



HANDLING RISK AND UNCERTAINTY IN COST-BENEFIT ANALYSIS: A PRACTITIONER'S GUIDE

January 2023

This publication is made possible by the support of the American People through the United States Agency for International Development (USAID) and was prepared by Integra Government Services International LLC for the Learning, Evaluation, and Analysis Project (LEAP III) Activity.

HANDLING RISK AND UNCERTAINTY IN COST-BENEFIT ANALYSIS: A PRACTITIONER'S GUIDE

Contract Title: LEAP III: Learning, Evaluation and Analysis Project III
Contract Number: GS-10F-083CA / 7200AA18M0004
Activity Number: LEAP III 2022 – 1009.1069
Submitted: January 18, 2023 (Revised October 20, 2023)
Contractor: Integra Government Services International LLC
1100 Vermont Avenue NW, Suite 750
Washington, DC 20005
Sub-contractor: Limestone Analytics LLC
USAID Office: Center for Economics and Market Development
COR: Erin Shutty, kqutub@usaid.gov
Prepared by: Bahman Kashi (Limestone Analytics) and Eric Hyman (USAID)
Photo Credit: Measuring Risk by Andrey Popov

DISCLAIMER

This report is made possible by the support of the American people through USAID. The authors' views expressed in this publication do not necessarily reflect the views of USAID or the United States Government.

TABLE OF CONTENTS

TABLE OF CONTENTS	3
ACRONYMS	5
I. INTRODUCTION	6
1.1 OVERVIEW	6
1.2 IMPORTANCE FOR USAID	6
1.2 CRITICAL ASSUMPTIONS	7
2. METHODS FOR THE ANALYSIS OF RISK AND UNCERTAINTY	8
2.1 EXPECTED VALUE ANALYSIS	9
2.2 SENSITIVITY ANALYSIS	10
2.3 SCENARIO ANALYSIS	13
2.4 MONTE CARLO SIMULATIONS	14
3. STEPS IN RISK AND UNCERTAINTY ANALYSIS	16
3.1 IDENTIFICATION OF CRITICAL ASSUMPTIONS	17
3.2 SENSITIVITY ANALYSIS	17
3.3 ASSESSMENT OF CO-MOVEMENTS	19
3.4 SELECTION OF THE PROBABILITY DISTRIBUTIONS	19
3.5 IDENTIFICATION OF THE RESEARCH QUESTIONS AND METHODS	22
3.6 ANALYSIS OF RISK AND UNCERTAINTY	24
MODEL PREPARATION	24
SCENARIO ANALYSIS	25
MONTE CARLO SIMULATIONS	25
3.7 COMMUNICATION (ASSUMPTIONS, METHODS, FINDINGS)	26
ANNEX A: MORE ON EXCEL DATA TABLES	29
TROUBLESHOOTING EXCEL DATA TABLES	31
ANNEX B: USE OF THE CRYSTAL BALL SOFTWARE	32
INSTALLATION	32
STEPS FOR USING CRYSTAL BALL SOFTWARE	32
COPY-AND-PASTING ASSUMPTIONS	35
ANNEX C: GLOSSARY	36

ANNEX D: PARAMETERIZATION EXAMPLE	37
INFLATION RATE	37
NATURAL GAS PRICE	37
ANNEX E: ESTIMATING CORRELATION	39
ANNEX F: CASE STUDY	40
STEP 1: IDENTIFICATION OF CRITICAL ASSUMPTION	40
STEP 2: SENSITIVITY ANALYSIS	41
STEP 3: ASSESSMENT OF CO-MOVEMENTS	43
STEP 4: SELECTION OF PROBABILITY DISTRIBUTIONS	43
STEP 5: IDENTIFICATION OF THE RESEARCH QUESTIONS AND METHODS	44
STEP 6: ANALYSIS OF RISK AND UNCERTAINTY	44
STEP 7: COMMUNICATION (ASSUMPTIONS, METHODS, FINDINGS)	46
REFERENCES	47

ACRONYMS

BCR	Benefit-cost ratio
CBA	Cost-benefit analysis
CEA	Cost-effectiveness analysis
IRR	Internal rate of return
NPV	Net present value (present value of net benefits)
USAID	United States Agency for International Development

I. INTRODUCTION

I.1 OVERVIEW

The purpose of this practitioner’s guide is to provide supplementary guidance for analyzing uncertainty and risk in cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA). Subsequent use of the term “cost-benefit analysis” or the acronym “CBA” in this report may pertain to either of these tools. This guide will be most useful for practitioners already familiar with CBA concepts and models in spreadsheet programs, such as Microsoft Excel or Google Sheets.

CBA models rely on data and assumptions subject to risk and uncertainty at the time of the analysis. Marchau et al. (2019) defined uncertainty as the gap between available knowledge and the knowledge decision-makers need to make the best policy choice. Risk is a type of uncertainty, the part that is calculable and quantifiable using probabilities and losses. The degree of uncertainty and risk, and the choice of risk analysis method depend on the intervention (sector, location, scale, and other design parameters) and the availability and reliability of the underlying data.

Major CBA textbooks, including Florio (2014), Jenkins et al. (2011), Boardman et al. (2018), and Sartori (2014), discussed the theory and methods for integrating risk analysis in CBA. This practitioner’s guide summarizes the main concepts and applications. It includes a seven-step process for conducting risk analysis along with details concerning the choice of software and data sources. The guide is accompanied by a spreadsheet and a slide deck for training.

I.2 IMPORTANCE FOR USAID

The analysis of risk and uncertainty in CBAs can identify the sources of uncertainty, qualitatively or quantitatively measure the level of confidence in the results, suggest measures for managing and mitigating risk, and report the risk and uncertainty associated with a decision. These insights are valuable for development activities and USAID’s Program Cycle.

USAID’s new Risk Appetite Statement stated that the Agency has a high-risk appetite in its programs and approaches for achieving development outcomes. The statement established a lower risk appetite level for six other risk categories -- fiduciary, reputational, legal, security, human capital, information technology, and operational (USAID 2022a).

Since USAID works in some of the most fragile countries and changing contexts, risk and uncertainty are unavoidable considerations in the Agency’s work. Having a high appetite for programmatic and development outcome risks does not imply that there is less need to analyze these risks and assess uncertainty. In fact, it increases the need to understand, reduce, and compensate for programmatic and development outcome risks. The analysis is also essential to avoid increasing fiduciary, reputational, legal, security, human capital, information technology, and operational risks beyond USAID’s lower risk appetite level.

ADS Chapter 20I also highlighted the need to account for risk and uncertainty in designing, implementing, and monitoring activities (USAID 2022b). However, the use of CBA is optional at USAID. USAID staff or partners can use CBA to inform the design, implementation, monitoring, or evaluation of activities and interventions. When CBAs are conducted, this practitioner’s guide can facilitate the quantitative and qualitative assessment of risk and uncertainty.

I.2 CRITICAL ASSUMPTIONS

A CBA model calculates decision criteria, such as the net present value (NPV), based on a series of assumptions. *Critical assumptions* can significantly affect the decision criteria; in other words, they are important drivers of the results. Figure 1 provides a general process flow for a CBA model.

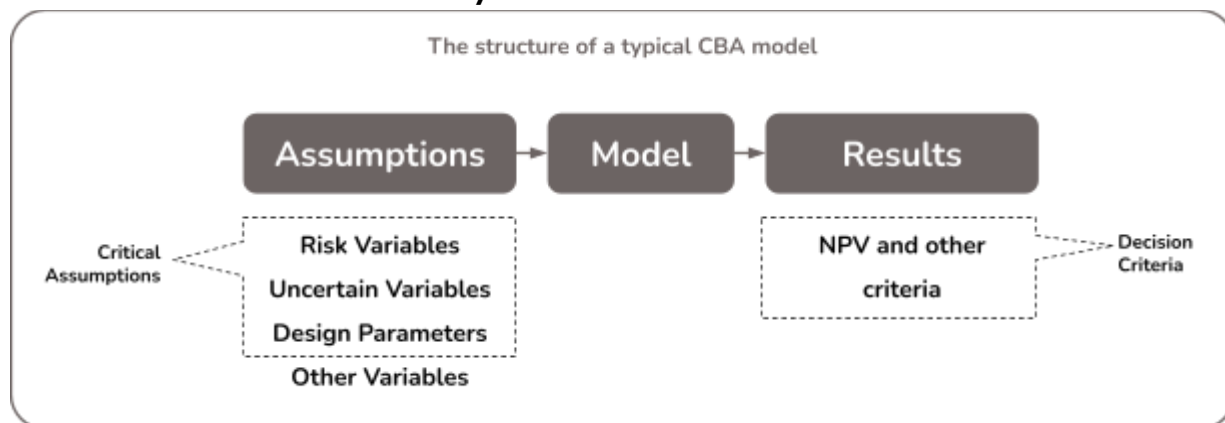
There are three main types of critical assumptions:

- **Risk variables:** Assumptions subject to known variability, including possible outcomes and the likelihood of occurrence. Examples include market prices of traded inputs and products.
- **Uncertain variables:** Assumptions that are subject to unknown variability. Examples include the timing and magnitude of an extreme weather event or another natural disaster in a particular location.
- **Design parameters:** Assumptions that reflect an analytical or design decision. Examples include the discount rate for the time value of money, the scale of the intervention, and the length and intensity of an activity.

Sometimes, these distinctions are less clear-cut. For example, the participation rate of small-scale farmers in a farmer training intervention could be analyzed as a risk or uncertain variable.

Although changes in the value of every assumption can affect the decision criteria, not all assumptions are risk or uncertain variables. To constitute a risk or uncertain variable, the range of variability would have to affect the decision criteria substantially. Otherwise, they do not need to be analyzed as risk or uncertain variables. For example, fertilizer market prices can be highly volatile, but if small-scale farmers do not buy much fertilizer or the government controls or subsidizes the prices, these prices would not affect the total costs in an agriculture CBA.

FIGURE 1. Elements of a Risk Analysis s in a CBA Model



2. METHODS FOR THE ANALYSIS OF RISK AND UNCERTAINTY

Section 2 discusses the four main tools for analyzing risk and uncertainty in CBAs:

- **Expected value analysis:** Uses probabilities in CBA calculations.
- **Sensitivity analysis:** Focuses on how changes in one or two assumptions affect the results.
- **Scenario analysis:** Reports decision criteria under more or less favorable scenarios.
- **Monte Carlo simulations:** Uses assumed probability distributions in many random scenarios to estimate the mean values and ranges for the decision criteria.

These four types of tools can be complementary and are often used together. Expected value analysis brings probabilities into the CBA formulas to report expected values for decision criteria (Boardman et al. 2018). Scenario analysis can either be a separate method (Jenkins et al. 2011) or part of a sensitivity analysis (Florio 2014).

