future or fully characterizing all future uncertainties, but on bounding uncertainty (Schoemaker 1991), generally looking at best and worst cases of uncertainties.

Scenario analysis may be particularly useful in situations where potential outcomes require different types of models. For example, in response to a regulation, regulated entities may decide to incur compliance costs and continue producing their product in the United States, or they may decide to move production to another country. Estimating the costs of each scenario requires a fundamentally different model. Trends or issues combined with uncertainties are the key components of scenarios.

The essential steps a scenario analysis include:

1. Define the general scope of the analysis including the time frame and key issues that will contribute to the uncertainties;
2. Identify the key uncertainties and the relationships among uncertainties since not all combinations may occur;
3. Construct initial scenario themes – some example scenarios may be the extreme worlds where all positive or all negative elements occur together. Other scenarios may focus on the most important scenarios;
4. Check for consistency and plausibility – for example, do the scenarios combine outcomes of uncertainties that are plausible? Generally two to four scenarios are sufficient (Schoemaker 1991); and

5. Name the scenarios. A scenario is a story. Make sure the scenarios describe different futures rather than variations on one theme.

For example, during the crafting of the 2012 proposed rule designating critical habitat for the northern spotted owl pursuant to the Endangered Species Act, stakeholders expressed considerable disagreement about how the regulation would be implemented in the future. Based on considerable research and extensive interviews, the U.S. Fish and Wildlife Service (USFWS) decided to model three different plausible scenarios in its RIA, including: (1) a scenario resulting only in minor administrative costs and negligible benefits; (2) a scenario where timber harvests on Federal land increase, resulting in minor costs and significant quantifiable benefits; and (3) a third scenario resulting in significant costs in terms of forgone timber harvests and uncertain, non-quantified benefits (IEc 2012). The results of all three scenarios are given equal weight in the RIA in terms of presentation, allowing the decision-maker to use his judgment about the most likely outcome when deciding on the configuration of the final rule.

In a second example, the EPA sought to regulate the management of hazardous waste pharmaceuticals at healthcare facilities pursuant to the Resource Conservation and Recovery Act. Incremental costs and benefits of a proposed regulation are measured relative to a baseline scenario (i.e., the world without the regulation). Due to significant uncertainty about baseline waste disposal practices, and the sensitivity of the results of assumptions about these practices, EPA modeled two different baseline scenarios: (1) it assumed full compliance with existing RCRA regulations, and (2) it assumed partial compliance with existing regulations (EPA 2015). The partial compliance baseline reflects the likelihood that many facilities may not manage their hazardous waste pharmaceuticals in a manner consistent with existing requirements. In this case, the full compliance scenario is presented in the report's executive summary, and the alternative baseline scenario is presented to demonstrate the sensitivity of the impact estimates to assumptions about baseline conditions.

Advantages

Scenarios can provide a sound conceptual framework for the analysis. Depending on the level of detail required in the analysis, scenario analysis may be combined with other tools (including decision trees and Monte Carlo simulation). Scenario analysis provides a good structuring tool and a good communications tool.

Limitations

The goal of a scenario analysis is not to consider all combinations of all uncertainties. Without careful consideration and articulation, critical scenarios or uncertainties could be missing.

If one of the scenarios is intended to be a "bounding" scenario, care must be taken to understand exactly what that scenario represents because it is often possible to imagine worse possible scenarios. Additionally, upper and lower bounds need to be plausible to be helpful to support the decision, so a truly worst-case scenario, if it is extremely unlikely, may be a distraction. If a bounding scenario is used and justified as a conservative estimate of an effect, the analysts must be careful in the remaining analysis not to ignore unmeasured qualitative effects that were not captured. Finally, care must be taken when reporting bounding scenarios so that a decision-maker does not misinterpret the resulting range of estimates as corresponding to the 95th and 5th confidence bounds recommended in *Circular A-4*. As shown in Monte Carlo Simulations, combining 95th percentiles for multiple distributions does not provide an estimate of the 95th percentile for the combination of uncertain variables. In general, scenario analyses provide policymakers with a range of possible outcomes, but without a clear estimate of the likelihood of these outcomes.

4.3 EVENT TREES, PROBABILITY TREES, AND DECISION TREES

A tree diagram is a handy visual tool that can be used to represent probabilities for both dependent and independent uncertain events. Each path through the tree represents a possible scenario of the outcomes of the various events logically sequenced in time. The probability of the outcomes of the different scenarios defined by the tree are determined by multiplying the probability values of the connected branches for each path.

Whether a tree diagram is technically a decision tree, an event tree, or a probability tree depends on the characteristics of the nodes in the tree. Decision trees include decisions as well as chance nodes (probabilistic events). Decision trees are particularly useful to understand scenarios over time when there are 'downstream' decisions or options available to the decision-maker once some uncertain events have been resolved.

The expected value of a decision tree is calculated by starting at the right end and working backwards to the base. At each uncertainty node, the expected value of its branches is found, and at each decision node, the branch that maximizes/minimizes the expected value/expected cost is chosen. Using this approach, the best decisions and their expected values are found (Clemen 1996).

Event trees were primarily developed in the nuclear power industry to identify accident sequences or failure paths, so all the nodes in an event tree are chance nodes with binary outcomes of success or failure for each event (USNRC 1975, 1983). A probability tree is a similar, but more general, tool than an event tree. In a probability tree, event nodes can have more than binary success/failure branches.

All three types of tree diagrams are generally constructed using deductive logic, starting with an initiating decision or event, and then considering the occurrence or non-occurrence of other possible events. The probabilities associated with each additional event (or node) are conditional on all previous outcomes in the tree. Most people refer to the different variations of trees as decision trees even if there are no decisions (Leach 2006).

For example, Exhibit 2 provides a probability tree describing the sequence of uncertainties that are important when considering the potential beneficial outcomes of changes to mammogram reporting recommendations (based on FDA 2019a). The reporting requirement would provide more information about the density of the breast in the mammogram. If the image shows high density, a certain portion of these patients would be advised for follow-on ultrasound screening. A significant uncertainty is whether the patient follows the advice because benefits are realized only if the patient seeks treatment following screening and additional cancer cases are detected. Again, the probabilities of each of the branches is the product of the probabilities at each node moving from left to right. To illustrate, working from left to right along the top branch, the probability that an image shows high breast density, the patient is advised to undergo additional ultrasound screening, the patient follows the advice, and additional cancer is detected is 0.0116 percent (0.42 * 0.40 * 0.164 * 0.0042 = 0.000116).

In a second example, Exhibit 3 provides a tree used to explain the three levels of review in the preliminary RIA for the Premarket Tobacco Product Applications and Recordkeeping Requirements (US FDA, 2019b). This tree was not used to demonstrate dependencies among uncertain values like Exhibit 2, but instead to demonstrate uncertain paths through the review process. As shown in the exhibit, at each level of review, the agency determines if the applicant provided sufficient detail to proceed to the next level of review. To estimate the cost of time spent on application reviews, the number of applicants screened out at each level is an important variable. Using the tree structure to explain the process provides clarity and insight into the uncertainty associated with the time and cost variables.

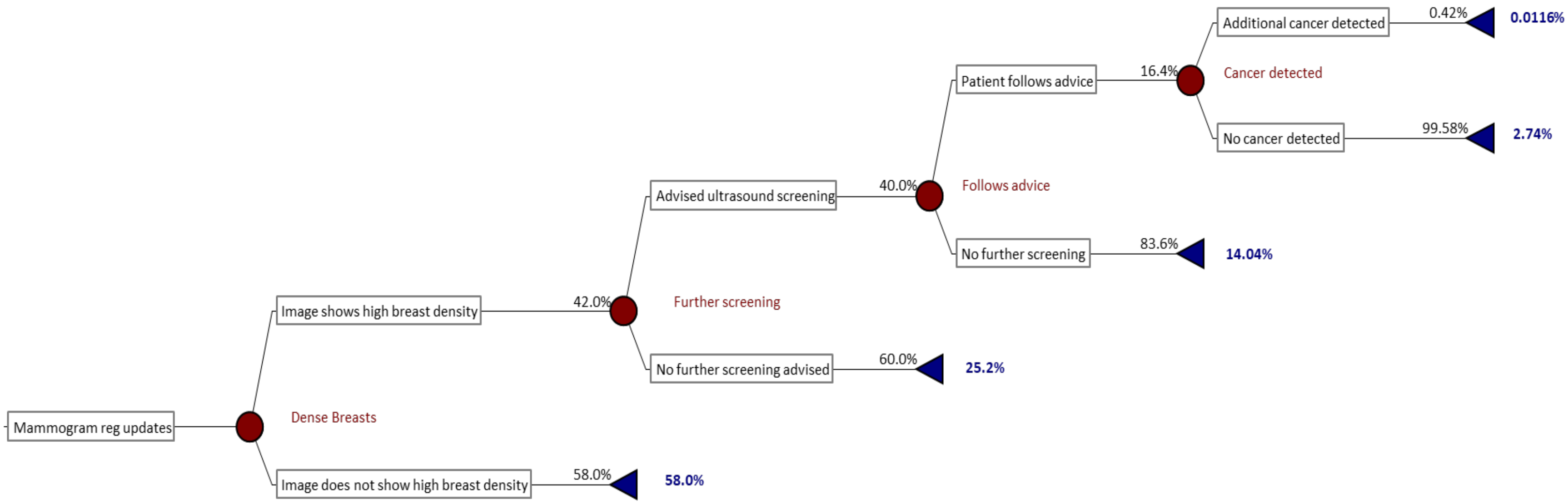
Advantages

- Analysts can concentrate upon one specific scenario at a time.
- Results provide a graphic that lays out the key uncertainties in the model.
- Trees can capture complex dependent temporal events or conditional probabilities.
- An exhibit with a decision tree provides clear insight into the interplay between uncertainties and decisions if decisions are included.

Limitations

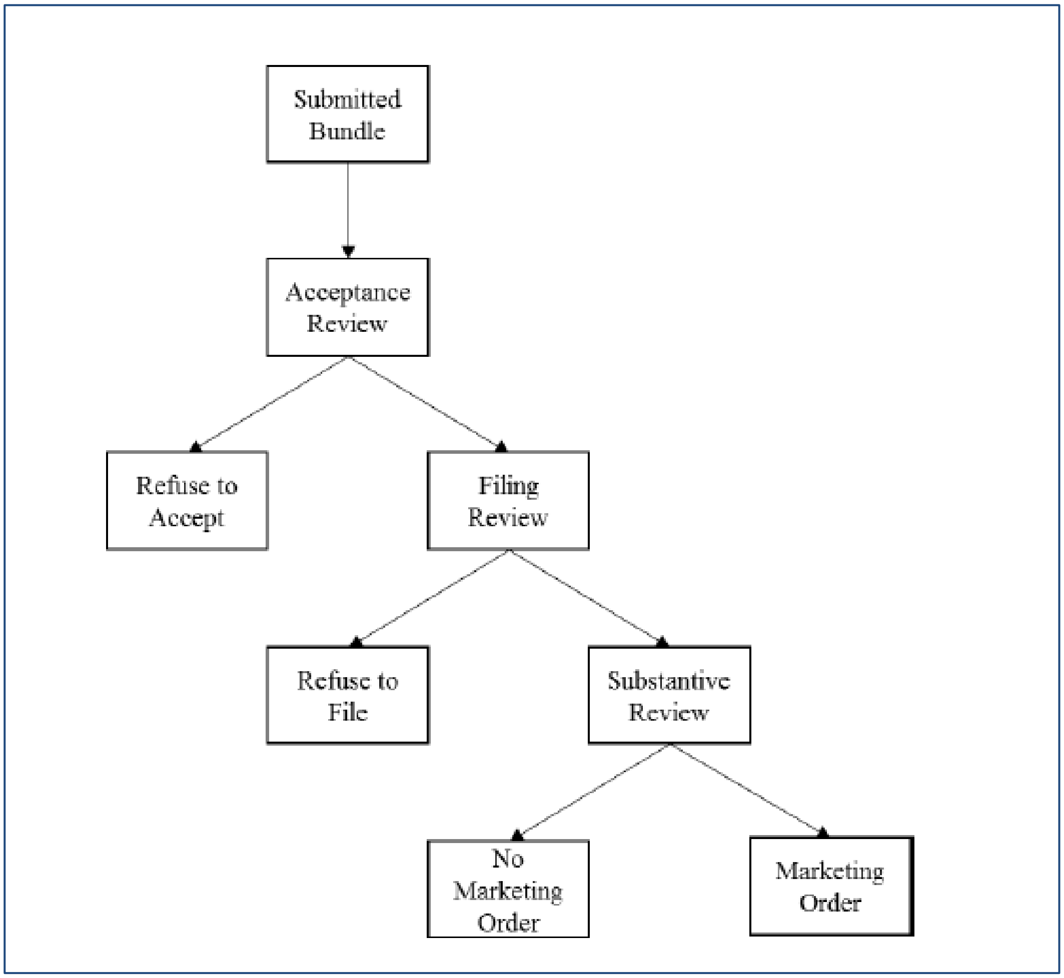
- A few decision nodes can quickly make trees large and complex.
- Trees are best when key uncertainty parameters are characterized by discrete values, such as the event happens or it does not, or the outcomes result in discrete levels such as low, medium and high values. Uncertainties that are continuous probability distributions would need to be discretized to be displayed in a decision tree.

EXHIBIT 2. EXAMPLE DECISION TREE EVALUATING REVISED MAMMOGRAM REPORTING RECOMMENDATIONS



Source: Created for this white paper based on FDA (2019a).