In combination, these tools can generate five common outputs:

1. The most critical assumptions and their viable range
2. The probability distribution for risk variables
3. The range of possible outcomes for each investment criterion
4. The probability distribution for investment criteria
5. The *switching values* for critical assumptions

Switching values are values of a critical assumption (design parameter, risk, or uncertain variable) that results in an inflection point between favorable and unfavorable outcomes, such as a positive and negative NPV (NPV of zero). Switching values are also known as the break-even points.

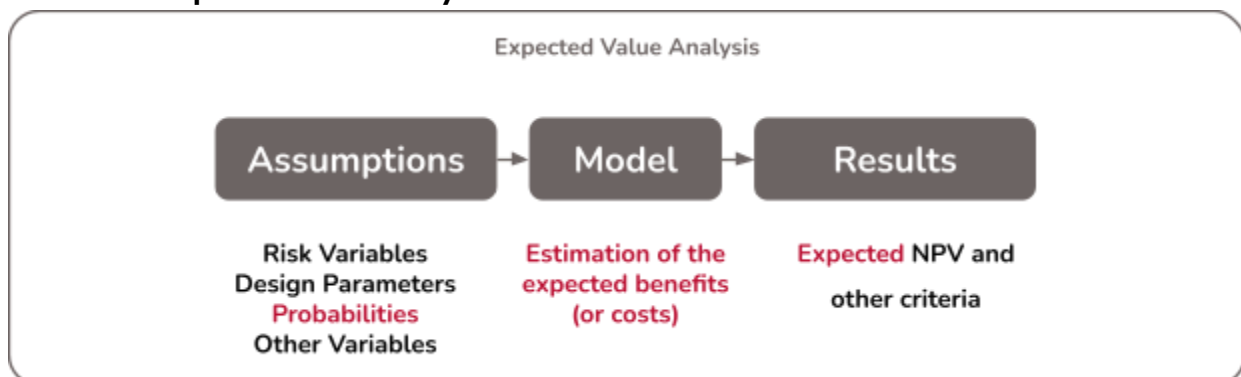
The four main tools for analyzing risk and uncertainty in CBAs differ in their capabilities. For instance, sensitivity analysis follows a deterministic structure and cannot produce probability distributions. Table I summarizes the advantages and disadvantages of these four methods.

TABLE 1: Four Methods for Risk Analysis		
Method	Advantages	Disadvantage
Expected Value Analysis	<ul style="list-style-type: none"> Used in CBA to assess the desirability of interventions that affect the probability of an event or the projected positive or negative impacts if the event occurs 	<ul style="list-style-type: none"> Does not estimate statistical confidence levels Not appropriate for risks that have a low-probability, but potentially catastrophic impacts
Sensitivity Analysis	<ul style="list-style-type: none"> Easy to implement¹ Can estimate switching values Easy to communicate 	<ul style="list-style-type: none"> Usually limited to changing one or two variables at a time Does not report estimate statistical confidence levels
Scenario Analysis	<ul style="list-style-type: none"> Easy to implement Considers co-movements² Easy to communicate 	<ul style="list-style-type: none"> Does not estimate statistical confidence levels
Monte Carlo Simulations	<ul style="list-style-type: none"> Considers co-movements Estimates confidence levels 	<ul style="list-style-type: none"> Complex to set up, run, and troubleshoot The appropriate type of probability distribution is generally not known and findings may vary substantially with different assumed probability distributions Easier to hide underlying biases and setup errors compared to other methods Requires specific software programs

2.1 EXPECTED VALUE ANALYSIS

Expected value analysis incorporates the probability and impact of an event in estimating costs or benefits. For example, better road designs and construction methods can reduce the probability and potential impact of traffic accidents. Similarly, introducing an ambulance service can reduce the mortality rate associated with traffic accidents. In either case, the probability of accidents is a part of the benefits estimations in these examples. Figure 2 summarizes expected value analysis, where probabilities enter as variables and calculations are updated to estimate expected values.

FIGURE 2. Expected Value Analysis



¹ The straightforwardness of implementation in Excel has made sensitivity analysis a popular debugging tool for practitioners.

² Correlates changes in the values of two or more parameters.

Table 2 shows the expected net benefits of vehicle accident reduction.

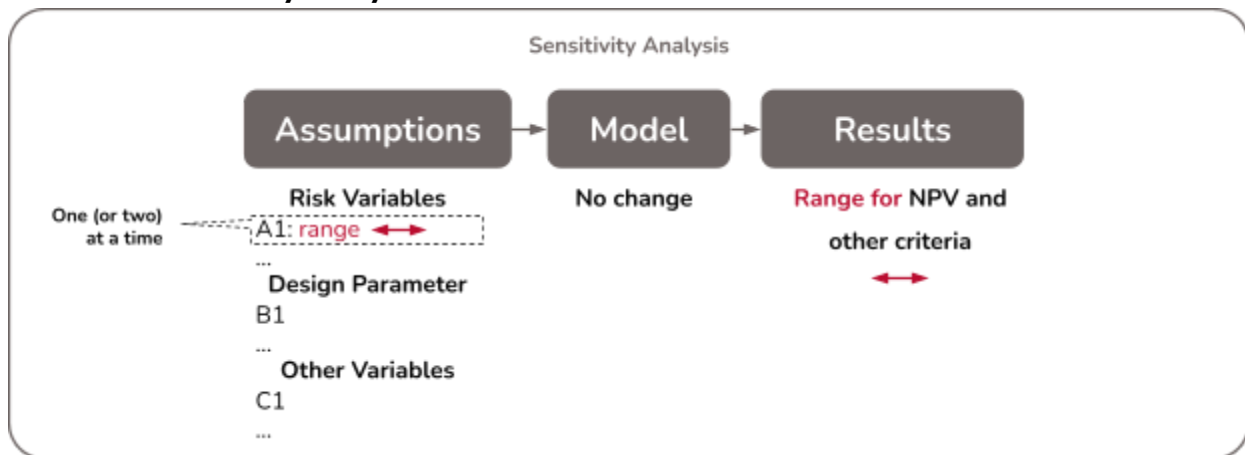
TABLE 2: The Expected Value of the Net Benefits of Accident Prevention						
Event	Cost	Probability without the Intervention	Annual Traffic (Vehicles)	Expected Value of Accident Damage Cost	Probability of Event with the Intervention	Expected Benefit of the Intervention
Traffic Accident	\$10,000	1 percent	200,000	$\$2,000,000 = (10,000 \times 0.01 \times 200,000)$	0.5 percent	\$1,000,000

Expected value analysis is widely used in CBAs for health interventions, disaster management, and efforts to reduce the cost of contingent events (such as sickness and natural disasters). Expected value analysis can also help decision-makers when a particular risk threatens a project's success. Some development assistance agencies encourage institutional reform as a high-risk, high-return investment. CBAs for institutional reform activities can estimate the expected benefits by multiplying the gains (conditional on success) by an assumed probability of success. The lowest probability of success that can justify the intervention may be the main criterion to report to decision-makers from this expected value analysis. See Belt et al. (2017) for an example of expected value analysis in a CBA.

2.2 SENSITIVITY ANALYSIS

Sensitivity analysis shows how changes in one or two variables at a time will affect the decision criteria. Multiple sensitivity analyses can be done sequentially to examine the effects of changing a larger number of variables. A sensitivity analysis is relatively easy to set up to test model formulas, identify the critical variables, and estimate switching values. However, sensitivity analysis typically only tests the sensitivity of decision criteria (or one decision criterion) to arbitrarily defined changes in one (or two) inputs at a time.

FIGURE 3. Sensitivity Analysis

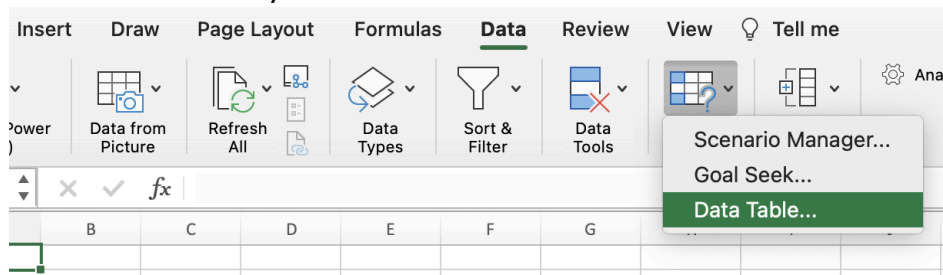


Sensitivity analysis can produce two types of information:

- **Range of values for decision criteria**
- **Switching value (break-even analysis)**

FIGURE 4. Sensitivity Analysis: Finding the Range of Values for Decision Criteria in Excel

Excel has built-in features for conducting sensitivity analysis. In the 2021 version of Excel, this feature is under the Data → What-if Analysis → Data Table menu.



There are three elements for data tables to function:

- **The input** must be an exogenous entry (E19 in this example).
- **The calculation** must be in a cell (E20).
- **The alternative values** to consider for the input.

Important note: For data tables to work, the input and the table must be on the same sheet.

Once you have a functioning spreadsheet where at least one calculation depends on one input, you can use the steps below to build a data table. These tables can be constructed vertically or horizontally:

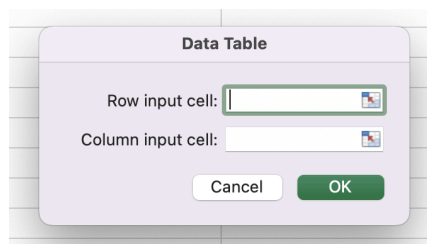
1. Place the alternative values in a column for a vertical table (or row for a horizontal table)
2. Create a *linked cell* to the calculation in the first cell of the second column (or second row) to complete the skeleton of the table. A linked cell is a simple formula that creates a reference to another cell. If the reference cell is C20, the linked cell will have the following content: =C20.
3. Select the table's area, including the alternative values and the linked cell.

	A	B
1		=E20
2	10	
3	20	
4	30	
5	40	
6	50	

or

	A	B	C	D	E	F
1		10	20	30	40	50
2	=E20					
3						
4						
5						
6						
7						

4. Click on the “Data Table” from the above menu
5. Excel will then open a new window asking for the input. For vertical tables, enter the reference for the input cell in “Column input cell.” Otherwise, enter the input cell reference in the “Row input cell.”
6. Press “OK,” and Excel will fill the empty cells with alternative values for the decision criterion in E20.