EXHIBIT 3. EXAMPLE EVENT TREE FOR REVIEW OF PREMARKET TOBACCO PRODUCTS



Source: FDA (2019b)

4.4 MONTE CARLO SIMULATION

In Monte Carlo simulation, key inputs in a model are identified in terms of a probability distribution, rather than a point estimate. The computer draws a value for each input based on the distributions and calculates the output metrics of interest. The simulation repeats the process a defined number of times, selecting new values for each input, running through the calculations, producing new values for each output, and storing those values. Generally, analysts set the simulation to repeat at least a few hundred times; several thousand iterations are more common. Upon completion of the runs, the simulation calculates statistics based on the simulated output values. With the simulation drawing input values at random from the specified probability distributions, and by performing a statistically significant number of trials, the results then show the probability of achieving a certain threshold for the output metrics. The computer "plays the game" enough number of times to develop a good idea of the distribution of possible results.

Monte Carlo simulation has been used since at least the 1950s and specialized programs for performing complex simulations that integrate with spreadsheet models are available (e.g., Crystal Ball and @Risk). Statistical packages such as STATA have the capability for Monte Carlo simulations, and there is a Monte Carlo package for the R language that provides tools to create simulations quickly and easily.

Steps (within the existing cost or benefit model) include:

1. Define probability distributions for key input values;
2. Define the output variables;
3. Sample input values;
4. Collect statistics on output values based on the random inputs; and
5. Repeat the simulation N number of times and calculate statistics for output values.

While formulas exist to calculate the exact number of iterations to run based on the output parameter of interest, the standard deviation of the output values, and the percent confidence interval desired for the parameter, generally the simulation needs to be run enough times to achieve output distributions that approach smooth, continuous distributions (i.e., “not lumpy”). Additionally, if an analyst is interested in the 50th percentile of an output distribution, this value will reach a stable value far quicker than percentiles towards the tails (i.e., the 90th percentile of a distribution). If the tails of the output distribution are of concern, more iterations will be needed. Ideally, the simulations are ended when outputs converge to a pre-specified threshold or criteria (e.g., the mean of the cost estimate does not change with additional runs by more than a predetermined percentage).

A review of several recent RIAs produced by Federal agencies suggests analysts commonly employ 10,000 simulations. For example, in the 2017 Soy Relabeling RIA (US FDA 2017), analysts set the Monte Carlo simulation calculations estimating product relabeling costs to repeat 10,000 times. FDA used a similar number of iterations in the 2019 Premarket Tobacco Application and Recordkeeping Requirements Analysis (US FDA 2019b).

For a discussion of the types of probability distributions commonly used in Monte Carlo simulation, and when each may be most applicable, see Appendix A. For a demonstration of the use of several different distributions in a Monte Carlo simulation, and the implications of these different choices on the modeling results, see Appendix B.

Another challenge for analysts is to document and convey the details of a Monte Carlo simulation at a level that is understandable and explains the important assumptions without onerous specifics and technical details. The U.S. Department of Transportation’s (DOT) Pipeline and Hazardous Materials Safety Administration (PHMSA) used Monte Carlo analysis in its *Final Regulatory Impact Analysis for Enhanced Rail Cars for High-Hazard Flammable Trains* (PHMSA, 2015). Exhibits 4, 5, and 6 provide examples of helpful tables from this study documenting the details of the Monte Carlo simulation.

Exhibit 4 provides the details of the Input Distributions used in the simulation. Exhibit 5 shows an example of the outputs of the model, and Exhibit 6 describes the output distribution, in this case, the total estimated damages from higher-consequence events over 20 years.

EXHIBIT 4. INPUTS IN PHMSA'S MONTE CARLO ANALYSIS OF LIKELY BENEFITS (PHMSA 2015)

Table EB11. Assumed Distribution of Input Values for Monte Carlo Analysis

Risk Variables	Probability Distribution	Point Estimate	Low	Mode	High
Number of Events	Triangular*	2	50%	100%	250%
Percent of Population Fatalities in HHFT	Triangular*	14%	75%	100%	110%
Percent of Population Fatalities in HTUAs	Triangular*	10%	75%	100%	110%
Baseline Non-Fatality Damages in HHFT	Triangular*	$263,310,064	75%	100%	125%
Baseline Non-Fatality Damages in HTUA	Triangular*	$190,860,845	75%	100%	125%
Random Selection of Population Density	Custom**				
Wetlands	Triangular*	$14,654,932	75%	100%	125%
Wetlands Probability	Uniform	6%	2%		10%

*We chose a triangular distribution for most elements because of its ease of use, only 3 parameters (lower limit, an upper limit and a mode) and allows for skewed distributions, like the percent fatalities above. Additionally, with definite lower and upper limits we can avoid extreme values.

**The random selection of a population density is based on custom distributions with points given by the population densities in HTUAs and the rest of the HHFT rail network as illustrated in the table titled *Population Densities along Crude Oil and Ethanol Rail Routes*. The fuel network rail link lengths are used within HTUA/non-HTUA as a scaling factor to develop the population density custom distributions for HTUAs and non-HTUA networks, assuming the probability of having an accident is proportional to the link length. PHMSA uses the average annual tons transported on the fuel network to select in which area the accident occurs and is a proxy for the usage of the rail track in the HTUAs versus the rest of the HHFT network. Thus, the probability of an accident occurring in the HTUA portion of the network is given by the ratio between the average annual tons transported in the HTUA network to the total tons transported in the entire HHFT network:
$8{,}009{,}851{,}408 / 49{,}800{,}510{,}555.7 = 0.1608$.
HTUAs account for 16.08 percent of the HHFT network. In each run of the simulation, a new, random number between 0 and 1 is generated. If this number is less than 0.1608, the incident is modeled as if taking place in a HTUA; otherwise, it is modeled as if taking place outside of a HTUA.

EXHIBIT 5. PHMSA MONTE CARLO ANALYSIS SAMPLE OUTPUTS (PHMSA 2015)

Table EB12. Sample Outputs of the Risk Model Showing 20-Year Values for Fatality and Non-Fatality Damages

Trial values	# of Events	Percent Fatalities	Population Density per half-square km	Scaling Factor	Non-Fatality Damages* (million)	Damages Non-Fatality (million)	Damages Fatality (million)	Wetlands (million)	Total Damages (million)
1	1.7	11.4%	53.0	38.9%	$266.0	$172.0	$107.7	$0.0	$279.7
2	4.1	14.1%	8.1	6.0%	$279.2	$68.2	$50.3	$0.0	$118.5
3	3.6	12.4%	226.0	165.9%	$251.5	$1,517.2	$1,096.4	$0.0	$2,613.6
4	1.8	13.6%	169.3	124.2%	$245.1	$533.6	$433.9	$0.0	$967.5
5	1.8	12.9%	280.8	206.1%	$246.9	$917.8	$702.1	$0.0	$1,619.9
6	2.2	14.2%	82.7	60.7%	$294.9	$389.0	$274.1	$0.0	$663.2
7	2.7	11.2%	67.9	49.9%	$229.2	$307.7	$220.2	$0.0	$527.9
8	2.2	13.0%	23.7	17.4%	$251.2	$97.7	$74.1	$0.0	$171.8
9	3.5	12.2%	218.2	160.2%	$317.1	$1,797.1	$1,012.9	$0.0	$2,810.0
10	2.2	13.0%	166.6	122.3%	$228.4	$612.9	$509.4	$0.0	$1,122.4
11	2.0	14.6%	51.7	38.0%	$274.7	$203.9	$158.7	$15.8	$378.5
12	2.4	12.8%	49.4	36.3%	$240.5	$212.5	$165.1	$0.0	$377.6
13	3.1	13.5%	2.3	1.7%	$299.9	$15.6	$10.2	$15.2	$41.0
14	2.9	14.0%	0.3	0.2%	$304.2	$2.2	$1.5	$0.0	$3.6
15	3.9	13.0%	26.4	19.3%	$265.4	$197.9	$141.7	$0.0	$339.6
16	4.2	9.9%	0.0	0.0%	$198.2	$0.0	$0.0	$0.0	$0.1
17	1.3	13.6%	94.5	69.4%	$249.2	$226.8	$181.1	$0.0	$407.9
18	1.6	14.0%	27.4	20.1%	$271.6	$85.2	$64.2	$0.0	$149.4
19	2.3	10.8%	0.6	0.4%	$197.1	$1.9	$1.5	$0.0	$3.5
20	1.5	12.8%	0.2	0.1%	$300.7	$0.5	$0.3	$0.0	$0.9
21	2.0	9.4%	1,499.1	1100.3%	$194.8	$4,197.8	$2,960.5	$0.0	$7,158.3
22	2.5	11.7%	54.8	40.2%	$260.8	$265.6	$173.6	$0.0	$439.2
23	3.7	12.5%	640.5	470.1%	$273.9	$4,815.7	$3,222.0	$0.0	$8,037.6
24	3.2	13.8%	395.9	290.6%	$283.4	$2,667.3	$1,893.5	$0.0	$4,560.9
25	1.6	13.7%	493.1	361.9%	$272.9	$1,628.3	$1,191.6	$0.0	$2,819.9
26	1.3	12.2%	182.0	133.5%	$271.3	$475.1	$311.8	$0.0	$786.9
27	3.4	14.5%	406.1	298.1%	$257.0	$2,640.7	$2,175.7	$0.0	$4,816.4
28	4.0	9.0%	4,028.9	2957.1%	$162.7	$19,224.2	$15,454.4	$0.0	$34,678.7
29	1.9	11.8%	25.9	19.0%	$224.7	$81.0	$62.3	$0.0	$143.4
30	4.6	12.6%	429.8	315.5%	$281.5	$4,093.1	$2,677.3	$0.0	$6,770.5

Note: *The values represent the Lac-Mégantic baseline non-fatality damages and vary randomly according to the triangular distribution of non-fatality damages set forth in table titled *Assumed Distribution of Input Values for Monte Carlo Analysis*.

EXHIBIT 6. PHMSA MONTE CARLO ANALYSIS RESULTS (PHMSA 2015)

Table EB13. Distribution of Total Estimated Damages From Higher-Consequence Events Over 20 Years

Percentile	# of Events	Percent Fatalities	Population Density per half-square km	Scaling Factor relative to Lac-Mégantic	Damages Non-Fatality (million)	Damages Fatality (million)	Wetlands (million)	Total Damages (million)
0%	1.0	7.5%	0.0	0.0%	$0	$0	$0	$0
5%	1.5	9.2%	0.0	0.0%	$0.0	$0.0	$0.0	$0.1
10%	1.6	9.9%	0.2	0.1%	$0.6	$0.4	$0.0	$1.3
15%	1.8	10.5%	0.8	0.6%	$3.4	$2.5	$0.0	$8.7
20%	1.9	11.2%	3.0	2.2%	$13.8	$10.1	$0.0	$24.7
25%	2.0	11.7%	7.5	5.5%	$35.5	$26.3	$0.0	$62.8
30%	2.1	12.0%	14.7	10.8%	$70.9	$51.7	$0.0	$124.3
35%	2.2	12.3%	25.4	18.7%	$117.5	$85.2	$0.0	$203.7
40%	2.3	12.6%	37.8	27.8%	$182.6	$131.1	$0.0	$315.4
45%	2.4	12.8%	54.8	40.2%	$262.9	$190.5	$0.0	$452.6
50%	2.5	13.0%	77.0	56.5%	$360.8	$266.1	$0.0	$631.9
55%	2.7	13.1%	105.1	77.1%	$501.4	$361.7	$0.0	$872.7
60%	2.8	13.3%	137.2	100.7%	$668.5	$493.8	$0.0	$1,173.4
65%	3.0	13.5%	182.8	134.2%	$883.5	$639.5	$0.0	$1,521.5
70%	3.1	13.6%	232.3	170.5%	$1,152.8	$847.6	$0.0	$1,997.1
75%	3.3	13.8%	308.3	226.3%	$1,517.2	$1,095.1	$0.0	$2,621.8
80%	3.5	13.9%	410.6	301.4%	$1,997.6	$1,455.1	$0.0	$3,460.0
85%	3.7	14.1%	554.7	407.2%	$2,792.9	$2,039.2	$0.0	$4,843.7
90%	3.9	14.3%	827.8	607.6%	$4,206.6	$3,053.3	$0.0	$7,241.5
95%	4.2	14.5%	1,532.7	1125.0%	$7,292.1	$5,307.7	$13.1	$12,618.0
100%	5.0	15.2%	23,415.0	17186.2%	$111,443.4	$86,982.0	$18.2	$198,436.5

A final recommendation for developing a correct Monte Carlo simulation is to conduct all analysis within a common Monte Carlo "shell" that samples from the same set of uncertain performance parameters, propagates them through the suite of analyses, and collects the resulting performance measures as a vector of performance measure values. As the Monte Carlo shell iterates, these performance measure vectors accumulate in accordance with the probability distributions that are defined over the set of performance measures. In some cases, when separate Monte Carlo simulations are used, using the results of one simulation in the next can overstate the uncertainty of the effect of the rule, and these figures would carry through into cost and benefit estimates, again overstating the uncertainty.

Advantages

- Monte Carlo simulations can be a very flexible modeling tool that can accommodate many varying parameters and thus model a range of possible outcomes.
- They are relatively simple and intuitive.
- They can provide a quantified range of outcomes for situations with significant uncertainty.

Limitations

- The results are only as good as the model's inputs and assumptions, so analysts need to worry about the concept of "garbage in-garbage out." In other words, the results of an analysis are only good as the quality of the data and assumptions that go into the model. This limitation is best mitigated with clear details of the inputs and outputs as demonstrated in Exhibits 4 through 6.
- Relatedly, a simulation analysis may provide a false sense of rigor.
- It is not always obvious how many trials are needed for the simulation to converge, and it may require many computations to approximate a solution.

4.5 SENSITIVITY ANALYSIS

Sensitivity analysis can be helpful in determining which variables in an analysis are key sources of uncertainty. Sensitivity analysis performs best when it primarily focuses on one or a few variables at a time, keeping the remaining values constant.