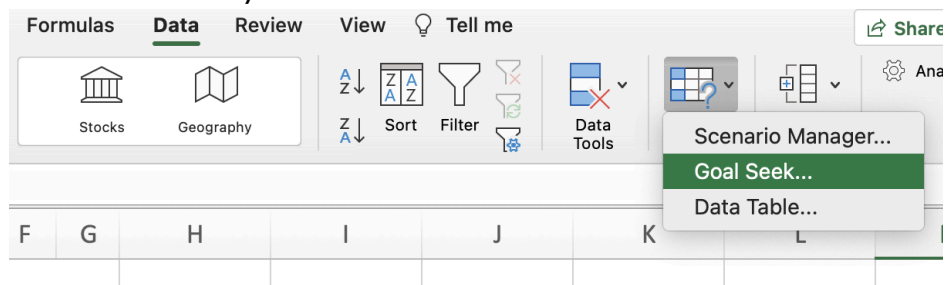


Annex A contains further practical notes on how to

- Create an Excel data table with multiple decision criteria
- Create an Excel data table with two inputs
- Format Excel data tables
- Troubleshoot data table issues

FIGURE 5. Sensitivity Analysis: Finding Switching Values in Excel

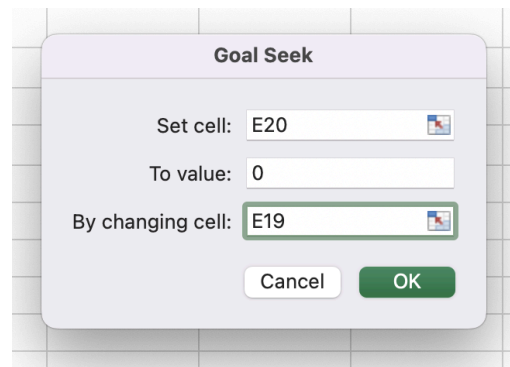
Excel has built-in features for finding switching values. In the 2021 version of Excel, this feature is under the Data → What-if Analysis → Goal Seek menu.



There are three elements for finding a switching value:

- **The input** must be an exogenous entry (E19)
- **The calculation** must be in a cell (E20)
- **The threshold** is the inflection point for a decision criterion (the calculation). For NPV, the threshold is zero.

Clicking on the Goal Seek menu opens a prompt requesting references or values for the three items. “Set cell” will look for the location of the calculation in your CBA. “To value” is the threshold for Goal Seek. “By changing cell” is the input. Clicking okay will force Excel to replace the input value (E19) with another value that results in an NPV (E20) to zero.



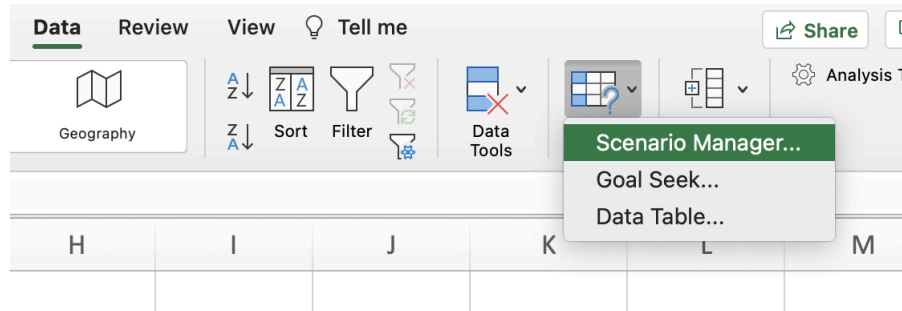
2.3 SCENARIO ANALYSIS

A *scenario* is a set of data and assumptions that defines the internal and external context of a probabilistic or deterministic scenario. An analyst can change some or all of the data and assumptions to estimate and compare the resulting decision criterion value under a different scenario. This tool is similar to sensitivity analysis, but scenario analysis allows changes in multiple inputs at the same time. Often, scenarios are designed to be more favorable (optimistic or best case) and less favorable (pessimistic or worst case).

Scenario analysis is a powerful analytical tool. It is easy to approve an investment if the decision criteria support it under a conservative, less favorable scenario. Similarly, it is easy to reject an investment when decision criteria dismiss it under an optimistic, more favorable scenario.

FIGURE 6. Scenario Analysis: Excel Scenario Manager Tool

Excel has a built-in feature for conducting scenario analysis. In the 2021 version of Excel, this feature is under the Data → What-if Analysis → Scenario Manager menu.

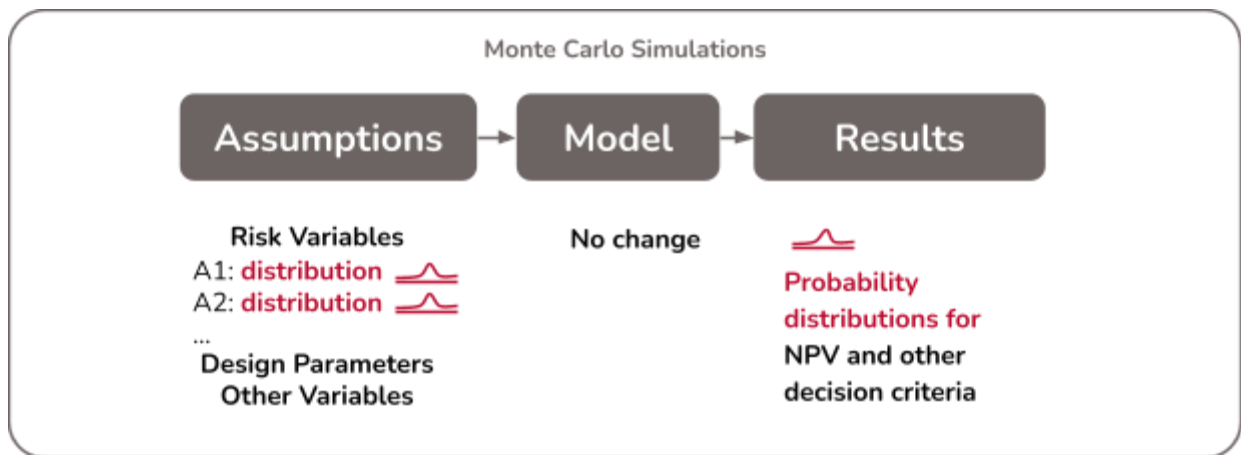


For each scenario, you need to select the range of inputs for which you would like to assume different values under each scenario. You can then name each scenario and assign the values for each input under each scenario. When this information is inserted in the Scenario Manager menus, you can use the summary button to have Excel prepare a simple report. After pressing the summary button, Excel shows a new window asking for the cell that contains the results (decision criteria). Placing the reference to results is in the new window and pressing ok will produce the results and present them as a new sheet in the same Excel workbook.

2.4 MONTE CARLO SIMULATIONS

Many parameters in a CBA model are subject to random or nonrandom variations. Monte Carlo simulations analyze the effects of varying assumptions with known or plausible probability distributions on the decision variables in many random samples. Each simulation run uses a specific value for each parameter based on the range and probability distribution. The software then estimates the outcomes for each simulation and the average and range of NPV values across the whole set of simulation runs. For example, Monte Carlo simulations can be used in an agricultural CBA to show how random weather variations affect the NPV. Compared to sensitivity analysis or scenario analysis, Monte Carlo simulations generate more information, such as the probability of obtaining favorable outcomes. Figure 7 shows the basic approach of a Monte Carlo analysis.

FIGURE 7. Monte Carlo Simulations



An easy way to understand this tool is to consider a fair coin toss with a 50 percent probability of heads or tails in a random distribution. A single coin toss will result in one outcome -- either heads or tails. However, if you flip a coin 10,000 times, the outcome will not necessarily be 5,000 heads and 5,000 tails. You might end up with 4,900 heads and 5,100 tails, purely by chance, and the probability of this or any other outcome happening can be estimated. With a known sample size, the average (mean) outcome and standard deviation of the population can be estimated.³ Because a random coin toss has a normal statistical distribution, the mean and standard deviation estimate the range of likely outcomes at a confidence level of 95 percent (or some other probability).

There are many examples where data gaps or inaccuracies have limited the use of Monte Carlo simulations or produced misleading results. The statistical basis and visual presentations of probability distributions can also give decision-makers an unrealistic impression that the analysis is objective and will produce accurate projections. Consequently, this tool should only be used when adequate data are available.

A Monte Carlo analysis consists of the following steps:

- I. Build a scenario by randomly selecting a set of values for each risk variable (based on a realistic probability distribution and assumed correlations).

³A standard deviation measures the amount of variation in a set of values. A low standard deviation indicates that the values tend to be close to the mean. When the standard deviation is high, the values are spread out over a wide range around the mean.

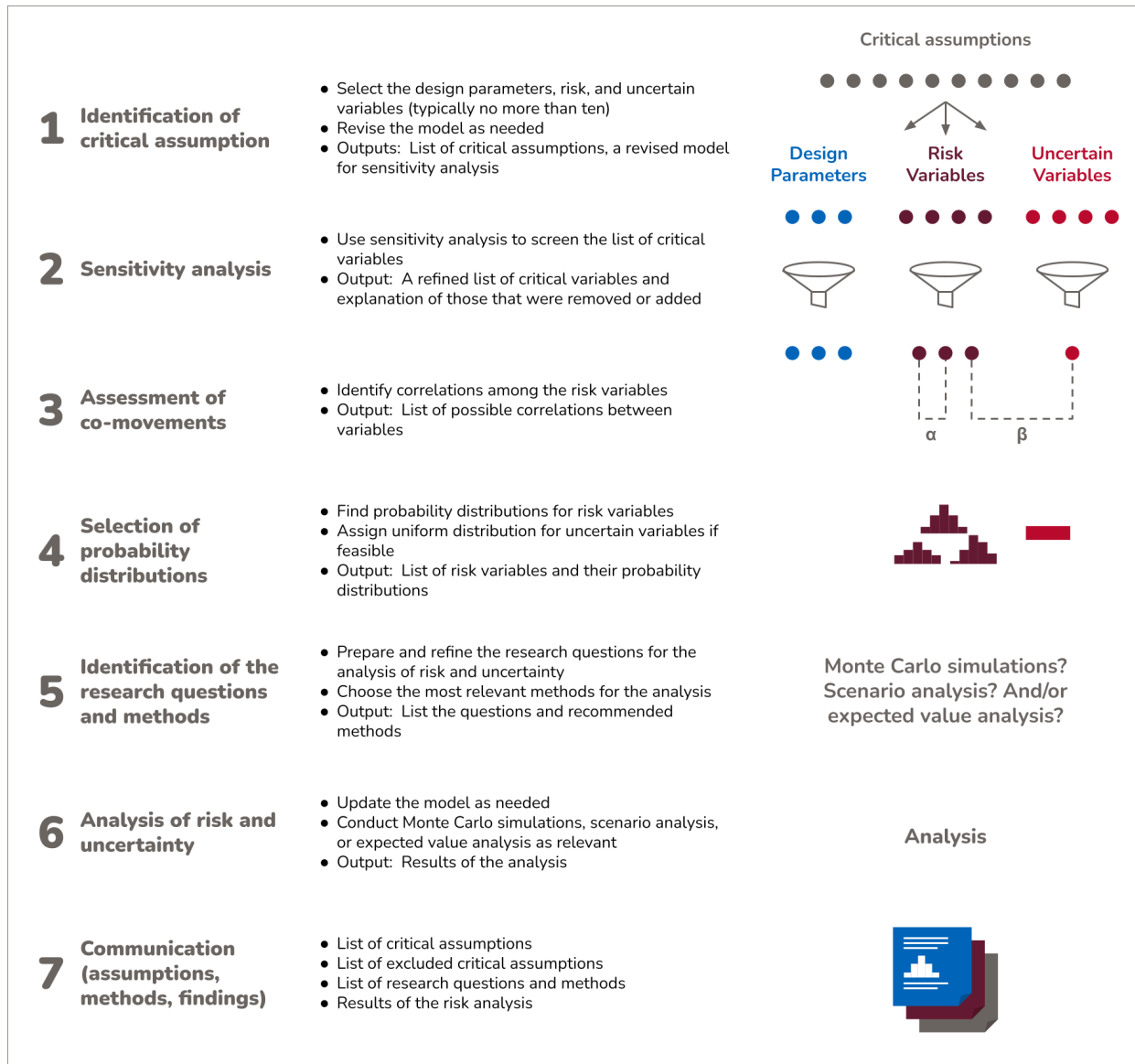
2. Plug the values into the model and record the values estimated for each decision criterion.
3. Repeat steps 1 and 2 until many values are recorded for the decision criteria.
4. Report the statistical properties of the sample for each decision criterion (e.g., expected value or the probability of being below or above a specific threshold value).