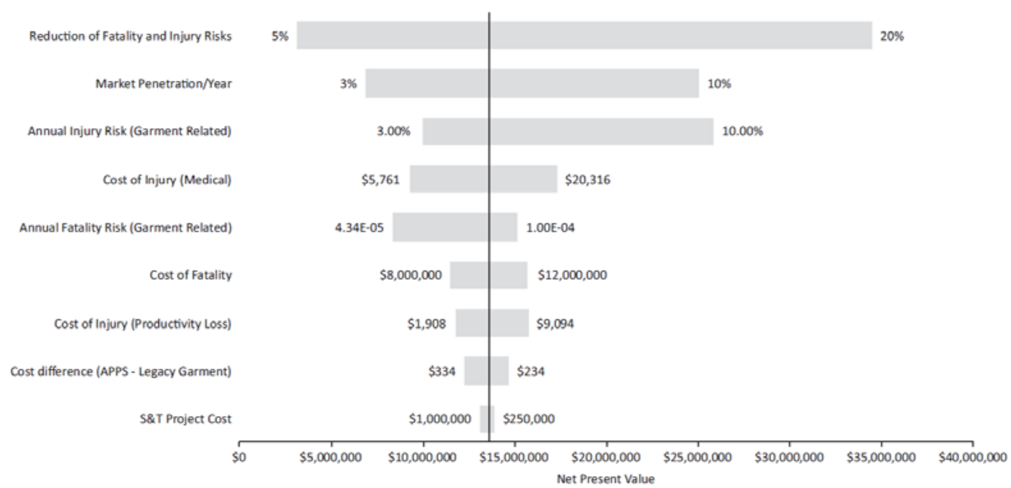
Sensitivity analysis tries to answer the question: "what makes a difference in this decision?" (Clemen 1996) A well done sensitivity analysis helps analysts confirm that they are solving the right problem and focusing on the most important uncertainties impacting the analysis. Identifying the most important uncertainties is critical information to provide to decision-makers.

A one-way sensitivity analysis is the most common approach for determining which variables have the biggest impact on the results of an analysis. Analysts insert different values for a single variable, re-running the model to calculate the results associated with each assumption. All other variables remain unchanged, so the analyst can clearly see the impact of the variable in question. This type of sensitivity analysis can be accomplished using simple spreadsheet tools or Monte Carlo simulation tools (see Section 4.4).

Where several variables are uncertain, analysts might also conduct a sensitivity analysis that considers different possible values for these variables simultaneously. This type of sensitivity analysis is most frequently accomplished using Monte Carlo simulation tools. A common way of displaying output that considers the sensitivity of a result to more than one variable is with a "tornado diagram" (Clemen 1996, Leach 2006).

An example of a tornado diagram is provided in Exhibit 7. This example is from Thrift and von Winterfeldt (2021)'s study examining the benefits and costs of DHS's investments in APPS for wildland firefighters. The goal of the APPS was to reduce risk and improve comfort for the wildland firefighter. The analysis estimated that the expected net benefits of the investment in terms of reduced fatalities and injuries would be $13.6 million (present value in 2019 dollars). However, as can be seen in Exhibit 7, the estimate of the expected value, depicted in this figure as a vertical line, was highly sensitive to several model parameters. The sensitivity analysis showed that net benefits could range from a 5th percentile estimate of $6.4 million to a 95th percentile of $43.7 million. The tornado diagram in Exhibit 7 shows the large range was due principally to the uncertainty about the reduction of fatality and injury risks, which analysts estimate could range from 5 percent to 20 percent of baseline risk, and the market penetration rates of the new garments, potentially ranging from 3 percent to 10 percent of the market. Given the importance of these two parameters, the authors provide additional information for decision-makers about the break-even values for each parameter and the plausibility of achieving these values.

EXHIBIT 7. EXAMPLE TORNADO DIAGRAM (THRIFT AND VON WINTERFELDT, 2021)



Advantages

- Identifying the critical variables in a sensitivity analysis helps analysts focus on the aspects of the model that matter (i.e., can change the recommendation). This information allows the decision-makers to focus on the plausibility of key assumptions.
- Analysts can examine the critical variables identified in the sensitivity analysis to determine if there is value in collecting more information about these variables. While the stylized value of information calculation described in Section 6 in many cases may not be possible, the process of considering both the costs of collecting the information and the benefits to the analysis in terms of potential changes for decision-making can provide important insights that may or may not justify additional information collection.

Limitations

- No optimal sensitivity analysis exists and to some extent determining the best sensitivity analysis is an art. For example, conducting sensitivity analysis of too many variables with similar effects on outputs may overwhelm or confuse decision-makers.
- Tornado diagrams are easiest to create and understand if variables are examined one at a time holding others constant, but this approach needs to be carefully implemented if variables are correlated. In cases where correlated variables make

the results of the one-way sensitivity analysis non-sensical, consider grouping correlated variables in a single bar.

5.0 UNCERTAINTY ELICITATION AND THE ROLE OF SUBJECT MATTER EXPERTS

Section 4.0 focuses on tools used to (1) identify key uncertainties in models estimating the benefits and costs of proposed regulations and (2) quantify uncertainty associated with model results. In this section, we explain how expert judgment can be used to fill data gaps and better characterize uncertainty associated with key parameters. In some cases, such judgment may be needed when a systematic review of the empirical literature finds that the available research is of poor quality or addresses situations that differ in important ways from the context addressed by the policy. In other cases, the review may find that the value needed by analysts has not been previously studied. In either situation, parameter estimates may be necessary to meet the requirements for conducting regulatory analysis.

Approaches for soliciting information from experts range from unstructured interviews to formal, structured expert elicitation, with semi-structured interviews, focus groups, expert panel meetings, and other approaches falling in between. More sophisticated approaches are likely to improve the validity of the estimates, but require more time and resources. At one extreme, informal phone interviews are relatively quick and inexpensive, but the resulting data may be less reliable. At the other extreme, formal structured expert elicitation could require significant time and effort to implement if the assessment task is complex.

There is no single, best approach for obtaining expert judgments. Rather, various approaches might be characterized as occurring on a continuum, with unstructured interviews being the most informal approach and formal, structured expert elicitation representing the opposite end of the spectrum. The criteria used to select the best approach for a given analysis include what is at stake with the rulemaking, as well as the amount and quality of data and resources available.

In the remainder of this section, our discussion focuses on best practices for conducting formal, structured elicitation.[9] Structured expert elicitation differs from less formal approaches in that it follows a systematic framework for obtaining experts' judgments about the value of a clearly defined source of uncertainty, where each expert is asked to provide a probability distribution (or more likely key parameters that can be fit to a probability distribution) characterizing his or her beliefs about the uncertainty. The types of issues considered in each step of such an elicitation can also inform the design and conduct of other elicitation approaches.

The literature describing best practices for eliciting expert judgments, including relevant choices and methods, along with their strengths and weaknesses, is extensive. Furthermore, consensus on key points, such as whether and how to combine judgments

[9] For a discussion of less formal approaches, such as semi-structured interviews and focus groups, see Newcomer et al. (2015).

from different experts, does not exist. Useful general references on structured expert elicitation include:

- *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Morgan and Henrion 1990);
- *Uncertain Judgements: Eliciting Experts' Probabilities* (O'Hagan et al. 2006);
- *Making Hard Decisions: An Introduction to Decision* Analysis (2nd ed.) (Clemen 1996); and
- *Expert Elicitation Task Force White Paper* (U.S. Environmental Protection Agency 2009).[10]

The general steps followed in a structured expert elicitation include:

1. **Develop background information and prepare for the elicitation.** This step includes identifying and clearly defining the variables for which expert judgment is needed; choosing the elicitation method; identifying the staffing needs; developing the elicitation protocol, and preparing a briefing book for the experts.
2. **Identify and Recruit experts.** This steps involves determining the number and type of experts required for the elicitation, and choosing and implementing a methodology for identifying and selecting experts.
3. **Conduct the elicitation.** In this step, analysts motivate and train the experts to provide judgments in terms of probability distributions and to avoid cognitive biases. Then, they conduct the elicitation.
4. **Report the results.** Analysts report judgments provided by each expert and make decisions and whether and how to combine estimates across experts.
5. **Document and verify the process.** Finally, analysts keep detailed records of the development and implementation of the elicitation and might consider conducting a peer review of the elicitation process.

In Appendix C, we provide additional detail regarding key considerations in each step.

The EPA's *Second Section 812 Benefit-Cost Analysis of the Clean Air Act* provides an example of the use of expert elicitation in an assessment of the costs and benefits of a series of Federal regulations.[11] The relationship between decreases in fine particulate matter ($PM_{2.5}$) and reductions in mortality was a key source of uncertainty in EPA's estimation of the benefits of the Act. The agency elicited distributions from 12 of the world's leading experts on this relationship. The results of the elicitation are shown in

[10] Additionally, the Nuclear Regulatory Commission (NRC), the U.S. Department of Energy, and the Electric Power Research Institute sponsored a study to develop recommendations for conducting expert assessments in the context of natural hazards. The resulting Senior Seismic Hazard Analysis Committee (SSHAC) process is described in NRC (1997) and NRC (2012), and updated in NRC (2016).

[11] See EPA (2011) for an overview of the benefit-cost analysis, IEc (2011) for a detailed discussion of the uncertainty analyses supporting the Second Section 812 Benefit-Cost Analysis, and Roman et al. (2008) for the peer-reviewed publication of the expert elicitation.

Exhibit 8. Each expert provided minimum, 5th, 25th 50th, 75th, 95th and maximum effect estimate, as shown in the box and whisker plots. The agency used to the results of the elicitation to inform the selection of a primary concentration-response function to estimate avoided premature mortality (the benefits) of reductions in exposure to $PM_{2.5}$, and in its quantitative uncertainty analysis.

EXHIBIT 8. UNCERTAINTY DISTRIBTUTIONS FOR THE $PM_{2.5}$ MORTALITY CONCENTRATION-RESPONSE COEFFICIENT PROVIDED BY EXPERTS (ROMAN ET AL 2008)

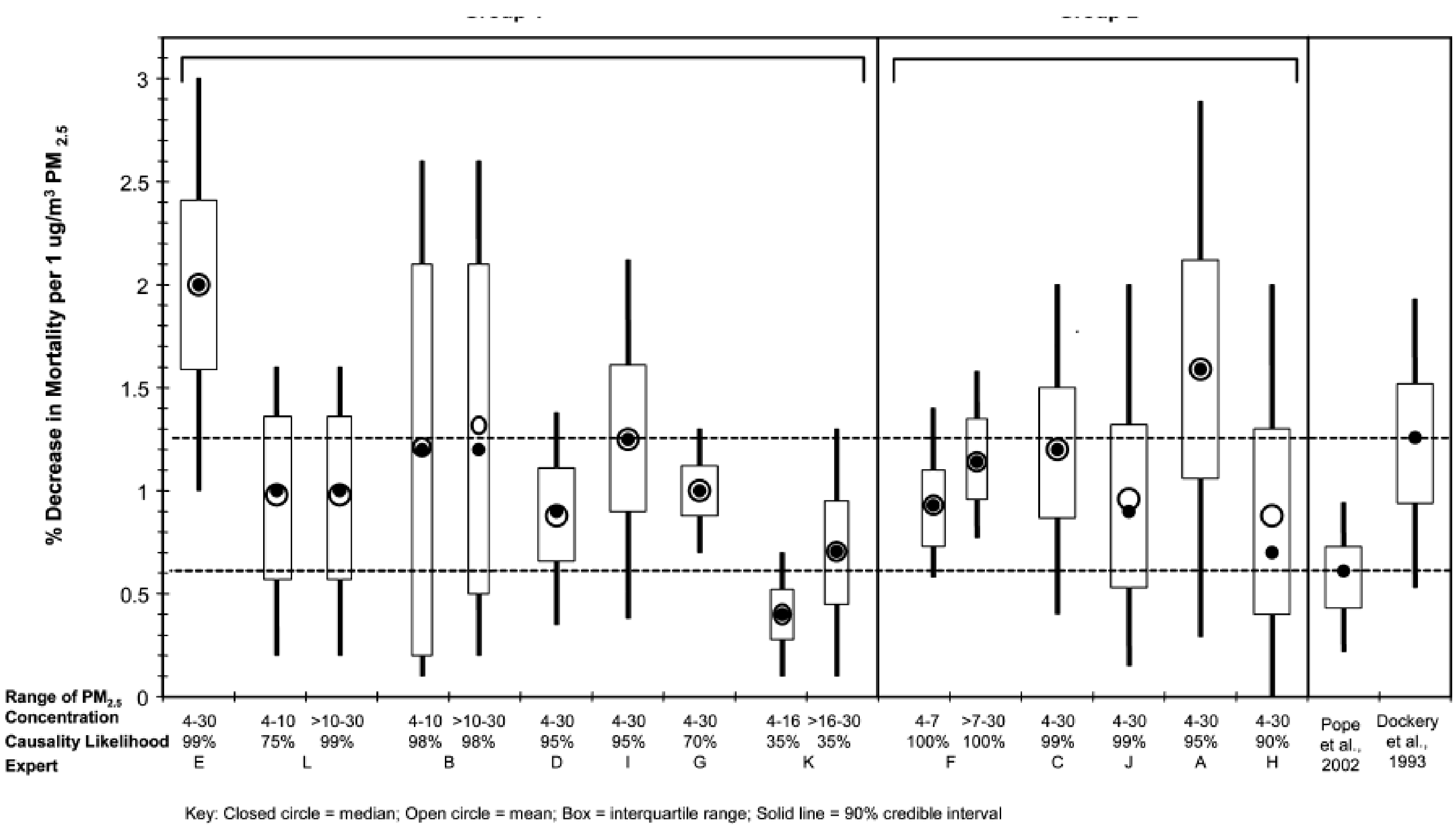


FIGURE 3. Uncertainty distributions for the $PM_{2.5}$–mortality *C*–*R* coefficient for annual average $PM_{2.5}$ concentrations of 4–30 $\mu g/m^3$
Note: Box plots represent distributions as provided by the experts to the elicitation team. Experts in group 1 preferred to give conditional distributions and keep their probabilistic judgment about the likelihood of a causal or noncausal relationship separate. Experts in group 2 preferred to give distributions that incorporate their likelihood that the $PM_{2.5}$–mortality association may be noncausal. Therefore, the expert distributions from these two groups are not directly comparable.

6.0 VALUE OF INFORMATION CALCULATIONS

When substantial uncertainty exists for a variable in a model, sometimes better information about the variable could increase one's understanding of the net benefits of the regulation. There is always a trade-off between a timely, cost-effective analysis and the inclination to continue to acquire better information to support decision-making. The challenge becomes identifying when investments in information collection are warranted because the benefits to the improvement of the analysis exceed the costs of the additional data collection.

In the context of a regulatory decision, it is helpful to understand how much benefit can be derived from the collection of additional data (e.g., exposure data) that will help elucidate the consequences associated with alternative regulatory policies. The value of information is calculated as the expected monetary difference between the estimate of the net social benefits of the regulation both with and without the additional information on the uncertainty (in this example, exposure data).

Further information is most likely to be valuable and worth additional collection when considerable uncertainty exists and uncertainty about the variable of interest could lead to different regulatory choices. For example, if there are high costs associated with a proposed rule and enormous uncertainty about the true impacts for society, then improving one's understanding of the uncertainties is often justifiable based on a value of information calculation.

It is common to depict a value of information calculation with a decision tree where the collection of information reverses the order of decision and chance nodes. Exhibit 9 replicates a stylized value of information calculation described in work for Health Canada (Leighty et al., 2000). In this example, there are three policy options: (1) take no further regulatory action (top branch), (2) institute specific pollution control measures based on the current information (middle branch), and (3) acquire additional information, which in this case is a database of human exposures to a particular toxin prior to deciding whether to adopt control measures (bottom branch).

The uncertainty is the current level of human exposure to a particular toxin and is shown with three possible outcomes: high, medium, and low. To evaluate the decision tree, probabilities must be assigned to each of the exposure states and payoffs must be estimated for each scenario or path through the tree at the endpoints. The value of information in this example estimates how valuable collecting this national database of exposures would be. If the database were collected, then the decision-makers could understand the likely exposures before they needed to make the control decisions.

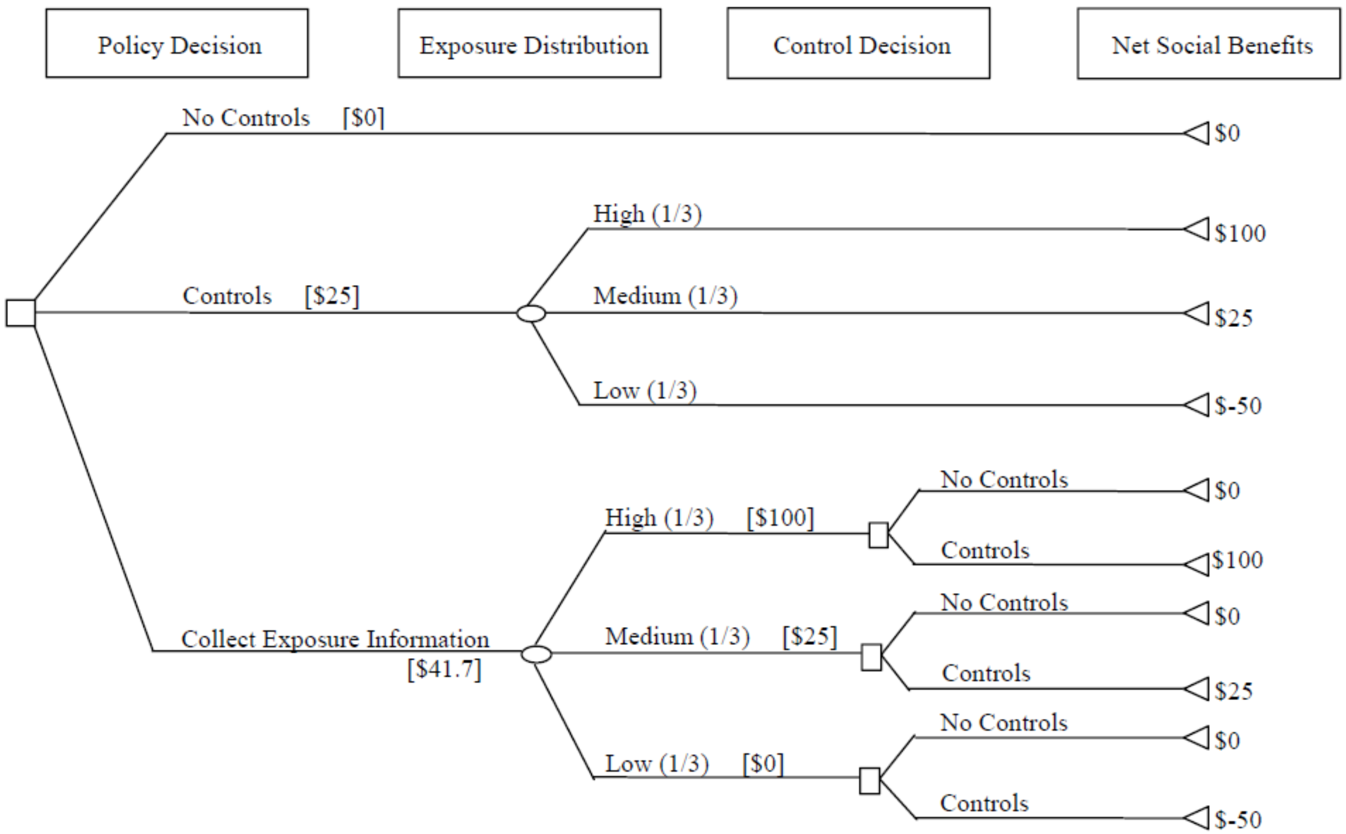
More complex analyses may consider the value of collecting information to reduce uncertainty in one component of a decision that involves many uncertain steps (i.e., the value of partial information). In this example, the value of implementing no controls is just the status quo, no net social benefits. In the control branch, there is uncertainty about the exposure distribution; if exposure is high, there are significant positive net social benefits. If exposure is low, the cost of the controls generates negative net social benefits, and if exposure is medium, there are net positive social benefits, but they are less than in the high condition. In this example, the probability of high, medium, and low exposure existing are equally likely. Creating an exposure database means knowing the exposure (high, medium, and low) before making the control decision and only choosing to implement the controls if the exposure is high or medium. Comparing the expected value of the middle branch to the bottom branch has an expected value of information of $16.7 units, which is greater than $0 and probably justifies the collection of the exposure information.

The approach used to collect more information to reduce uncertainty will vary depending on the data gaps. Analysts might undertake statistical analysis of existing datasets or spend additional time developing a broader literature review. As discussed in Section 5, subject matter experts can offer professional judgement to characterize the likely impacts, but this approach requires time and other resources. In some cases, conducting a well-designed survey can resolve uncertainty over respondents' preferences. An agency could also initiate a pilot study of a regulatory approach under consideration as an initial step before considering broader action. When considering the value of collecting this information in regulatory decision-making, it is appropriate to balance the potential benefits of reductions in uncertainty in the analysis with the costs of the collection, including expenditures to acquire data or to field a survey, as well as the effect of delaying regulatory action.

To learn more about VOI analysis, see the following references:

- *Decision Analysis: Introductory Lectures on Choices Under Uncertainty* (Raiffa 1968); and
- *Making Hard Decisions: An Introduction to Decision* Analysis (2^{nd} ed.) (Clemen 1996).

EXHIBIT 9. FRAMEWORK FOR VALUE OF INFORMATION: DECISION TREE



Expected Value of Perfect Information = ($41.7 - $25) = $16.7

*Numbers in parentheses are probabilities; numbers in brackets are expected values of net benefits at decision or chance nodes

Source: Leighty et al. (2000).

7.0 COMMUNICATION OF UNCERTAINTY

Clear, concise communication of uncertainty is critical for decision-making. At a minimum, every RIA should include a qualitative discussion of any significant sources of uncertainty. Depending on the stakes of a regulation (see Section 3.0), quantitative uncertainty analysis may also be warranted or necessary. Below, we provide examples from actual RIAs illustrating the communication of all of the types of analyses discussed in this white paper.

7.1 QUALITATIVE PRESENTATION OF UNCERTAINTY

Quantification of uncertainty is desired because it provides a clear indication of the likely direction and magnitude of impacts. However, if quantification is not possible, analysts must determine how to convey the uncertainty qualitatively. Additionally, they need to consider how to address potentially important non-quantified effects. Ignoring such effects may lead to poor decisions; but overemphasizing minor sources of uncertainty may also negatively impact decision-making. Clear presentation of uncertainty is needed to ensure potential sources of bias or uncertainty are appropriated weighted and considered.

At a minimum, every RIA should include a qualitative discussion of key sources of uncertainty and their potential influence on the results of the analysis. This discussion can be included in sections titled, "Limitations and Key Sources of Uncertainty" at the end of each chapter in an RIA, in a separate appendix, or both. Because RIAs often rely on dozens of assumptions and values, listing all of the uncertainties can be overwhelming. To assist decision-makers and the public, analysts should use tools, like tables and graphics, to highlight important effects, to ensure they are not overlooked. These tools should also help decision-makers understand which uncertainties are less likely to materially affect the results.

There are many ways to convey this information. Below, in Exhibit 10, we list several examples of benefit-cost analyses of Federal regulations providing high-quality discussions of uncertainty or illustrating useful approaches to addressing uncertainty that can be adopted for other analyses. For each example, we provide hyperlinks to the documents, which are all publicly-available via the internet, and a description of why we included them in this discussion.[12] This list is not exhaustive; there are hundreds of RIAs providing additional, excellent examples.

In summary, best practices for a qualitative discussion of uncertainty include:

---

[12] In some cases, a single document provides multiple examples of the presentation of uncertainty that are relevant to this white paper. In the exhibit, we attempt to highlight a range of different presentation techniques, and therefore generally focus on one to two examples in each document. In each case, we encourage readers to skim the entire document to see additional relevant examples of ways to present both quantified and non-quantified uncertainty.

- Clearly discuss key assumptions, uncertainties, and data limitations and make sure this discussion appears along with the presentation of the primary estimates of benefits and costs;
- Describe whether the assumptions could lead to overstating or understating benefits or costs and the likely direction of the potential bias;
- Discuss the potential magnitude of each assumption's impact on the analysis' estimates (e.g., whether it is major or minor); and
- Wherever possible, use a presentation format (e.g., colors or graphics) that helps readers easily comprehend the overall takeaway message on uncertainty.

EXHIBIT 10. EXAMPLES OF QUALITATIVE PRESENTATION OF UNCERTAINTY

TITLE/HYPERLINK	AGENCY	SECTION(S)	RELEVANCE
The Benefits and Costs of the Clean Air Act from 1990 to 2020: Final Report – Rev. A (EPA 2011)	EPA	Uncertainty discussion at the end of each chapter	This report to Congress estimates the costs and benefits of the regulations promulgated under the Clean Air Act. Each chapter presents a step in the overall analysis. At the end of each chapter, EPA provides a table listing key uncertainties associated with the analysis presented in that section. For each uncertainty, it discusses the potential direction of bias on the results and the relative significance of the uncertainty. Importantly, EPA's Science Advisory Board (SAB) provided substantial advice during the development of this report, and EPA's presentation of uncertainty was well-received by the expert reviewers.
Regulatory Impact Analysis: Mandatory Advance Electronic Data (AED) for International Postal Shipment Interim Final Rule (IEc 2021)	CBP	Appendix B	This interim final rule is intended to lessen the flow of illegal opioids into the United States. The report includes a comprehensive discussion of uncertainty in a separate appendix. The analysts add color to help draw out key conclusions regarding the relative importance of various sources of uncertainty.[1]
Food Labeling; Gluten-Free Labeling of Fermented or Hydrolyzed Foods Regulatory Impact Analysis (Final Rule) (FDA undated)	FDA	Table 1	*Circular A-4* (OMB 2003) requires agencies to provide an accounting statement in its RIAs summarizing monetized, quantified, and qualitative categories of costs and benefits. In this example, FDA integrates the presentation of its qualitative results (including the results of a break-even analysis) with its quantitative estimates.[2]
36 CFR 51 Concessions Contracts Revisions: Regulatory Impact Analysis and Initial Regulatory Flexibility Analysis (IEc 2020)	NPS	Exhibits 4-1 and 4-1	This proposed rule is intended to improve the way NPS solicits, evaluates, and administers contracts with firms providing services to visitors within National Park System units. Significant uncertainty exists about how concessioners are likely to respond to the proposed changes. The analysts use decision trees to map the possible outcomes and provide qualitative descriptions of the potential positive and negative consequences of each scenario.

Notes:
1. This appendix also includes quantitative sensitivity analysis addressing two key sources of uncertainty highlighted in the qualitative discussion.
2. This report quantifies the primary benefits of the proposed rule and uses break-even analysis to characterize other, non-quantified benefits.

7.2 QUANTIFIED ESTIMATES OF UNCERTAINTY

Ideally, analysts can characterize key sources of uncertainty quantitatively. In Exhibit 11, we provide examples of different ways to present the results of these types of analyses. This list is not exhaustive; readers are encouraged to look at other examples referenced throughout this white paper.