Various desktop computer software programs can be used for a Monte Carlo Analysis. One example is Oracle's Crystal Ball, an easy-to-use add-on for Microsoft Excel. Some USAID staff and contractors have used Oracle's Crystal Ball software and similar programs to quickly and easily run 10,000 Monte Carlo simulations on a laptop computer. At this time, Google Sheets does not have a comparable quality feature or reliable software add-on for Monte Carlo simulations. Annex B provides a step-by-step set of instructions for using Oracle Crystal Ball to conduct risk analysis on CBA models.

3. STEPS IN RISK AND UNCERTAINTY ANALYSIS

This section discusses seven practical steps to analyze risk and uncertainty in a cost-benefit analysis. Figure 8 provides a summary of the actions and outputs of each step.

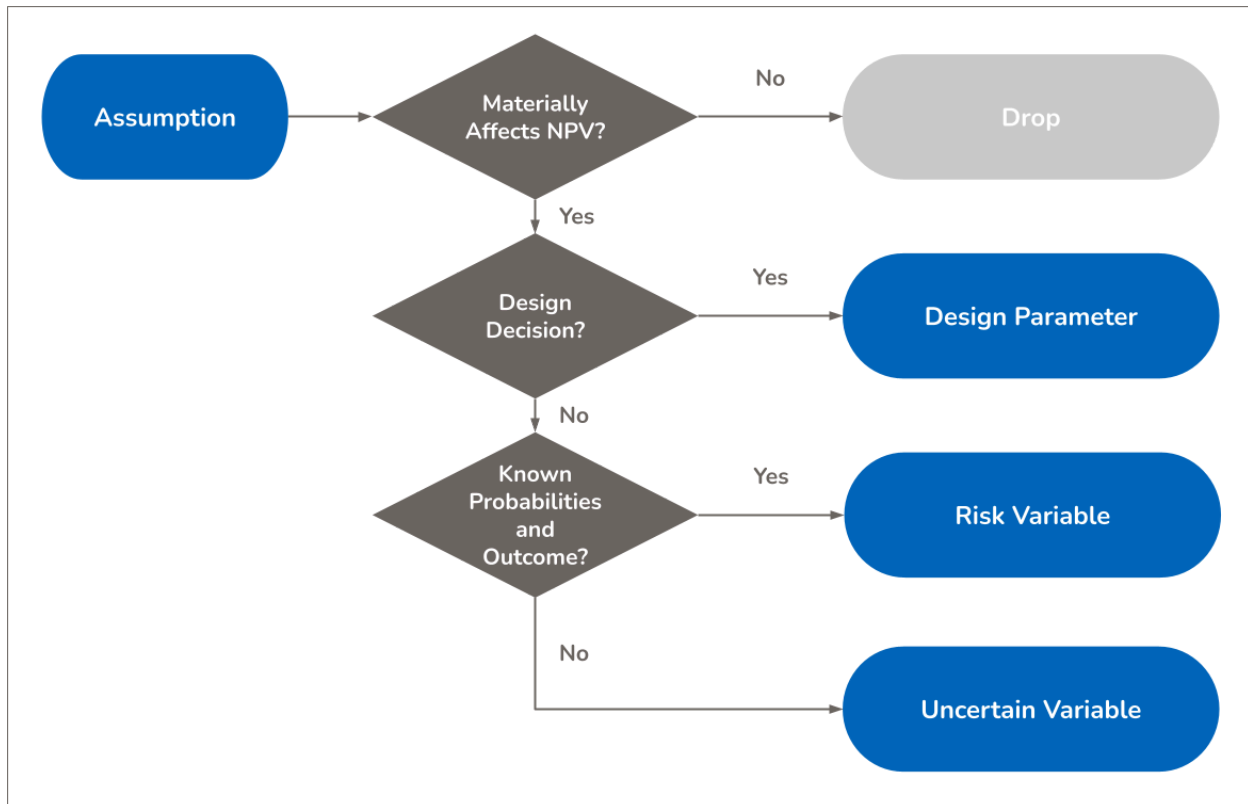
FIGURE 8. Seven Steps in Risk and Uncertainty Analysis



3.1 IDENTIFICATION OF CRITICAL ASSUMPTIONS

Risk and uncertainty should be considered early in the process of conducting a CBA. Begin by preparing a list of questions about the key assumptions that can threaten or increase the probability of success. Identify the most critical assumptions (design parameters, risk, and uncertain variables). Drop the assumptions that do not materially affect the NPV. Separate design decisions from the remaining assumptions. Separate the rest of the assumptions based on the existing knowledge about their probabilities and outcomes. When such knowledge exists, classify them as risk variables. Otherwise, they are uncertain variables. Figure 9 presents a flow chart for deciding the classification of critical assumptions.

FIGURE 9. Identification of Critical Assumptions



Although presented as the first step, identifying critical assumptions is an iterative process. The analyst or team may have to repeat this or other steps as they might identify new risks and design parameters later in subsequent steps through a qualitative or quantitative approach. For instance, the analysis can change the answer to the first question about the magnitude of the effect on NPV.

3.2 SENSITIVITY ANALYSIS

Prepare the CBA model for sensitivity analysis by ensuring that the model can accommodate variation in the value of the identified critical assumptions. For example, in an agricultural CBA, variation in crop yield should affect costs and revenues. If the model does not calculate harvesting and post-harvesting

costs as a function of the yield rate, then this model is not ready for the analysis of sensitivity to crop yield rate.

If some design parameters are handled as critical assumptions, it may be necessary to consider some changes in the approaches. For example, a CBA on whether to introduce a commercial service for potable water supply in a rural village may include the following assumptions:

- The price of water (revenue source for the private operator and a cost for the consumers).
- The amount of water consumed per individual or household (a major driver of the benefits).

If the price of water is included in the critical assumptions as a design parameter, the model should consider the relationship between consumption and the unit price. When the water price increases or decreases in the analysis, sales revenues may be affected by associated changes in the price of water and the average amount of water consumed.

In Excel, one-way data tables can be used to conduct sensitivity analyses for each critical assumption. Two-way data tables can pair related critical assumptions and assess their potential influence on decision criteria when they move together. Start by using expert or peer opinions to identify a defensible range of values for each risk and uncertain variable:

- Some values can be omitted from the range because they are not relevant or feasible. For example, the number of participants cannot be negative.
- Historical data are important in selecting the range of values, but their relevance and availability can vary. An appropriate time frame for the analysis of historical data will depend on the availability of data and its relevance in predicting future values (stability). Recent historical data may be skewed by events with a low likelihood of recurrence. However, systematic or persistent changes can increase future risks, making past experiences less relevant. For example, climate change may make prior weather and hydrological data obsolete. More extensive or more accurate data are needed if the consequences of an incorrect decision are potentially catastrophic.

Then, define a range of viable options for the design parameters. With these ranges in mind, develop sensitivity tables for each critical assumption. It may be possible to shorten the initial list of critical variables to reduce the complexity of the model. Assess the size of the potential impacts of changes in each critical assumption on the decision criteria. This assessment may involve a mix of quantitative and qualitative considerations, reflecting the judgment of the analyst or team. Some suggestions for deciding on the inclusion of a critical variable follow:

- It may or may not be easy to assess whether a certain percent change in a critical variable is likely to increase or decrease the NPV by the same proportion. However, there may be a nonlinear relationship between the risk and decision variables, depending on the size and nature of the change. In addition, the probability of a change of that magnitude may vary for different critical variables.
- An alternative approach is to consider whether the decision criterion exceeds a threshold value (such as a breakeven or zero NPV) within the range of identified values for the critical variable. If this threshold value is reached, the critical variable will be retained. If not, it can be deleted. This approach allows each critical variable to have an independent range, but it can ignore large absolute changes in the decision criteria when their base case value is far below or above the threshold.

- The threshold value and the absolute change approaches can be combined in making decisions on the defensible range for each critical variable.

Even if a risk variable, uncertain variable, or design parameter is excluded from the analysis, it should be discussed in the report and communicated to decision-makers.

3.3 ASSESSMENT OF CO-MOVEMENTS

Co-movements occur when some variables move together in the same or opposite direction. Causation and correlation both translate to co-movement. However, analysts must be able to distinguish between causation and correlation. For instance, sales revenues will likely increase when unit prices or the quantity of a good or service sold increases. However, real price increases may reduce the sales volume, depending on the price elasticity of demand. Conversely, increases in the quantity of a good or service available may reduce market prices. Regardless of the nature of relationship, these positive or negative co-movements should be built into the analysis. Failure to account for co-movements of variables can underestimate or overestimate the variability of decision criteria. Review the list of risk variables and decide whether there are any important co-movements to estimate or integrate. When they exist, either update the model to factor for these co-movements in the calculations or introduce them in the Monte Carlo simulation. When one of the variables that co-move is a decision parameter, then the co-movement must be introduced within the formulation of the model.