In summary, best practices for a quantitative discussion of uncertainty include:

- Clearly discuss key assumptions and uncertainties and analyze them (e.g., using sensitivity analysis) to determine which are likely to have the largest impact on the benefits or costs in the analysis;
- Focus the quantitative analysis on the most important sources of uncertainty, rather than on the parameters for which distributions are readily available. Doing so will help to avoid misconceptions about the true range of confidence intervals around estimates of costs and benefits;
- Choose the tool that will best characterize uncertainty. For example, model uncertainty may be better characterized using scenario analysis, while uncertainty associated with specific model parameters might be addressed using Monte Carlo simulations;
- Clearly describe the modeling approach, including the rationale for the choice of probability distributions used in the modeling, the number of draws if a simulation is used, and any correlations among variables that are included;
- If possible, quantify the statistical distribution of estimated impacts and provide key characteristics of the distribution including mean, median, standard deviation, variance, minimum, maximum, 5th percentile, and 95th percentile; and
- Describe any data limitations, including key non-quantified costs and benefits, along with the presentation of quantitative results.

EXHIBIT 11. EXAMPLES OF QUANTITATIVE PRESENTATION OF UNCERTAINTY

TITLE/HYPERLINK	AGENCY	SECTION(S)	RELEVANCE
Uncertainty Analyses to Support the Second Section 812 Benefit-Cost Analysis of the Clean Air Act (IEc 2011)	EPA	Entire report	This report supplements the main analysis described in Exhibit 10. The EPA conducted nearly a dozen separate quantitative uncertainty analyses for key steps in its analysis of the costs and benefits of the Clean Air Act. The results of sensitivity and scenario analyses are primarily provided in tabular form; however, it also uses different types of charts to illustrate relative differences across scenarios. The agency also used expert elicitation and presents the experts judgments using box and whisker plots. Finally, it uses mapping tools to highlight differences in pollution levels across the United States in response to different scenarios.[1]
Potential Public Health Effects of Reducing Nicotine Levels in Cigarettes in the United States (Apelberg et al. 2018) & How Could Lowering Nicotine Levels in Cigarettes Change the Future of Public Health? (FDA 2020)	FDA	Figure 1 in paper and figures on webpage	FDA issued an Advanced Notice of Proposed Rulemaking (ANPRM) seeking comments on standards lowering the nicotine level in cigarettes. It cites a study it funded, published in the *New England Journal of Medicine*, that relied on statistical models and expert elicitation to estimate potential health impacts. The study and the website provide graphics illustrating analysts' best estimates of various health improvements, with shading to indicate uncertainty bounds for each projection.
Regulatory Assessment and Initial Regulatory Flexibility Analysis for the Interim Final Rule: Importer Security Filing and Additional Carrier Requirements (IEc 2008)	CBP	Appendix C	CBP promulgated an interim final rule intended to help the agency identify high-risk ocean shipments to prevent smuggling and ensure cargo safety and security. Because the regulation had the potential to result in costs exceeding $1 billion in a single year, the agency performed probabilistic uncertainty analysis. The appendix includes exhibits describing the parameter values used in the Monte Carlo analysis, illustrating the results, and illustrating the sensitivity of the results to key assumptions.
Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Roder 12866 (Interagency Working Group on Social Cost of Carbon, United States Government 2013)	Interagency Working Group on the Social Cost of Carbon	Entire document	This Technical Support Document provides estimates of the social cost of carbon for use in U.S. regulatory analysis based on the output of three different integrated assessment models (IAMs) and three different discount rates. Each of the IAMs is subject to key assumptions and sources of uncertainty, as is the choice of discount rate. The presentation of recommended values, and how they should be used by analysts, is simple and clear. The document also provides a comprehensive discussion of assumptions and decisions made by the working group to develop these recommendations.

Notes:

1. This supplemental volume also includes the types of qualitative discussion of uncertainty discussed in Exhibit 10.

8.0 REFERENCES

Apelberg, B.J., S.P. Feirman, E. Salazar, C. G. Corey, B. K. Ambrose, A. Paredes, E. Richman, S. J. Verzi, E.D. Vugrin, N.S. Brodsky, and B.L. Ronstron. 2018. "Potential Public Health Effects of Reducing Nicotine Levels in Cigarettes in the United States." *New England Journal of Medicine*. 378:1725-1733.

Clemen, R.T. 1996. *Making Hard Decisions: An Introduction to Decision Analysis (2nd ed.)*. Belmont, CA: Duxbury Press at Wadsworth Publishing Company.

Clemen, R.T. and R.L. Winkler. 1999. "Combining Probability Distributions From Experts in Risk Analysis." *Risk Analysis.* 19:187-203.

Clinton, W.J. 1993. "Executive Order 12866: Regulatory Planning and Review." *Federal Register*. 58(190):51735-51744. http://www.whitehouse.gov/omb/inforeg_regmatters

Cooke, R.M. 1991. *Experts in Uncertainty: Opinion and Subjective Probability in Science*. New York, NY: Oxford University Press.

Cooke, Roger M. and L.H.J. Goossens. 1999. *Procedures Guide for Structured Expert Judgment.* Prepared for the Commission of European Communities Directorate-general XI (Environment and Nuclear Safety).

European Food Safety Authority. 2014. "Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment." *EFSA Journal* 12(6):3734. Available at http://www.efsa.europa.eu/en/efsajournal/pub/3734.htm

Goodwin, P. and G. Wright. 2004. "Chapter 9: Biases in Probability Assessment." *Decision Analysis for Management Judgment (3rd Ed.)*. Chichester, England: John Wiley & Sons.

Hammitt, J. 2011. *Expert Judgment (Expert Elicitation)*. Presentation at Society of Risk Analysis Workshop "Synthesizing Evidence: An Introduction to Systematic Reviews, Meta-Analysis, and Expert Elicitation," December 4, 2011.

Hora, S. 2007. "Eliciting Probabilities From Experts." In W. Edwards, R.F. Miles Jr., and D. von Winterfeldt (Eds.), *Advances in Decision Analysis: From Foundations to Applications*. New York, NY: Cambridge University Press.

Hora, S., N. Dodd, and J. Hora. 1993. "The Use of Decomposition in Probability Assessments of Continuous Variables." *Journal of Behavioral Decision Making.* 6(2):133-147.

Industrial Economics, Incorporated. 2021. *Regulatory Impact Analysis: Mandatory Advance Electronic Data (AED) for International Postal Shipment Interim Final Rule.* Final report prepared for U.S. Department of Homeland Security, U.S. Customs and Border Protection. February 24, 2021. Available at https://www.regulations.gov/document/USCBP-2021-0009-0002

Industrial Economics, Incorporated. 2020. *36 CFR 51 Concessions Contracts Revisions: Regulatory Impact Analysis (RIA) and Initial Regulatory Flexibility Analysis (IRFA).* Draft report prepared for the National Park Service, Commercial Services Program. June 16, 2020. Available at https://www.regulations.gov/document/NPS-2020-0003-0006

Industrial Economics, Incorporated. 2012. *Economic Analysis of Critical Habitat Designation for the Northern Spotted Owl.* Draft Report prepared for the U.S. Fish and Wildlife Service. May 29, 2012. Available at https://www.regulations.gov/document/FWS-R1-ES-2011-0112-0031

Industrial Economics, Incorporated. 2011. *Uncertainty Analyses to Support the Second Section 812 Benefit-Cost Analysis of the Clean Air Act.* Final report prepared for James DeMocker, Office of Air and Radiation, U.S. Environmental Protection Agency. March 2011. Available at https://www.epa.gov/sites/default/files/2016-01/documents/uncertaintyfullreport.pdf

Industrial Economics, Incorporated. 2008. *Regulatory Assessment and Final Regulatory Flexibility Analysis for the Interim Final Rule: Importer Security Filing and Additional Carrier Requirements.* Final report prepared for the U.S. Department of Homeland Security, U.S. Customs and Border Protection. November 6, 2008. Available at https://www.regulations.gov/document/USCBP-2007-0077-0200

Interagency Working Group on Social Cost of Carbon, United States Government. 2013. *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866.* Revised November 2013. Available at https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/inforeg/technical-update-social-cost-of-carbon-for-regulator-impact-analysis.pdf

Knol, A.B., P. Slottje, J.P. van der Sluijs, and E. Lebret. 2010. "The Use of Expert Elicitation in Environmental Health Impact Assessment: A Seven Step Procedure." *Environmental Health.* 9(19).

Leach, P. 2006. *Why Can't You Just Give Me The Number? An Executive's Guide to Using Probabilistic Thinking to Manage Risk and to Make Better Decisions*. Probabilistic Publishing.

Leighty, W., J. Neumann, H. Roman, E. Ruder, and T. Walker. 2000. "Value of Information Analyses," Memorandum prepared by Industrial Economics, Incorporated for Health Canada.

Makridakis, S., R. Hogarth, and A. Gaba. 2009. *Dance with Chance*. Oxford: One World Book.

Morgan, M. G. 2014. "Use (and Abuse) of Expert Elicitation in Support of Decision Making for Public Policy. *PNAS* 111(20): 7176-7184.

Morgan, M.G. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge, UK: Cambridge University Press.

Newcomer, K.E., H.P. Hatry, and J.S. Wholey. 2015. *Handbook of Practical Program Evaluation (4th Ed.).* San Francisco, CA: Jossey-Bass.

O'Hagan, A., C.E. Buck, A. Daneshkhah, J. Eiser, P. Garthwaite, D. Jenkinson, J. Oakley, T. Rakow. 2006. *Uncertain Judgements: Eliciting Experts' Probabilities*. Chichester, England: John Wiley & Sons Ltd.

Obama, B. 2011. "Executive Order 13563: Improving Regulation and Regulatory Review." *Federal Register*. 76(14):3821-3823. Available at http://www.whitehouse.gov/omb/inforeg_regmatters

Paté-Cornell, M.E. 1996. "Uncertainties in risk analysis: Six levels of treatment." *Reliability Engineering and System Safety*. 54:95-111.

Raiffa, H. 1968. *Decision Analysis – Introductory Lectures on Choices Under Uncertainty*. Longman Higher Education Division (a Pearson Education Company).

Roman, H.A., J.K. Hammitt, T.L. Walsh, and D.M. Stieb. 2012. "Expert Elicitation of the Value Per Statistical Life in an Air Pollution Context." *Risk Analysis*. 32: 2133-2151.

Roman, H.A., K.D. Walker, T.L. Walsh, L. Conner, H.M. Richmond, B.J. Hubbell, and P.L. Kinney. 2008. "Expert Judgment Assessment of the Mortality Impact of Changes in Ambient Fine Particulate Matter in the U.S." *Environmental Science and Technology* 42: 2268-2274.

Schoemaker, P. J.H. 1995. "Scenario planning: a Tool for Strategic Thinking." *Sloan Management Review*. 36(2):25-40.

Schoemaker, P. J.H. 1991. "When and How to Use Scenario Planning: a Heuristic Approach With Illustration" *Journal of Forecasting*. 10(6): 549-564.

Tannenbaum, D, C. R. Fox, and G. Ülkümen. 2017. "Judgment Extremity and Accuracy Under Epistemic vs. Aleatory Uncertainty." *Management Science*. 63(2):497-518.

Thrift, S.M. and D. von Winterfeldt (2021) "Risk-Informed Benefit-Cost analysis for Homeland Security R&D: Methodology and an Application to Evaluating the Advanced Personal Protection System for Wildland Firefighters," *Journal of Benefit Cost Analysis* 12(2):335-366.

Tversky, A. and D. Kahneman. 1974. "Judgments Under Uncertainty: Heuristics and Biases." *Science*. 185:1124-1131.

U.S. Environmental Protection Agency. 2021. "Uncertainty and Variability", EPA Expobox, Available at https://www.epa.gov/expobox/uncertainty-and-variability

U.S. Environmental Protection Agency. 2015. *Regulatory Impact Analysis of EPA's Proposed Regulations for the Management of Hazardous Waste Pharmaceuticals at*

Healthcare Facilities. June 2015. Available at https://www.regulations.gov/document/EPA-HQ-RCRA-2007-0932-0412

U.S. Environmental Protection Agency. 2011. *The Benefits and Costs of the Clean Air Act from 1990 to 2020: Final Report – Rev. A.* Prepared by the Office of Air and Radiation. April 2011. Available at https://www.epa.gov/sites/default/files/2015-07/documents/fullreport_rev_a.pdf

U.S. Environmental Protection Agency. 2009. *Expert Elicitation Task Force White Paper: External Review Draft.* Prepared for the Science and Technology Policy Council. Available at http://yosemite.epa.gov/sab/sabproduct.nsf/fedrgstr_activites/Expert%20Elicitation%20White%20Paper?OpenDocument

U.S. Environmental Protection Agency. 2000. *Arsenic in Drinking Water Rule Economic analysis.* EPA 815-R-00-026. Available at https://www.epa.gov/dwreginfo/support-documents-final-arsenic-rule

U.S. Food and Drug Administration. 2020. "How Could Lowering Nicotine Levels in Cigarettes Change the Future of Public Health?" Webpage content current as of 5/01/2020. Available at https://www.fda.gov/tobacco-products/public-health-education/how-could-lowering-nicotine-levels-cigarettes-change-future-public-health

U.S. Food and Drug Administration. 2019a. *Mammography Quality Standards Act; Amendments to Part 900 Regulations: Preliminary Regulatory Impact Analysis, Initial Regulatory Flexibility Analysis, Unfunded Mandates Reform Act.* Docket No. FDA-2013-N-0134. Available at https://www.fda.gov/media/122866/download

U.S. Food and Drug Administration. 2019b. *Premarket Tobacco Product Applications and Recordkeeping Requirements.* FDA-2019-N-2854. Available at https://www.fda.gov/about-fda/economic-impact-analyses-fda-regulations/premarket-tobacco-product-applications-and-recordkeeping-requirements-proposed-rule-preliminary

US Food and Drug Administration. 2017. *Food Labeling: Health Claims; Soy Protein and Coronary Health Disease.* FDA-2017-N-0763. Available at https://www.fda.gov/about-fda/economic-impact-analyses-fda-regulations/food-labeling-health-claims-soy-protein-and-coronary-heart-disease-proposed-rule

US Food and Drug Administration. Undated. *Food Labeling: Gluten-Free Labeling of Fermented or Hydrolyzed Foods.* FDA-2014-N-1021. Available at https://www.fda.gov/about-fda/economic-impact-analyses-fda-regulations/food-labeling-gluten-free-labeling-fermented-or-hydrolyzed-foods-regulatory-impact-analysis-finale

U.S. Health and Human Services. 2021. "Appendix D: Updating Value Per Statistical Life (VSL) Estimates for Inflation and Changes in Real Income." Prepared by L.A. Robinson, J.R. Baxter, and W. Raich for the Office of the Assistant Secretary for

Planning and Evaluation. April 2021. Available at https://aspe.hhs.gov/reports/updating-vsl-estimates

U.S. Health and Human Services. 2016. *Guidelines for Regulatory Impact Analysis.* Prepared by L.A. Robinson, J.K. Hammitt, J.R. Baxter, L. Ludwig, M. Black, and M. Welsh for the Office of the Assistant Secretary for Planning and Evaluation. Available at https://aspe.hhs.gov/reports/guidelines-regulatory-impact-analysis

U.S. Office of Management and Budget. 2003. *Circular A-4*. Available at https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A4/a-4.pdf

U.S. Pipeline and Hazardous Materials Safety Administration. 2015. *Final Regulatory Impact Analysis: Hazardous Materials: Enhanced Tank Car Standards and Operational Controls for High-Hazard Flammable Trains: Final Rule*. Available at https://www.regulations.gov/document/PHMSA-2017-0102-0016

U.S. Nuclear Regulatory Commission. 1975. *Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*. WASH-1400 (NUREG-75/014). Washington, District of Columbia. Available at https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr75-014/index.html

U.S. Nuclear Regulatory Commission. 1983. *PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants*. NUREG/CR-2300. Available at https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr2300/index.html

U.S. Nuclear Regulatory Commission. 1997. *Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts: Main Report.* Prepared by R.J. Budnitz, G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. NUREG/CR-6372.

U.S. Nuclear Regulatory Commission. 2012. *Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies.* Prepared by Annie M. Kammerer and Jon P. Ake. NUREG-2117, Rev. 1.

U.S. Nuclear Regulatory Commission. 2016. "White Paper: Practical Insights and Lessons Learned on Implementing Expert Elicitation." Prepared by Jing Xing and Stephanie Morrow, Office of Nuclear Regulatory Research. October 13.

Wallsten, T.S. 1986. *Meanings of Nonnumerical Probability Phrases.* Final Report for the Period 16 August 1983 Through 15 August 1986. L.L. Thurstone Psychometric Laboratory Research Memorandum No. 67. Chapel Hill, NC.

APPENDIX A: COMMONLY USED PROBABILITY DISTRIBUTIONS

This section describes the main characteristics of several probability distributions commonly used in Monte Carlo simulations.

A.1 UNIFORM DISTRIBUTION

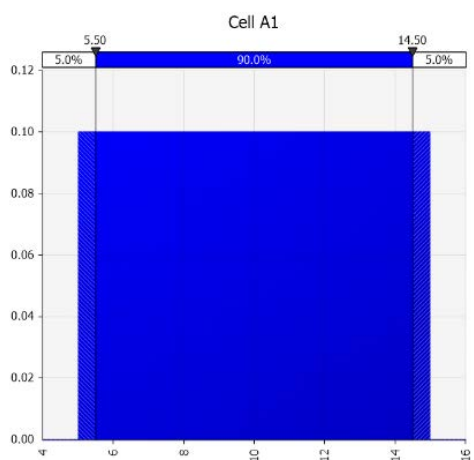
The uniform distribution (Exhibit A-1) is used when the uncertainty of interest is a bounded, continuous variable where all outcomes have the same probability. It is one of the simplest means of representing uncertainty in a model input because it is defined by only two parameters, the minimum and maximum value.

It is appropriate to use the uniform distribution when it is feasible to identify a range of possible values but when it is not feasible to identify that some values within the range are more likely to occur than others.

Parameters are usually determined using subjective reasoning, asking experts to provide the minimum and maximum values possible for the uncertainty of interest.

For example in the 2017 Soy Protein Relabeling analysis (US FDA 2017), a uniform distribution with minimum value of 200 and maximum value of 300 is used to model the likely number of products affected by the rule.

EXHIBIT A-1. EXAMPLE UNIFORM DISTRIBUTION



A.2 TRIANGULAR DISTRIBUTION

The triangular distribution (Exhibit A-2) is used for quantifying uncertainty with a rough three point estimation. It requires three parameters: a minimum, a maximum, and a most likely (mode) value. It is convenient and easy to understand and is appropriate when values near the middle of the range of possible values are considered more likely to occur than values near either extreme.

The triangular distribution does not need to be symmetric around the mean; it may be right or left skewed. The arbitrary triangular shape can help convey the message that the details of the shape are not precisely known which may help prevent an overinterpretation of results or a false sense of confidence.

For example in the 2017 Soy Protein Relabeling analysis (US FDA 2017), triangular distributions are used to estimate the costs to relabel different types of products. Exhibit A-3 shows the parameters values used for each triangular distribution in the Soy Protein Relabeling RIA for the six different products studied.

EXHIBIT A-2. EXAMPLE TRIANGULAR DISTRIBUTION

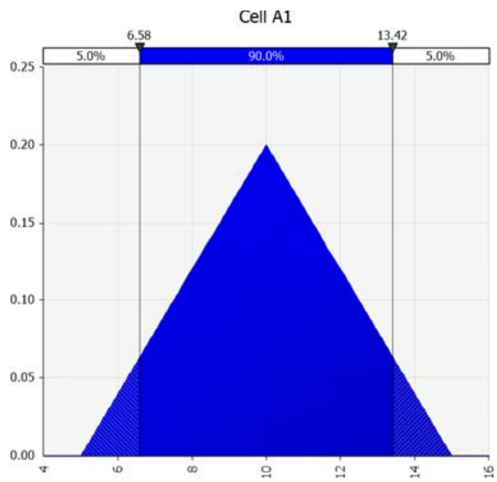


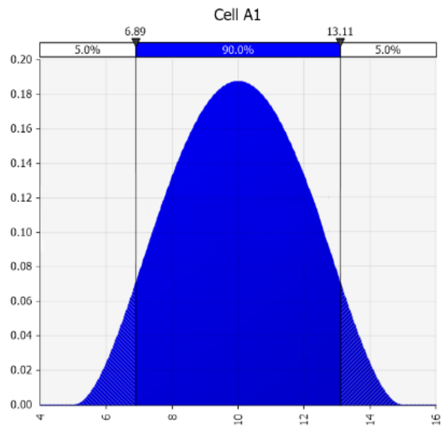
EXHIBIT A-3. PARAMETER VALUES USED FOR EACH TRIANGULAR DISTRIBUTION IN FDA (2017)

Type	Low	Mean	High
Soy Milk	$ 665	$ 1,233	$ 2,047
Protein Supplements	$ 1,484	$ 2,798	$ 4,722
Tofu	$ 810	$ 1,498	$ 2,507
Frozen Soybeans	$ 1,484	$ 2,798	$ 4,722
Canned Soup	$ 1,268	$ 2,324	$ 3,933
Snacks/Other	$ 189	$ 363	$ 544

A.3 PERT

The PERT distribution (Exhibit A-4) uses the same three parameters as the triangular: minimum, most likely value, and maximum. Like the triangular distribution, it places more emphasis on the most likely value over values around the minimum and maximum estimates. Unlike the triangular distribution, the PERT distribution constructs a smooth curve. Depending on the values provided, the PERT distribution can provide a close fit to the normal or lognormal distributions.

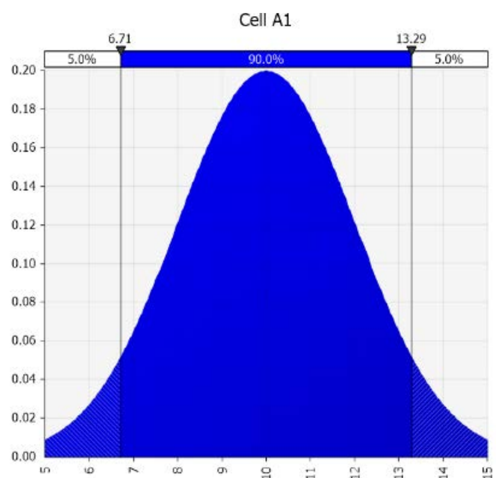
EXHIBIT A-4. EXAMPLE PERT DISTRIBUTION



A.4 NORMAL DISTRIBUTION

Because of the central limit theorem, which results in a normal distribution for additive quantities, the normal distribution (Exhibit A-5) is commonly an appropriate distribution to use in modeling. The distribution is characterized by two parameters: the mean and the standard deviation. For many uncertainties, the normal distribution is theoretically inappropriate because negative values are allowed. This problem can be ignored in many applications if the coefficient of variation, which is the ratio of the standard deviation to the mean, is less than about 0.2 because negative numbers would be very unlikely to be drawn from the distribution by the simulation.

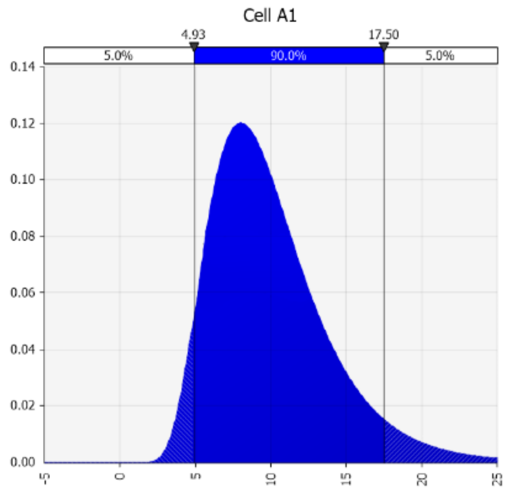
EXHIBIT A-5. EXAMPLE NORMAL DISTRIBUTION



A.5 LOGNORMAL DISTRIBUTION

The lognormal distribution (Exhibit A-6) results when the logarithm of the random variable is described by a normal distribution. The lognormal distribution is often found to be a good representation of physical quantities that are non-negative and positively skewed such as pollutant concentrations, stream flows, or explosion intensity. The parameters mu and sigma of the lognormal distribution correspond to the mean and standard deviation of the variable of interest's *logarithm*, something experts may have trouble intuiting.

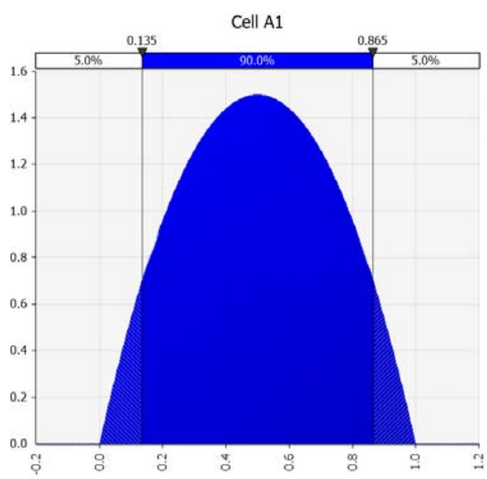
EXHIBIT A-6. EXAMPLE LOGNORMAL DISTRIBUTION



A.6 BETA DISTRIBUTION

The beta distribution (Exhibit A-7) provides a flexible means of representing variability over a fixed range, generally between 0 and 1, but additional parameters can be assessed to change the range of endpoints. The beta distribution is commonly used to represent variable probabilities or proportions.

EXHIBIT A-7. EXAMPLE BETA DISTRIBUTION



APPENDIX B: DISTRIBUTION SELECTION IN MONTE CARLO ANALYSIS

In this section, we illustrate how the selection of probability distributions can influence the outcome of an RIA. We draw upon the information included in Appendix A (Commonly Used Probability Distributions) to illustrate the sensitivity of example net benefits calculations to the characterization of uncertainty in key values. This example is not meant to be prescriptive – analysts may have reason to select alternative distributions from those presented in this appendix.

In the remainder of this section, we estimate the benefits and costs of a hypothetical rulemaking intended to reduce the risk of premature death in the United States. To characterize uncertainty in key values, we conduct a Monte Carlo simulation using the @Risk extension to Microsoft Excel.[13] These results are compared against a deterministic analysis in which best, or central, estimates are used without consideration of underlying uncertainty.