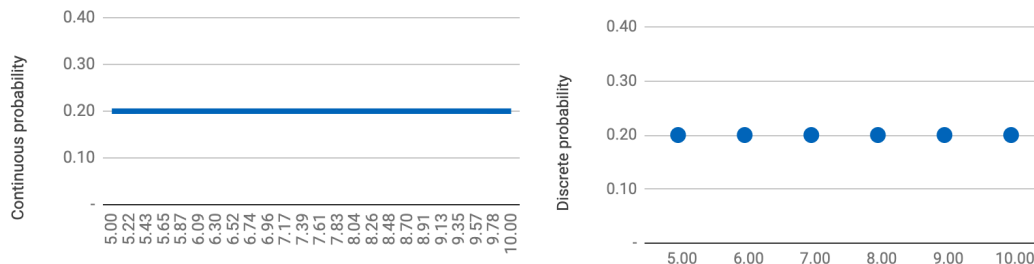
There are many alternative methods for estimating correlations and other statistical properties. When in doubt, consult with a statistician for help choosing analysis methods and estimating correlations, expected values, measures of variability, and other statistical properties.

3.4 SELECTION OF THE PROBABILITY DISTRIBUTIONS

Some risk variables have reasonably well-known variability. The next step is to select the most appropriate probability distribution for the risk variables. The most common probability distributions for a Monte Carlo analysis include:

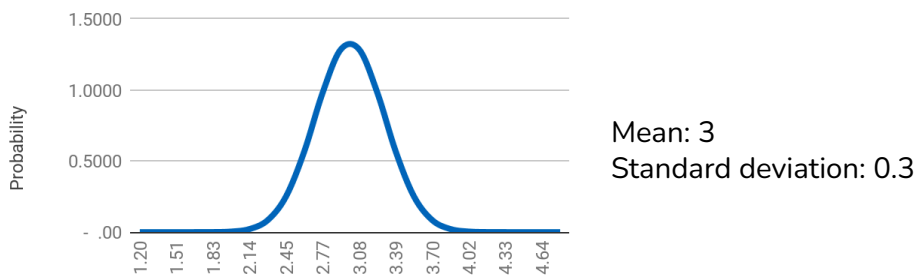
- **Uniform:** Assumes the same probability over a range of possible values. A uniform distribution may be used when minimum and maximum values can be specified, but there is no information about the probability of different values within the range due to the lack of a theoretical basis for a more specific assumption and high uncertainty about specific potential values. Figure 10 illustrates continuous and discrete uniform probability distributions with a minimum value of 5 and a maximum value of 10. The mean of uniform distribution is halfway between its minimum and maximum values.

FIGURE 10. Continuous and Discrete Uniform Probability Distributions



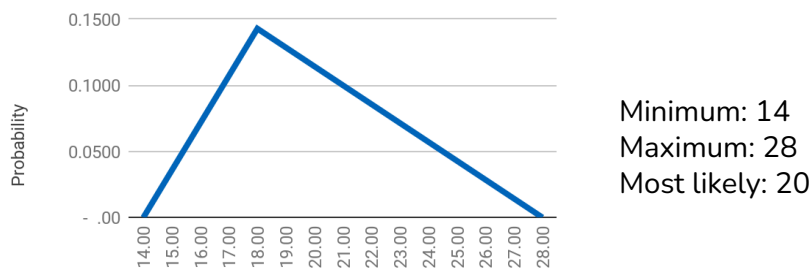
- Normal:** A normal (Gaussian) distribution is bell-shaped and symmetric around its central point, the mean, the value with the highest probability of occurrence. It is open-ended on both sides since there is a nonzero probability that values at either tail ends may occur. A normal distribution is common for many variables in nature. If the underlying population has a normal distribution, a random sample with a large sample size (typically 35 or more for each stratification) is also likely to have a similar shape. The mean is a measure of the central tendency, and the standard deviation measures the dispersion of results around the mean. With a normal distribution, a statistical formula based on the mean, standard deviation, and sample size can be used to estimate the confidence level and confidence interval. The confidence level is the probability that the actual value of a sampled parameter falls within a specified range of values. The confidence interval indicates the upper and lower bounds of the estimate you would expect to find at a specific probability. Figure 11 illustrates a normal probability distribution.

FIGURE 11. Normal Probability Distributions



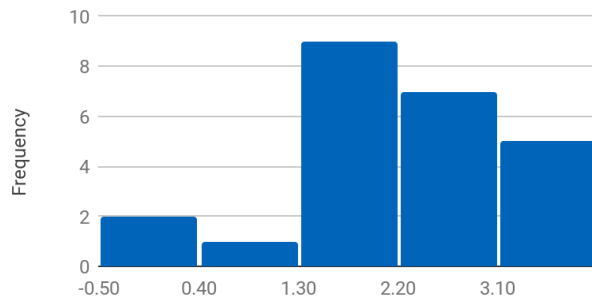
- Triangular:** A triangular distribution has a minimum value, maximum value, and a most likely value. This distribution is symmetric if the most likely value is halfway between the minimum and maximum values. With a triangular distribution, values in the tails have a higher probability of occurrence for values than in the normal distribution. Figure 12 illustrates a nonsymmetric triangular distribution.

FIGURE 12. Triangular Probability Distributions



- **Histogram:** A histogram (step distribution) shows the frequency of data items in successive numerical intervals of equal size. Usually, the independent variable is shown on the horizontal axis and the dependent variable is along the vertical axis. A histogram can be produced by spreading data observations into several buckets covering the range of possible values. The more observations in a bucket, the higher the probability for that bucket. Dividing the frequency of a bucket by the total number of observations will produce the probability for a given bucket. Figure 13 illustrates a histogram.

FIGURE 13. Step Probability Distributions (Histogram)



Florio (2014) discusses these four probability distributions and other types such as Weibull, exponential, Bernoulli, and binomial.

When sufficient data on a variable are available, it is relatively easy to select an appropriate probability distribution. When specific data do not exist, they can often be transferred or extrapolated from other sources. For example, although there will not be any historical data on the net revenues of a new product, it may be possible to estimate the production costs based on input requirements and unit costs, and projected sales revenues. There may be more uncertainty around the projected sales revenues than the production costs that could be simulated with an assumed probability distribution.

In more complex cases, Florio (2018) and Jenkins et al. (2011) recommended that CBA practitioners use their judgment in selecting probability distributions. However, some economists are not experts in statistics and may be unaware of the implications of choosing one probability distribution over another. If sufficient resources are available, it may be desirable to consult a statistician or technical expert before selecting less common or more complex probability distributions.

Probability distributions can be estimated from mean values and variability measures, such as a confidence interval, a p-value, variance, or standard deviation. With some assumptions, these variability measures can be converted to standard errors, which are useful in describing a normal distribution in Monte Carlo simulations.

It may be possible to find relevant probability distributions or summary statistics to estimate them from previous CBAs (especially ex-post CBAs), performance or impact evaluations, or systematic reviews that summarize the quantified evidence on specific treatments. Availability of information on probability distributions may vary across sectors; systematic reviews are often done in health. Primary or secondary data may be available for producing a histogram.

In forecasting values for a risk variable over time, the estimation process can produce a trend (known rate of increase or decrease) and an error term (deviation from the forecast value). In such

circumstances, the relevant distribution for risk analysis would be the distribution of the error term since the trend can be in the deterministic model.

Annex D shows how to estimate a probability distribution for natural gas prices and the inflation rate.

In searching for or estimating a probability distribution, it is important to identify the correlation coefficients for any co-movements of variables. Annex E discusses an example of estimating a correlation coefficient between two variables.

3.5 IDENTIFICATION OF THE RESEARCH QUESTIONS AND METHODS

Next, consider the questions and methods for the risk and uncertainty analysis. The method of choice might require additional changes to the model. The questions below may help structure this review:

- What are the probabilities of success and failure for the investment?
- What are the most critical sources of uncertainty?
- What are the tradeoffs between risks and net returns for the intended beneficiaries and other stakeholders?
- How much risk is acceptable for the intended beneficiaries and other stakeholders? Risk tolerance and loss aversion may vary with their incomes and other characteristics.
- What approaches can be used for reducing or managing risks? How feasible are these approaches? What are the recommended approaches?
- What is the value of reducing risk and uncertainty? Boardman et al. 2018 proposed the application of the concept of quasi-option value.

It may not be possible or feasible to reduce some risks. Risks can also be managed through transferring to other parties by purchasing insurance, establishing risk sharing or pooling arrangements, government or donor indemnification, or providing guarantees to sources of financing. A CBA can help in assessing the financial and economic viability and sustainability of the selected risk management approaches. Such analysis makes a CBA more complicated because it may need additional stakeholders and cashflows contingent on whether a risk occurs. For example, the cost of a catastrophic event can be borne by investors or an insurance company, depending on the existence of an insurance contract. For the CBA model to inform the decision on purchasing insurance, the model needs to include the cost of the insurance and the probability and the cost of the catastrophic event as parameters in calculating the investment criteria.

Monte Carlo simulations are a good option when probability distributions are available for most risk variables. If the probability distributions are unknown, scenario analysis may be preferable. Even when it is possible to run Monte Carlo simulations, scenario analysis might be easier to communicate to decision-makers and other stakeholders. Therefore, these choices are not mutually exclusive.

Expected value analysis can help when risk probabilities are well understood or reasonably known, but this tool is more problematic or subjective when the probabilities are unknown due to high uncertainty. When the probabilities are known or can be estimated, they are applied to the projected costs and benefits. For example, if an event has a 10 percent probability of occurring, then the expected cost of the event is the product of the cost and 10 percent. In its simplest form, expected value analysis can only be used to estimate the mean value for investment criteria such as NPV. Monte Carlo simulations can complement expected value analysis by the probability of obtaining favorable or unfavorable outcomes.

There is another use case for expected value analysis. When uncertainty is high, and a single uncertain variable plays a critical role in making the NPV positive or negative, expected value analysis can still help by reporting threshold values for the critical uncertain variable. For instance, the expected value analysis can show how high the probability of an adverse event should be before an investment to reduce the risk or impact is economically justifiable. Some additional examples:

- In deciding whether an institutional reform is justified, the most critical uncertain variable might be the chance of success. While it is impossible to predict this chance, most experts can agree that the chance is above a certain threshold, such as one percent. If the expected value analysis shows that the NPV is positive even at a one percent chance of success, the analysis will support the decision to invest in the reform. Similarly, if the expected value analysis shows that the NPV turns positive only when the chance of success passes 90 percent, it supports rejecting the investment.
- In deciding whether insurance against flooding is necessary for an electricity generation investment, the chance of having a flood becomes the most critical uncertainty factor. While finding a probability of having a flood is difficult, most experts can agree whether the chance is above or below some thresholds.

Table 3 summarizes the reasons to consider and avoid each method for analyzing risk and uncertainty.