To quantify and monetize the benefits of this rule, we estimate the number of premature deaths avoided and value these deaths using estimates of the value per statistical life (VSL). In this example, both variables are uncertain. Analysts have estimated that the number of premature deaths could range from 20 to 80, with a best estimate of 40 deaths annually. Similarly, the *Guidelines* provide a low, central, and high estimate for VSL. These values are summarized in Exhibit B-1 and are assumed to reflect values for one year (in this case, 2021). Monetary values are expressed in 2014 dollars to facilitate comparison with HHS' *Guidelines* (see Table 3.1).[14]

EXHIBIT B-1. KEY PARAMETERS FOR BENEFITS ESTIMATION

PARAMETER	LOW	CENTRAL / BEST	HIGH
Premature deaths avoided	20	40	80
Value per statistical life (2014$, 2021 income levels)	$4.8 million	$10.3 million	$15.6 million

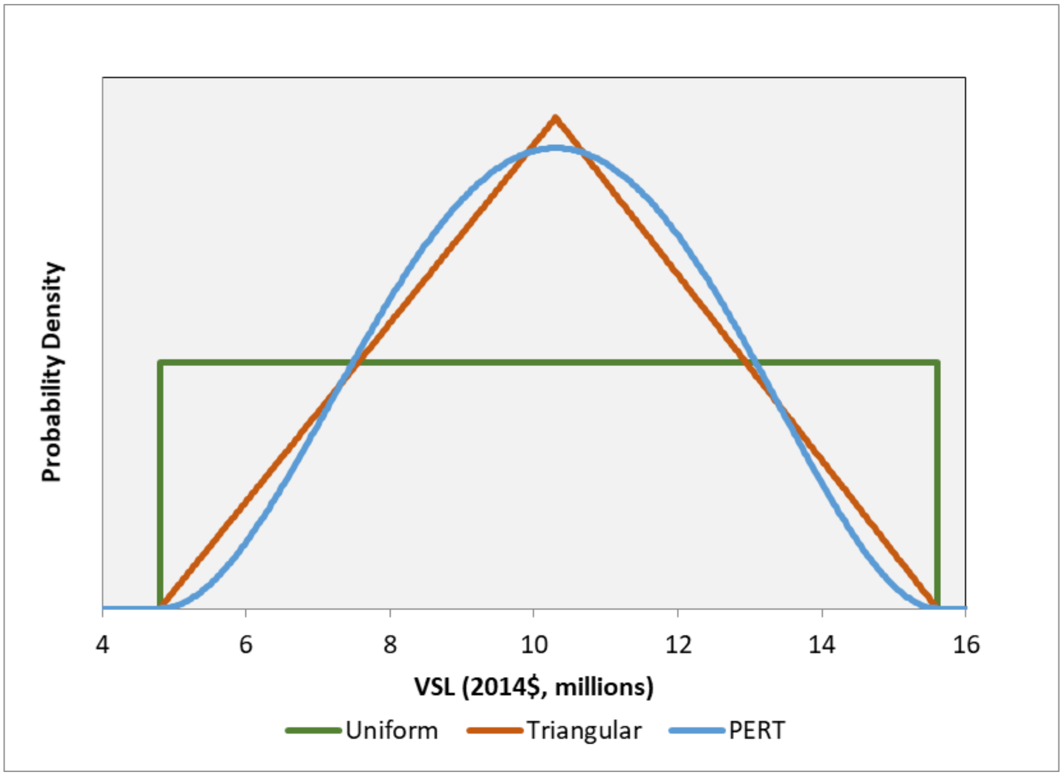
Next, we fit probability distributions to the low, central, and high estimates presented in Exhibit B-1. For each variable (deaths avoided and VSL), we fit three distributions: uniform, triangular, and PERT. These three distributions assign zero probability density outside of the range defined by the low and high estimates. That is, these values are treated as "bounds" inside of which we are certain the true parameter lies. In practice, analysts may wish to select alternative distributions. For example, we note that the tails of

[13] @Risk is a software add-in to Microsoft Excel that allows users to conduct Monte Carlo simulations. It is produced and sold by the Palisade Company (www.palisade.com).

[14] For a discussion of adjustments made to the VSL to account for inflation and changes in real income, see Chapter 3 of the *Guidelines*, or an expanded discussion contained in HHS (2021).

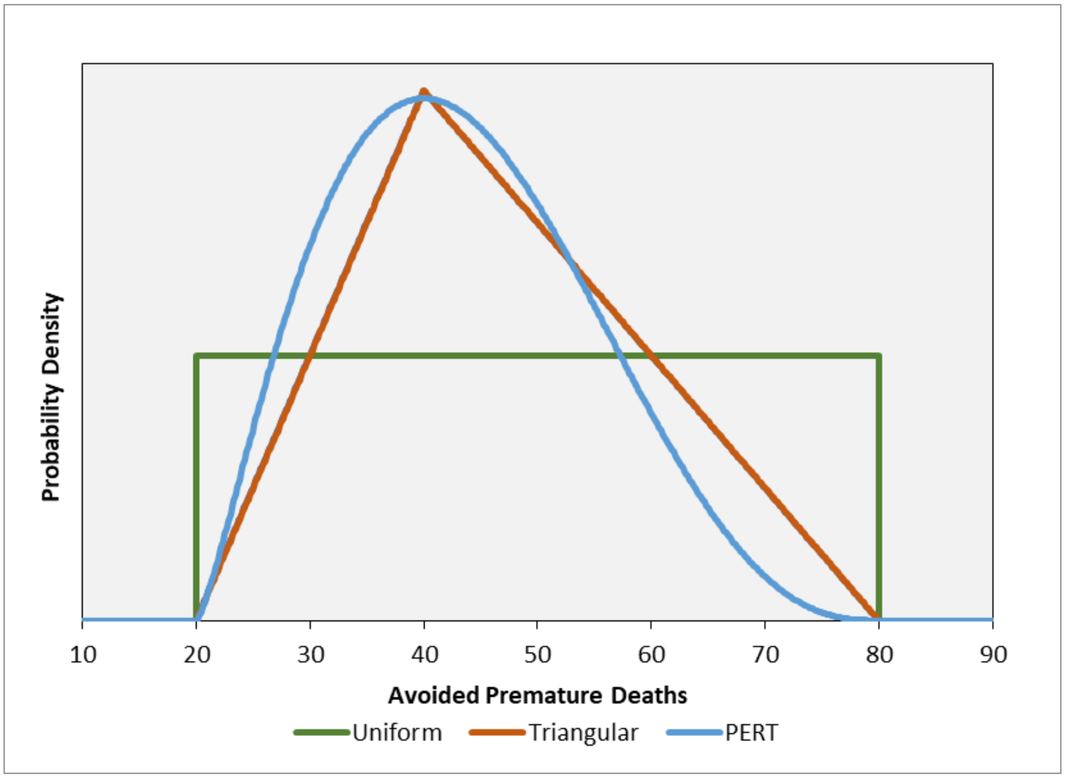
the VSL distribution likely extend beyond these bounds.[15] The resulting probability distributions are displayed in Exhibits B-2 and B-3.

EXHIBIT B-2. VSL SAMPLE PROBABILITY DISTRIBUTION



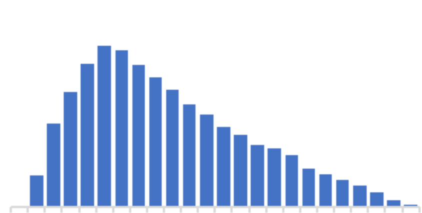
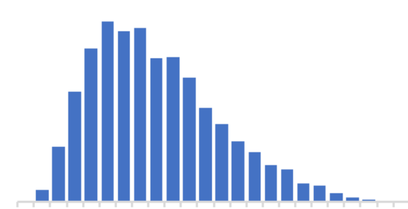
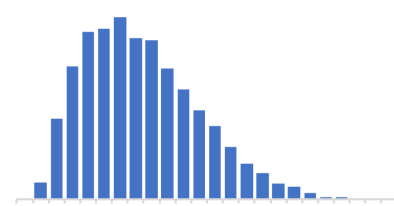
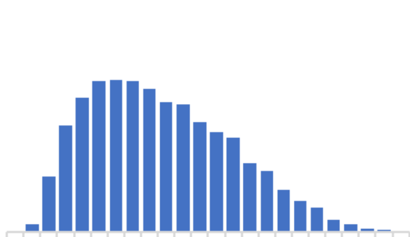
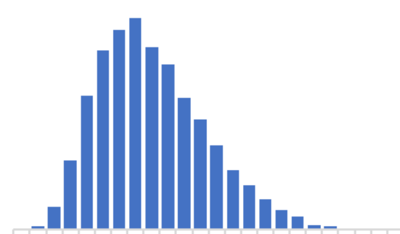
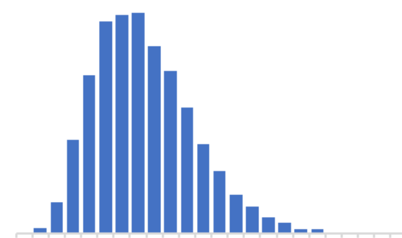
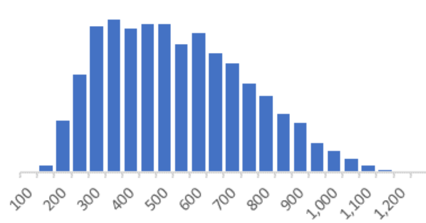
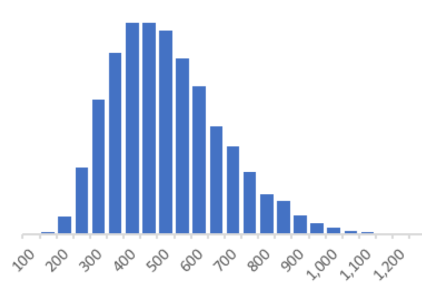
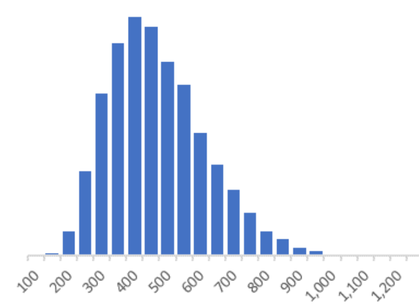
[15] The low and high VSL estimates used by HHS are each derived from individual studies estimating willingness to pay for mortality risk reductions. These studies acknowledge and quantify uncertainty in their best estimates. Additionally, the HHS VSL estimates are based on few studies selected through a criteria-driven literature review for their high quality and applicability to HHS regulatory analysis (see Robinson and Hammitt 2016). Other peer-reviewed research, while likely less applicable to the HHS context or of lower quality, provide alternative VSL estimates that fall outside of the range presented in HHS' *Guidelines*.

EXHIBIT B-3. AVOIDED PREMATURE DEATHS SAMPLE PROBABILITY DISTRIBUTION



Following parameterization of these distributions, we conduct a Monte Carlo simulation to estimate the benefits of this rule. We use @Risk to simulate 10,000 draws from each of the distribution. Total benefits are estimated by multiplying one VSL draw by one avoided deaths draw. Importantly, these distributions are assumed to be independent. That is, the value selected for VSL does not influence the value selected for avoided deaths, and *vice versa*. Combining three VSL distributions and three premature death distributions yields nine total benefits estimates, depicted in Exhibit B-4.

EXHIBIT B-4. ESTIMATED BENEFITS BASED ON DIFFERENT COMBINATIONS OF DISTRIBUTIONS

	UNIFORM VSL	TRIANGULAR VSL	PERT VSL
UNIFORM AVOIDED DEATHS			
TRIANGULAR AVOIDED DEATHS			
PERT AVOIDED DEATHS			

Notes: These histograms depict the relative frequency of simulated total benefits ($2014, millions). 10,000 simulations using @RISK.

As illustrated in Exhibit B-5, the selection of probability distribution influences the shape of the resulting distribution of simulated benefits. For example, using uniform distributions for both VSL and avoided deaths (upper left panel) results in a right-skewed distribution with fat tails (i.e., significant probability density at the extreme values). In contrast, combining PERT distributions (bottom right) results in a less skewed distribution with thinner tails. To better understand how these distributions compare across key summary statistics, we present the mean and 95 percent confidence intervals for each distribution in Exhibit B-5.

EXHIBIT B-5. SUMMARY STATISTICS FOR SIMULATED TOTAL BENEFITS (2014$, MILLIONS)

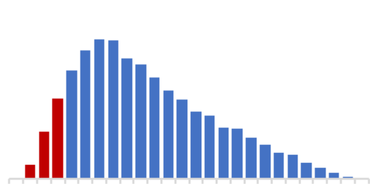
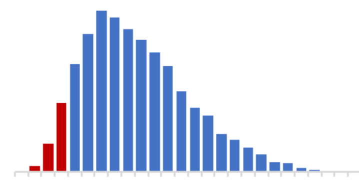
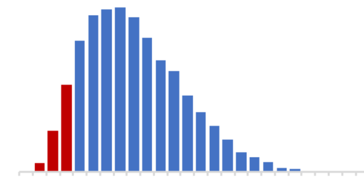
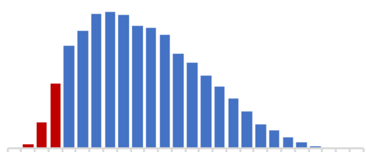
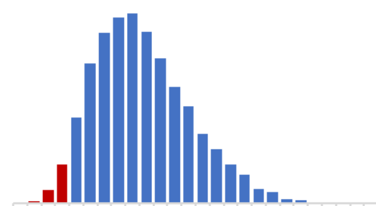
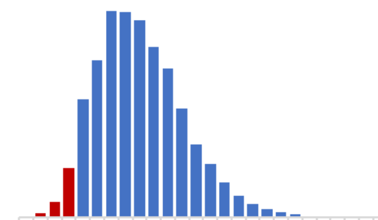
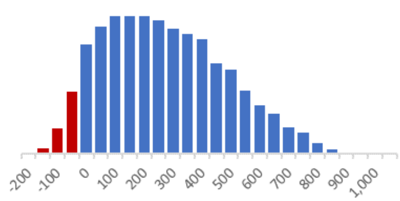
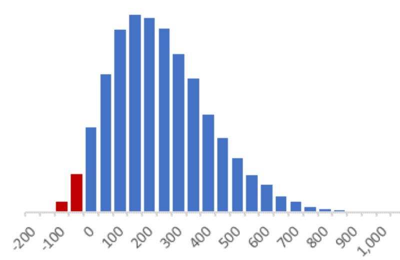
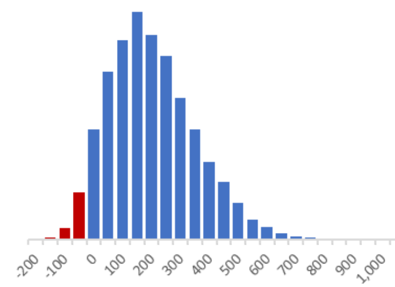
DISTRIBUTION FOR AVOIDED PREMATURE DEATHS	DISTRIBUTION FOR VSL		
	UNIFORM	TRIANGULAR	PERT
UNIFORM	510 (158 – 1,048)	476 (183 -933)	442 (173 – 844)
TRIANGULAR	512 (183 – 957)	477 (217 – 858)	444 (205 – 783)
PERT	514 (190 -951)	479 (223 – 842)	445 (214 – 772)

Simulated benefits range from a mean of $442 million to $514 million depending upon the choice of probability distributions. In this example, selecting a PERT distribution for VSL yields a lower mean benefit estimate relative to triangular and uniform distributions. Mean estimates are higher for triangular VSL distributions, and uniform VSL distributions result in the highest mean benefit estimates ($510 to $514 million). Uncertainty bounds similarly differ by distribution. Uniform distributions, which place more weight near the lower and upper bounds than most conventional distributions, produce the widest confidence intervals.