TABLE 3: Considerations When Selecting Methods		
METHOD	Reasons to consider	Reasons to avoid
Sensitivity Analysis	<ul style="list-style-type: none"> ● Helps with checking the model and identifying the critical assumptions ● When producing switching values is important 	<ul style="list-style-type: none"> ● None
Expected Value Analysis	<ul style="list-style-type: none"> ● Whenever the likelihoods are known for the events that drive the costs and benefits ● When a single and uncertain event acts as the main decision factor 	<ul style="list-style-type: none"> ● When there is more than one major source of uncertainty
Scenario Analysis	<ul style="list-style-type: none"> ● When there are uncertain variables in the picture ● When communication is important 	<ul style="list-style-type: none"> ● None
Monte Carlo Simulations	<ul style="list-style-type: none"> ● When it is important to quantify the confidence level in the results of the CBA ● When evaluating alternative risk mitigation solutions 	<ul style="list-style-type: none"> ● When the majority of variability is driven by uncertain variable (with unknown likelihoods)

3.6 ANALYSIS OF RISK AND UNCERTAINTY

MODEL PREPARATION

A CBA model may need to be modified before conducting Monte Carlo simulations or scenario analysis. Think about the following items when preparing the model:

1. Are all parameters and formulas in place to answer the questions for the risk and uncertainty analysis? If not, change the model parameters, data, formulas, or outputs as necessary. New decision criteria could include confidence intervals or threshold values.
2. Include expected value analysis by calculating expected costs and benefits, as needed.
3. When a forecast of a risk variable includes a time trend, it may be best to introduce a deviation term in the formulation of the CBA for risk analysis. In a deterministic model, the expected deviations are zero. However, a probabilistic analysis allows assigning a probability distribution to the deviation term.

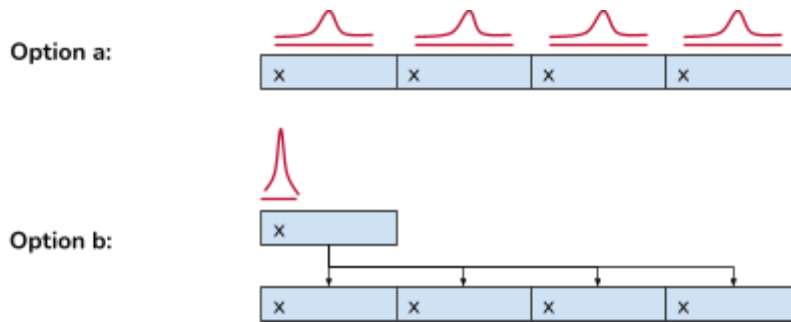
The example below shows how price variations can be included directly in a formula or through a deviation parameter. The price has the subscript “t,” meaning that it can differ in each period. Therefore, if the price is a risk variable, it will require a different probability for each period since the expected value should differ from one period to another. However, introducing a deviation parameter, the second formula enables the application of one single distribution with an expected value of zero for all years.

Without deviation parameter: $Sales_t = Quantity_t \times Price_t$

With deviation parameter: $Sales_t = Quantity_t \times Price_t \times (1 + deviation_t)$

4. If the model uses a risk variable in different years, it may be necessary to apply the probability distribution in each period independently. Otherwise, a probabilistic analysis can exaggerate the differences if the variation is applied once for all periods in the time horizon. As a result, a model may need to be revised to allow annual, independent variations. For example, if there are no trends in the real price of a cost item, the deterministic model has one entry for it which is used for all formulas. However, if the price’s variability is such that it could independently deviate around the expected value from one period to another, then one of the following two approaches can be used.
 - a. Apply the distribution independently for each period: Create separate entries for the unit price in each period and apply the probability distribution independently.
 - b. Apply the distribution once but reduce the variability: In the case of normal distribution, divide the standard deviation (σ) with the square root of the number of periods (\sqrt{n}) to reduce the spread and avoid exacerbating the variability.

FIGURE 14. Applying Probability Distributions for Multiple Periods



5. CBA models often build around a single deterministic case based on average or typical values. Deterministic models may yield unrealistic results when less typical or extreme cases occur (low probability events at either side of the tails of a probability distribution). A deterministic CBA would not consider that a participant may stop market farming if it is not sufficiently profitable in some scenarios or years. Unless a quitting trigger is included in the model, the analysis might presume that a farmer will continue operating at a loss over the entire time horizon. As a result, the financial and economic NPVs may show larger losses than are realistically likely to occur. Structuring a CBA model to accommodate important go/no go decisions and contingencies for risks would be complicated. However, analysts should decide whether additional complexities are worth modeling and be aware of the limitations of using an oversimplified model.

SCENARIO ANALYSIS

Start a scenario analysis by identifying the parameters and variables that you expect will have the largest impact on the decision criteria. Then, combine different parameters and variables to create a range of scenarios that are more or less favorable or reflect a range of risk levels or types.

Prepare a narrative to document the scenarios used in the analysis and explain where they fall on the less favorable to more favorable continuum and how the assumptions were derived. The narrative should also justify the number of scenarios analyzed and the range of values included. Is the most pessimistic scenario sufficiently pessimistic or still somewhat optimistic?

Proponents of a project or investment tend to provide assumptions that are too rosy in general or unrealistic for the targeted client group. For example, the only available crop yield data may come from government agricultural research stations with better agroclimatic locations and soils, greater use of high-quality inputs, and workers with more expertise than low-income farmers can provide. As a result, crop yields at agricultural experiment stations can be two or three times what low-income farmers typically obtain. Other assumptions often overly optimistic in CBAs include the capacity use rate of industrial or agro-processing facilities, product prices, and early sales and sale growth rates.

MONTE CARLO SIMULATIONS

Monte Carlo simulations involve four main steps:

1. Selecting the probability distributions for risk variables
2. Setting the decision criteria
3. Running the simulations (generating data on decision criteria)
4. Reporting results

Annex B explains these four steps in more detail. A Monte Carlo analysis can have errors in its design or formulas that will produce incorrect or confusing results. The following considerations can help in identifying errors or explaining unexpected results:

1. **The expected value of the probability distribution for a decision criterion should generally be the same as the deterministic case values**, but differences can occur for the following reasons:
 - a. *The deterministic model uses incorrect values for risk variables.* If the expected value of the probability distribution applied to one or more risk variables does not match the values included in the deterministic case, then the probability distributions introduce new information on the risk variables absent in the deterministic case. To fix this, ensure that the values for all risk variables in the deterministic case match the expected values from distributions used for these variables in Monte Carlo simulations.
 - b. *A risk variable may have an asymmetric impact on a CBA model due to the nature of the investment.* For example, construction delays may postpone the start of operations and the timing of revenues. A delay in the construction of one component can delay production and revenue. However, if a component is ahead of schedule, it might not be possible to begin operations until others have been completed. Therefore, while the probability of delivering a component later or earlier might be the same, early delivery might not have any impact on the timeline of revenue. As a result, the expected value of the decision criteria could differ from their base case values.
2. **There may be calculation errors in the simulations.** Calculation errors can occur when simulations are based on values from a random number generator. Ratios or other formulas that include divisions will generate errors if a denominator equals zero. Functions that involve iterative calculations, such as the internal rate of return function (IRR), can produce numbers that may be incorrect, particularly if the net cashflows switch back and forth between negative and positive values more than once during the time horizon. These problems can be reduced by avoiding using iterative functions like the IRR as a decision criterion in Monte Carlo simulations and setting up simulation runs to skip runs that include calculation errors such as division by zero. Skipping calculation errors is a menu option in Crystal Ball simulation settings. It is also good practice to add tests in formulas to avoid dividing by zero.

3.7 COMMUNICATION (ASSUMPTIONS, METHODS, FINDINGS)

As Table 4 shows, an analysis of risk and uncertainty can generate additional information that cannot be obtained in a deterministic analysis.

TABLE 4: Decision Criteria Produced by CBA	
DETERMINISTIC ANALYSIS	ANALYSIS OF RISK AND UNCERTAINTY
<ul style="list-style-type: none"> • NPV, IRR, ... • Other quantitative estimates such as cost per unit of effectiveness 	<ul style="list-style-type: none"> • Expected NPV, IRR, ... • Switching values • Best case and worst case scenarios for decision criteria • Probability of failure or success • Value of risk mitigation products/solutions

Some recommendations for communicating the results of the analysis of risk and uncertainty follow:

1. **Start with the elephant in the room:** Focus on risk variables that add a high degree of variation to the decision criteria results and uncertain variables that matter to your audience. Examples: When success in completing an investment as planned is the main driver of benefits, present the lowest probability of success at which the CBA's NPV turns from negative to positive. This information would be more important for decision-makers than the NPV estimated at any given probability of success.
2. **Document and explain the assumptions clearly and comprehensively:** Information about defined ranges for sensitivity analysis, selected values for each scenario (from less to more favorable), derived probability distributions for risk variables, and Monte Carlo simulation parameters should be available for review.
3. **Identify the weaknesses in the full report, Executive Summary, and presentation materials:** Every CBA model illustrates a simplified version of the world based on projections about each input variable. Risk analysis, similarly, has limitations in the way it brings in additional information and expands the analysis.
4. **Remain conservative with risk analysis:** It is generally recommended to be conservative when conducting a CBA – account for all costs but only include defensible measures of benefits. Risk analysis is similar; try to stay conservative with statistics and scenarios and avoid downplaying the risk.

Table 5 includes suggested language for communicating the results of analyzing risk and uncertainty.

TABLE 5: Suggested Language for Communicating the Results

RESULTS	SUGGESTION LANGUAGE
Expected NPV	<ul style="list-style-type: none"> • When a single event dominates the calculation of the expected decision criteria, describe the event and its probability of occurrence along with the decision criteria. Example: <ul style="list-style-type: none"> ○ Assuming a 1 percent probability of a mild earthquake per year, the NPV is \$1,000 (12 percent discount rate), and the IRR is 17 percent.
Switching values	<ul style="list-style-type: none"> • Acknowledge that only one variable changes over time in the sensitivity analysis. Example: <ul style="list-style-type: none"> ○ Everything else held constant; this intervention is justified (NPV>0) as long as the market price of maize does not fall below \$10 per unit.
Best case and worst case scenarios for decision criteria	<ul style="list-style-type: none"> • Acknowledge the spread and ensure assumptions are readily accessible. If a single critical value drives the most variation from one scenario to another, consider spelling its value out. <ul style="list-style-type: none"> ○ The intervention's NPV is \$40 and \$1,000 under the assumptions for the less favorable and more favorable scenarios, respectively. The population growth is assumed to be at 0 percent and 5 percent per year in the worst-case and best-case scenarios, respectively.
The probability of failure or success	<ul style="list-style-type: none"> • Report the expected value of the decision criteria and the probability of crossing the threshold. <ul style="list-style-type: none"> ○ The expected NPV is \$10,000. The probability of having a negative NPV is 6.7 percent. OR ○ The expected NPV is -\$10,000. The probability of having a positive NPV is 42 percent
Value of risk mitigation products/solutions	<ul style="list-style-type: none"> • Explain the risk management or mitigation product/solution and its value in the narrative <ul style="list-style-type: none"> ○ Insuring against the accident costs \$1,000, and the expected cost of the accident is \$500 (5 percent probability of occurrence at the cost of \$20,000) ○ A minimum grant of \$1,000 per farmer is required to reduce the probability of a loss to 0 percent hence increasing the program participation from 400 to 2,000 farmers. This incentive payment increases the program cost by \$2,000,000. However, it is expected to generate resilience benefits valued at \$4,000,000.

ANNEX A: MORE ON EXCEL DATA TABLES

Note that instructions for sensitivity analysis using simple data tables in Excel (with one input and one decision criterion) is in section 2.2 of the report.

FIGURE 15. Sensitivity Analysis: Creating an Excel Data Table with Multiple Decision Criteria

Multiple decision criteria can be included in a data table by adding more columns (or rows) with linked cells to a vertical (or horizontal) data table..

	A	B	C
1		=E20	=E21
2	10		
3	20		
4	30		
5	40		
6	50		
7			

or

	A	B	C	D	E	F
1		10	20	30	40	50
2	=E20					
3	=E21					
4						
5						
6						
7						

The rest of the process is identical to creating a simple Excel data table (see section 2.2).

FIGURE 16. Sensitivity Analysis: Creating an Excel Data Table with two Inputs

Excel data tables can show the sensitivity of one decision criterion to two critical assumptions. These tables will have the alternative values for one input in the first column and the alternative values for the second input in the first row. The linked cell for the calculation sits in the first cell of the first column. As with simple data tables, the entire table needs to be selected before clicking on the data table menu. Once prompted to locate the input reference, place the reference that corresponds to alternative values in the first column in “Column input cell” and the one that corresponds to the first row in “Row input cell.” Press OK, and Excel will fill the empty cells.

	A	B	C	D	E	F
1	=E20	10	20	30	40	50
2	1					
3	2					
4	3					
5	4					
6	5					
7						

or

Data Table

Row input cell:

Column input cell:

FIGURE 17. Sensitivity Analysis: Formatting Excel Data Tables

Formatting data tables can help with more accessible communication of their content. Below are steps for suggested formatting:

1. This two-way data table shows the sensitivity to the product of A and B, where A and B are set to 30 and 3 at the start. This table needs better formatting.

	A	B	C	D	E	F
1	90	10	20	30	40	50
2	1	10	20	30	40	50
3	2	20	40	60	80	100
4	3	30	60	90	120	150
5	4	40	80	120	160	200
6	5	50	100	150	200	250
7						

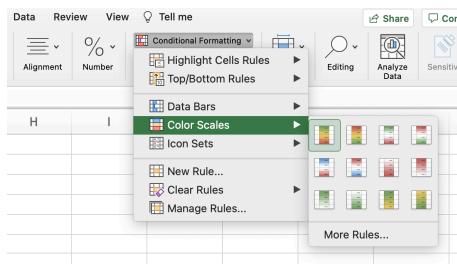
2. The first issue is the lack of clarity about the labeling and styling of calculations, alternative input values, and alternative decision criteria values. Basic cell formatting and descriptions can address this issue.

	A	B	C	D	E	F	G
1			A				
2	A*B=	90	10	20	30	40	50
3	B	1	10	20	30	40	50
4		2	20	40	60	80	100
5		3	30	60	90	120	150
6		4	40	80	120	160	200
7		5	50	100	150	200	250
8							

3. Highlight the base case values.

	A	B	C	D	E	F	G
1			A				
2	A*B=	90	10	20	30	40	50
3	B	1	10	20	30	40	50
4		2	20	40	60	80	100
5		3	30	60	90	120	150
6		4	40	80	120	160	200
7		5	50	100	150	200	250
8							

4. Lastly, you can select the range of values and use Excel's built-in conditional formatting feature to highlight the areas that need more attention.



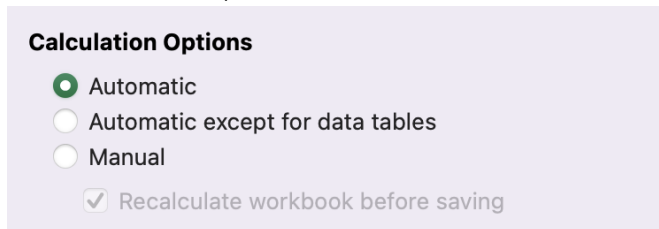
	A	B	C	D	E	F	G
1			A				
2	A*B=	90	10	20	30	40	50
3	B	1	10	20	30	40	50
4		2	20	40	60	80	100
5		3	30	60	90	120	150
6		4	40	80	120	160	200
7		5	50	100	150	200	250
8							

Important note: using custom colors and formatting can reduce the model's accessibility. Only use the built-in styles in Excel unless USAID or your unit provides spreadsheet styling guidelines.

TROUBLESHOOTING EXCEL DATA TABLES

The most common issues in working with Excel data tables and their underlying causes are listed below.

- Excel errors: Input cell reference is not valid
 - The input cell and calculation cells are not in the same sheet
 - The input cell is a formula rather than an entry
- Data table content does not change, shows a constant figure
 - The formula does not link to the input. Test whether manually changing the input in its original place changes the calculation results.
 - Excel allows the user to suppress data tables calculations to speed up the calculations in large spreadsheets. You can check this setting under Excel's preferences. Look for the Calculation menu (the exact name can be different based on your version of Excel)



- Excel does not allow partial deletion or updating of data tables
 - It is not possible to delete part of the data in a data table. You need to delete all of the data (except for alternative values and the linked references from the table) before you can recreate the content.

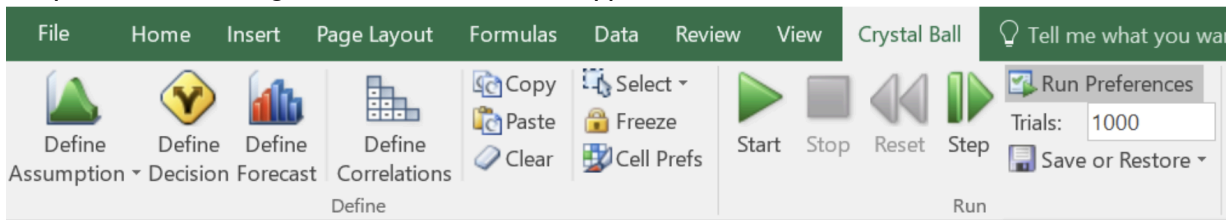
ANNEX B: USE OF THE CRYSTAL BALL SOFTWARE

INSTALLATION

This annex includes a series of practical steps for using Crystal Ball software to conduct Monte Carlo simulations.

FIGURE 18. Crystal Ball Software Menu

Oracle Crystal Ball is an add-on application for Excel. This add-on is currently only available for Windows-based personal computers. Pricing information is available at <https://www.oracle.com/applications/crystalball/>. After Crystal Ball has been installed on your computer, the following additional menu tab will appear in Excel.



STEPS FOR USING CRYSTAL BALL SOFTWARE

FIGURE 19. Crystal Ball Step 1: Selecting the Probability Distributions

Crystal Ball has a series of built-in probability distributions (shown at right). You can either select one of the built-in distributions or define a custom distribution. Select a cell and click on the “define assumption” button to add a probability distribution to a cell that contains a critical variable. Name your assumptions to make your results easier to interpret and report.

A cell with an assigned probability distribution will have a different background color. Click the “clear” button in the menu bar to delete the associated Crystal Ball data.

To define a correlation between two variables, select one variable and click the “correlate” button. Then select other critical variables that already have an assigned probability distribution and choose a correlation coefficient between -1.0 and 1.0. Since each correlation is between two variables, defining a correlation for one variable automatically associates it with the other variable.

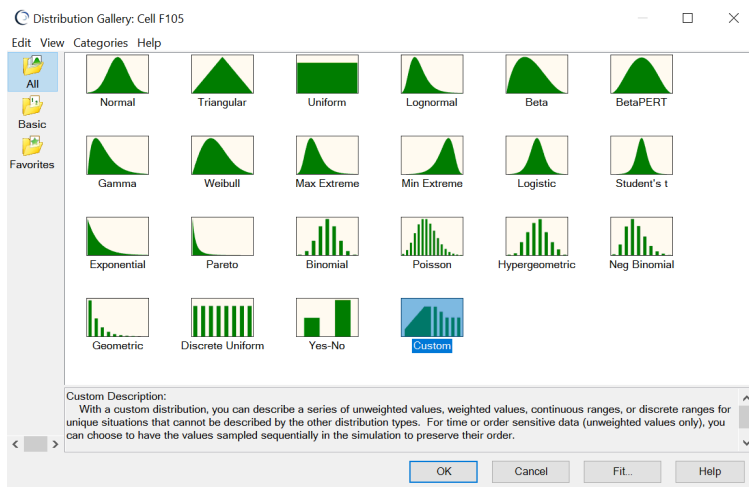


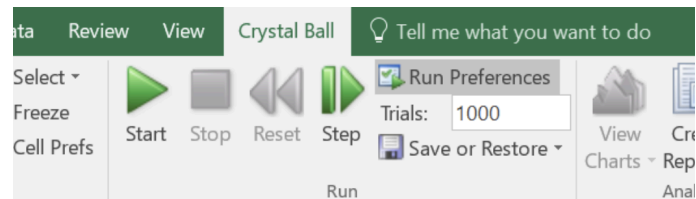
FIGURE 20. Crystal Ball Step 2: Setting the Decision Criterion

Choose the decision criterion by selecting the cell with the calculation and clicking on “define forecast.” Next, name the decision criteria and specify the units. Decision criteria will have a different background color. To delete the associated Crystal Ball information from the cell, click the “clear” button in the Crystal Ball menu.

FIGURE 21. Crystal Ball Step 3: Simulation Runs

After the assumptions (risk variables) and forecasts (decision criteria) have been defined, click “trials” to set the desired number of simulation runs. The probability that the simulation results reflect the underlying population correctly increases with the number of runs. A larger number of runs also increases the potential accuracy of the summary statistics. However, running a larger number of simulations will require more processing power and take more time. Nevertheless, a typical, recent personal computer can easily handle 10,000 simulation trials in Crystal Ball.

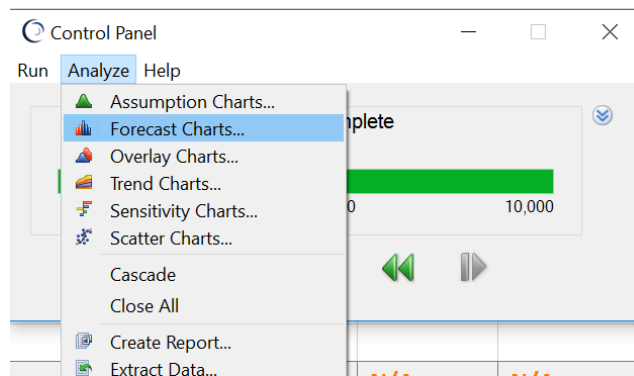
In testing the setup and functioning of the simulations, it is fine to start with a smaller number of trials (such as 100). After confirming the setup, you can increase the number of simulation trials.