To illustrate the potential impact of these patterns on *net* benefits, we simulate a cost distribution ranging from $200 to $400 million (PERT distribution, $250 million best estimate). While this range falls below the mean benefits of the rulemaking, the tails of the distribution overlap with portions of the benefits distribution, resulting in net costs for some of the 10,000 draws. We assume independence between the cost and benefit distributions; however, analysts may need to consider whether this assumption is appropriate in other contexts.[16] Simulated net benefits are summarized in Exhibit B-6, which provides histograms to show the relative frequency of net benefits draws in bins representing increments of $50 million.

[16] For example, costs and benefits may be correlated if each component is influenced by the effectiveness of a regulatory measure (e.g., the number of regulated firms adopting a safer technology influences exposure rates).

EXHIBIT B-6. ESTIMATED NET BENEFITS BASED ON DIFFERENT COMBINATIONS OF DISTRIBUTIONS

	UNIFORM VSL	TRIANGULAR VSL	PERT VSL
UNIFORM AVOIDED DEATHS			
TRIANGULAR AVOIDED DEATHS			
PERT AVOIDED DEATHS			

Notes: These histograms depict the relative frequency of simulated net benefits ($2014, millions). Values in red depict net costs. Values in blue depict net benefits.

As demonstrated in Exhibit B-6, the selection of benefits distributions can influence the resulting probability of a rulemaking producing net benefits. In this example, costs exceed benefits in 8 to 18 percent of draws (depicted in red) depending upon the selected distributions. Simulated mean net benefits are presented in Exhibit B-7, along with the percent of draws producing net benefits.

EXHIBIT B-7. SIMULATED MEAN NET BENEFITS (2014$, MILLIONS) AND PERCENT OF DRAWS PRODUCING NET BENEFITS

DISTRIBUTION FOR AVOIDED PREMATURE DEATHS	DISTRIBUTION FOR VSL		
	UNIFORM	TRIANGULAR	PERT
UNIFORM	243 (83.9%)	209 (85.9%)	175 (82.3%)
TRIANGULAR	244 (87.2%)	211 (91.1%)	177 (88.3%)
PERT	247 (87.6%)	213 (91.9%)	178 (89.4%)

We estimate net benefits across all nine distribution sets. These net benefits range from $178 to $247 million, with relative magnitudes mirroring those summarized for total benefits: uniform VSL distributions produce the greatest net benefits ($243 to 247 million), followed by triangular ($209 to $213 million) and PERT distributions ($175 to $178 million). Given the wider uncertainty bounds produced by the uniform distributions (see previous Exhibit B-5), these results also produce the highest number of simulated net cost draws (i.e., instances where costs exceed benefits). In contrast, the PERT and triangular distributions are more likely to produce positive net benefits due to their thinner tails.

Relative to a deterministic net benefits calculation, the distributions applied in this appendix communicate much of the uncertainty associated with the results. Applying the central estimates for VSL and deaths, an analyst would otherwise estimate benefits of $412 million, costs of $250 million, and net benefits of $162 million. While close to some of the results included in this appendix, the "best estimate" falls below the means of all simulated distributions and fails to depict the possibility that costs could exceed benefits.

In summary, distribution choice can factor heavily into the resulting estimates of net benefits. In addition to affecting the mean estimates of net benefits, distribution choice will affect the uncertainty surrounding the mean value. For example, uniform distributions place equal weight in the tails of the distribution relative to most continuous distributions (e.g., normal, log-normal, PERT, triangular), which assign greater weight to one portion of the distribution. In our example, uniform characterization of uncertainty results in higher net *mean* benefits but also suggests a higher likelihood of costs exceeding benefits.

While we estimate net (mean) benefits for the nine combinations of probability distributions in our example, the choice of distribution can be important enough to reverse the expected sign of the net benefits. This suggests that, in cases where benefits and costs are more similar, careful selection of distributions is particularly important. For additional discussion of the choice and selection of distributions, see Appendix A.

APPENDIX C: KEY CONSIDERATIONS WHEN CONDUCTING FORMAL, STRUCTURED EXPERT ELICITATION

Exhibit C-1 summarizes the key steps in the conduct of formal, structured expert elicitation (also described in Section 5.0 of this paper).[17] In this appendix, we describe key considerations for each step. In some cases, these considerations are also relevant to less formal elicitation processes.

EXHIBIT C-1. GENERALIZED STEPS FOR STRUCTURED EXPERT ELICITATION

STEP	DESCRIPTION
1	Develop background information and prepare for the elicitation
2	Identify and recruit experts
3	Conduct the elicitation
4	Report results
5	Document and verify the process

STEP 1: DEVELOP BACKGROUND INFORMATION AND PREPARE FOR THE ELICITATION

The process begins with a series of important choices, including: (1) how to define and structure the questions; (2) whether to conduct an individual or group elicitation; (3) how to staff the elicitation team; and (4) how to assemble, disseminate, and refine the background information. Choices made during the background and preparation phase require careful attention because they can have important consequences for the elicitation itself.

Among the most important choices is defining the uncertain value that will be the focus of the elicitation. Ideally, the elicitation should involve estimating a quantity. The value of interest must be well-specified, in the sense that it could be resolved, at least in principle, by some experiment or measurement.[18] In addition, it should be one for which there is a basis for forming and justifying judgments. When crafting the questions, it is also important to consider relationships between the uncertain value and other variables. The elicitation team may find it useful to develop a decision tree to illustrate the relationships between key factors that influence the value of interest, as introduced in

---

[17] Several approaches are described in the available literature. While the steps in each approach are not identical (e.g., certain activities may be undertaken earlier or later in the process), the broad outline is similar across approaches.

[18] Experts in elicitation refer to this criterion as the "clairvoyant test," which requires that an omniscient being with complete knowledge of the past, present, and future could determine whether a specified value is correct of incorrect (Morgan and Henrion 1990). The idea is that the value being elicited should be clear and not subject to interpretation. For example, if an expert is asked to predict the likelihood that the following day will be cloudy, he may define "cloudy" differently than the individuals conducting the elicitation. Instead, the question should be constructed precisely (e.g., the likelihood that Region A will experience complete cloud cover for more than 50 percent of daylight hours) so that there is no room for misinterpretation.

Section 4.3 and discussed in more detail in related texts (Clemen 1996, Morgan 2014, EPA 2009).

Additionally, a choice must be made about whether to construct the questions using an "aggregated" or "disaggregated" approach. In an aggregated approach, the uncertain value is obtained through a single question (e.g., how many deaths would be prevented if the new labeling rule is implemented), whereas in the disaggregated approach, the quantity is obtained through a series of questions (e.g., what is the likelihood that consumers will respond to labels; what change in consumption is likely if consumers respond)? Disaggregation (or "decomposition") of a question is often recommended as a way to "divide and conquer" a complex problem (Clemen 1996, Hora 2007). However, the disaggregation process can be time-consuming, and it is possible to over decompose an uncertain value and make the assessment task more difficult (Hora et al. 1993). Therefore, an elicitation team just find a balance between an overly aggregated value definition and an overly granular one.

STEP 2: IDENTIFY AND RECRUIT EXPERTS

Selecting experts is one of the most important steps in the elicitation process, as the outcome of the elicitation depends on their personality, experience, and technical background (O'Hagan et al. 2006). There are three main choices: (1) *who* should be selected as an expert; (2) *how* should the experts be selected; and (3) *how many* experts should be selected.

Experts should satisfy two principal criteria: they must have substantive expertise relevant to answering the question, and they must be able to think about quantifying their judgments of the uncertain value using subjective probability. They should also be free of financial or personal conflicts of interest and other characteristics that may make them appear to lack impartiality (O'Hagan et al. 2006, U.S. EPA 2009). The experts should also represent a balanced range of opinions, particularly when the stakes are high (Knol et al. 2010, U.S. EPA 2009).

There are a variety of approaches for nominating and selecting experts, most of which involve assessing the breadth and influence of a potential experts' body of published work and/or recommendations or nominations from respected organizations or persons. The number of experts needed for a given elicitation depends on a number of variables, including: the complexity of the task, the diversity of scientific points of view, the resources of the project, regulatory limitations (i.e., the Paperwork Reduction Act), and the objectives of the elicitation (i.e., whether results will be combined or not) (Hammitt 2011).[19]

[19] Under the Paperwork Reduction Act, OMB approval is needed to collect the same or similar information from more than nine individuals. Seeking such approval requires substantial time and effort.

STEP 3: CONDUCT THE ELICITATION

Conducting the elicitation is the most complicated part of the process, requiring skilled normative and subject matter experts (SMEs) to guide the questioning.[20] A substantial body of literature focuses on particular aspects of the elicitation activities, including methods for motivating and training experts; and tools for eliciting probability distributions.[21]

In structured expert elicitation, experts are asked to characterize, in quantitative terms, their uncertainty associated with a given value. This uncertainty is generally described in terms of probabilities. Morgan and Henrion (1990) note that most people find it easiest to express probabilities qualitatively, using words such as "credible," "likely," or "extremely improbable." However, they cite evidence from the literature that the interpretation of the quantitative value associated with these types of qualitative terms varies widely across individuals and depends on context. Thus, researchers agree that using qualitative descriptors in elicitations generally produces unreliable results (Morgan and Henrion 1990 and Wallsten 1986, as quoted in U.S. EPA 2009). By contrast, quantitative probabilities provide a clear sense of likelihoods and a consistent interpretation that facilitates both comparisons between experts and synthesis of responses across experts.

The individual facilitating the elicitation must establish a rapport with the experts to provide an incentive for developing careful, thoughtful responses (Clemen 1996, O'Hagan et al. 2006) and help them understand how their judgments will be used and why the information is important. Experts, particularly scientists, may be hesitant to express uncertain opinions that may or may not be "correct" (Clemen 1996, O'Hagan et al. 2006), therefore, it is also important to assure them that uncertainty is natural.

Judgments may be elicited in several forms, including discrete values (e.g., probabilities, percentages, odds) and continuous probability distributions. Morgan and Henrion (1990) note that, in practice, most techniques for eliciting continuous distributions rely on asking a series of discrete assessments of specific points in the distribution, which are used to estimate a continuous distribution. Several tools may be used to help experts visualize probabilities and make judgments, such as marking a point on a scale of zero to one, a probability wheel, or a reference lottery (Morgan and Henrion 1990, Goodwin and Wright 2004).

Despite experience working with numbers and even statistics, many experts, including scientists, lack extensive experience using subjective probability to quantify their judgments. Thus, the expert elicitation community is in agreement with regard to the importance of providing training to the experts prior to eliciting judgments (Clemen 1996, O'Hagan et al. 2006, EPA 2009, EFSA 2014, Cooke and Goossens 1999). Training

---

[20] These individuals are part of the elicitation team and are distinct from the experts providing judgments. Both can be internal HHS staff. The normative expert should have training in best practices for conducting an elicitation.

[21] Activities related to motivating and training experts can also occur during a pre-elicitation workshop or at the beginning of the elicitation session (Cooke and Goossens 1999, O'Hagan et al. 2006, EPA 2009).

might include asking the experts to practice characterizing uncertainty using questions where the answers are known to the facilitator but not the experts. Another common element of training focuses on educating experts about common heuristics and biases, such as representativeness, availability, anchoring and adjusting, and motivational bias, based on the pioneering work of Tversky and Kahneman (1974) (for a discussion of each concept, see Clemen 1996).

Finally, a substantial body of literature is devoted to the questions of whether expert judgments should be elicited individually or in a group setting. Key strengths of group elicitations include avoiding the need to decide whether and how to aggregate individual judgments, allowing experts to pool their knowledge, and providing opportunity for experts from different disciplines to interact. However, a key drawback is the potential for a few outspoken or more senior individuals to dominate the discussions and a tendency for overconfidence. Additional general discussion of the strengths and weaknesses of individual versus group methods can be found in Clemen and Winkler (1999) and O'Hagan et al. (2006).

STEP 4: REPORT THE RESULTS

The elicitation team should report the judgments provided by each expert, including rationale, dependence on ancillary assumptions, or any other important information. If judgments are elicited individually, then the team must choose whether and how to pool the judgments. Options generally include a Bayesian approach, linear opinion pooling, and Cooke's Classical Model. For additional discussion of the strengths and limitations of each, see Clemen and Winkler (1999), O'Hagan et al. (2006) and, Cooke (1991).

STEP 5: DOCUMENT AND VERIFY THE PROCESS

The elicitation team should thoroughly document the expert elicitation to ensure transparency and replicability. This information may be included as an appendix to the RIA or in a stand-alone report that is referenced in the RIA and made available to the public. In addition, depending on whether the elicitation will be used in a "high stakes" rule (see Section 3.2), the elicitation team may wish to consider a peer review of the process.

For reference, examples of recent studies using structured expert elicitation to characterize the uncertainty associated with key parameters used in benefit-cost analysis of Federal regulations include elicitations of mortality risks associated with particulate matter (Roman et al., 2008) and the value per statistical life (Roman et al. 2012).