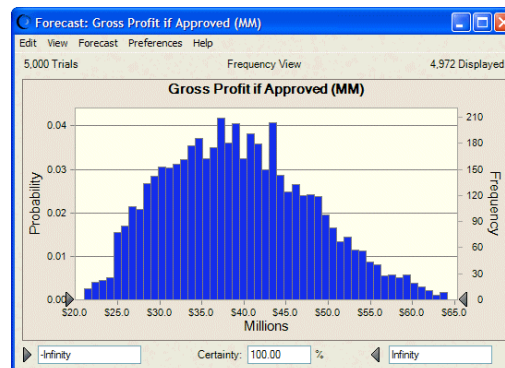
Run simulations by clicking the “start” button. Crystal Ball uses histograms to visually display the probability distribution for the decision criterion. A *histogram* is a graph that shows the frequency of numerical data using rectangles. The rectangle’s height (vertical axis) represents the distribution frequency of the variable. The histograms will look “bumpier” when fewer simulation trials are used. You can stop, reset, and continue the simulation while running the simulations. If you stop the run, you can see the assumption values and calculations for the current simulation run. You can monitor simulations one by one using the “step” button after stopping the run.

FIGURE 22. Crystal Ball Step 4: Generating Charts and Reports

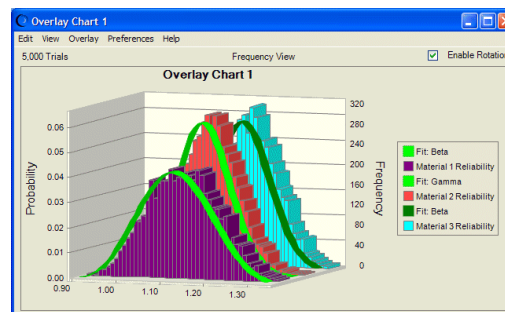
After the simulation runs have been completed, Crystal Ball can automatically generate charts for the assumptions. It can also prepare charts on the results in various visual formats -- forecast charts, overlay charts, trend charts, sensitivity charts, and scatter charts.



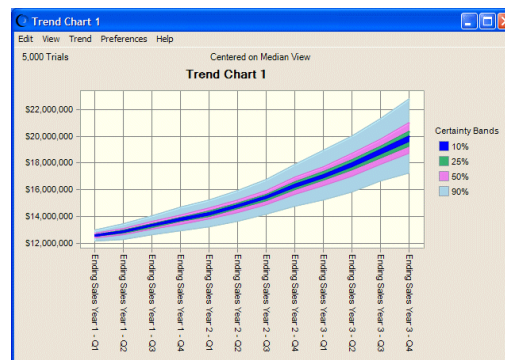
- *Forecast charts* show the probability distribution for the decision criterion. Forecast charts can be displayed as histograms, cumulative histograms, or bar or line graphs. Users can also define a range on the chart, and Crystal Ball will estimate the probability of being within that range.



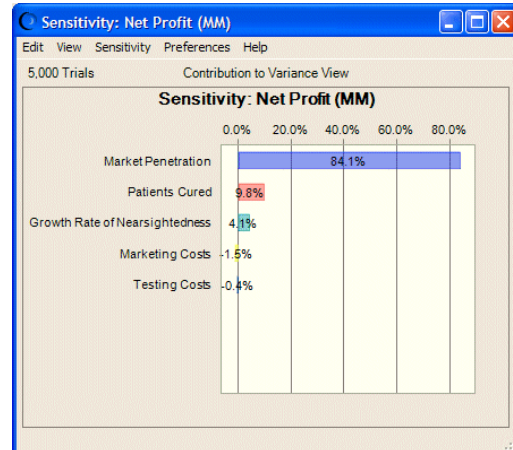
- *Overlay charts* can place multiple forecasts on one chart. Overlay charts help compare forecasts that have similar scales and the same threshold, such as the NPV, from alternative perspectives.



- *Trend charts* can present the certainty range for multiple forecasts on one chart. This chart is most useful when the selected forecasts form a time series, such as the net cash flow in year 1, year 2, and so on.

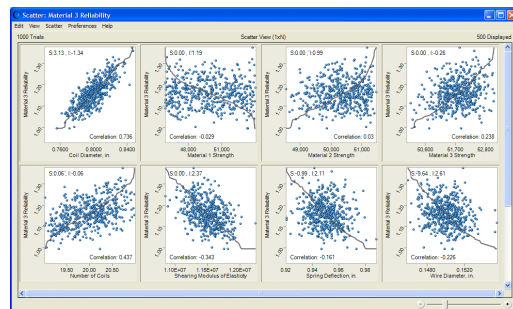


- *Sensitivity charts* (sometimes called tornado charts) are useful in identifying the risks that contribute the most to the variation in a decision criterion. In other words, they identify the most critical source variations (risks).



- *Scatter chart* plots correlation between pairs of variables (assumptions and forecast).

Find the charts by clicking on the “view charts” button. Delete observations using the “reset” button and rerun the simulations.



Note: Software screenshots from Oracle

COPY-AND-PASTING ASSUMPTIONS

The Monte Carlo simulation setup might require the application of the same probability distribution on many cells. An example is the price of an input with the same distribution but applied independently for each year. It is possible to copy and paste Crystal Ball data using the copy-and-paste buttons in Crystal Ball’s menu. Please note that Excel’s copy-and-paste function cannot duplicate Crystal Ball assumptions.

Crystal Ball allows for using formulas when setting a name for an assumption. This feature is helpful when copy-pasting assumption data, where the name can automatically update from one cell to another using dynamic references to other cells that contain the period count or other qualifiers.

ANNEX C: GLOSSARY

TABLE 6: Glossary	
TERM	DEFINITION
Deterministic Case	The most likely scenario of data and assumptions for analysis in the absence of risk and uncertainty.
Break-Even Analysis	Break-even analysis estimates the values of critical assumptions that will result in an NPV of zero. Businesses often conduct financial analyses to estimate break-even prices and break-even unit sales. Break-even values are often called switching values since they are inflection points between negative and positive NPVs.
Break-Even Price	The price of a good or service that would result in a present value of net revenues of zero at a particular level of production.
Break-Even Quantity	The volume of a good or service produced that would result in a present value of net revenues of zero at a particular unit price.
Critical Assumptions	design parameters, risk, and uncertain variables.
Decision Criteria	Decision criteria are the outputs of CBA models. The most common decision criterion is NPV.
Design Parameter	A design parameter for an intervention may include such factors as the scale, duration, intensity, approach, and targeting.
Risk	A type of uncertainty, the part that is calculable and quantifiable using probabilities and losses. Marchau et al. (2019)
Risk Variable	An assumption in CBA that 1) is variable, 2) its variability is known, and 3) its underlying variability can drive the CBA to produce substantially different results.
Uncertain Variable	An assumption in CBA that 1) is variable, 2) its variability is unknown, and 3) its underlying uncertainty can drive the CBA to produce substantially different results.
Uncertainty	The gap between available knowledge and the knowledge decision makers would need to make the best policy choice. Marchau et al. (2019)
What-if Analysis	What-if analysis examines the effects of changing one or more parameters on the decision criteria findings.
Monte Carlo Analysis	Monte Carlo analysis or simulation is the process of repeated random sampling out of a process to generate a sample for the output of the process. A CBA model is a process and repeated assignment of random values for one or more of its assumptions can produce a sample population for the decision criteria such as the NPV. Software programs such as Oracle Crystal Ball facilitate Monte Carlo simulations on Excel.

ANNEX D: PARAMETERIZATION EXAMPLE

Advanced forecasting models exist for various economic and non-economic indicators. While CBA practitioners do not need to be statistical analysts or forecasting experts, some skills in the area of parameterization can help practitioners in finding estimates, qualifying existing estimates, and identifying the areas where external help is necessary for parameterizing a risk variable.

This annex provides step-by-step instructions for using historical data to inform the probability distribution of two risk variables.

1. *The US inflation rate:* the inflation rate can introduce financial risk for private sector actors, such as farmers and utility companies. Therefore, the inflation rate can impose indirect risk on the economic benefits and development outcomes of investment. Based on economic theory, it is reasonable to assume that there can never be long-term trends in the inflation rate. A downward trend will lead to deflation, which is hard to support theoretically or observe in real life. An upward trend in inflation will result in the collapse of a currency. Therefore, trends in the inflation rates can only be temporary when countries recover from a high or low inflation period. However, we can use historical data to learn about the inflation rate's variability around the expected value. The US inflation rate has been relatively stable over the period studied in this example. Therefore, we do not need to worry about outliers—events with an extremely low probability of recurring.
2. *The price of natural gas:* many studies have forecasted trends in the price of natural gas. The CBA can use existing literature to inform the trend. This example focuses on using historical data to forecast the variability of the price.

Historical data for the US inflation rate and the price of natural gas is available on the websites of the World Bank and Energy Information Administration (EIA), respectively. The spreadsheet accompanying this guide contains the data and the formulas applied following the below steps to estimate a histogram distribution for each risk variable.

INFLATION RATE

1. Choose the number of bins for histogram distribution: the general rule of thumb is to have just enough bins so that two to the power of the number of bins exceeds the number of observations.
2. Find the minimum and maximum values: these values set the starting point of the first bin and the ending point of the final bin.
3. Construct the bins: the size of each bin is the gap between the minimum and maximum value divided by the number of bins.
4. Estimate probability for each bin: for each bin, count the number of observations in each bin and divide the number by the total number of observations.

NATURAL GAS PRICE

1. Estimate the price index: market prices are in current (nominal) currency and include the effect of currency devaluation over time (inflation). The first step is calculating price indices for converting these prices to constant (real) currency.
2. Convert nominal prices to real prices: divide the nominal prices by the price index.
3. Fit a trend line: the trend is another source of variability that can be modeled using a regression.

4. Calculate percentage deviation: the absolute deviation is the distance between the observed values and the trend line. Dividing the difference by the observed value produces the percentage deviation.
5. Choose the number of bins for histogram distribution: same as the case for inflation.
6. Find the minimum and maximum values: same as the case for inflation.
7. Construct the bins: same as the case for inflation.
8. Estimate probability for each bin: compared to inflation, in this case, the variable is standard deviation, which requires an expected value of zero. Since the dataset is limited in size, the observations in bins are not perfectly balanced, and the expected value can be non-zero at first. The best approach to address this problem is to manually, or using Excel's Solver), adjust the frequencies (probabilities) until the expected value falls as close as possible to zero. Please note that the sum of the probabilities must still come to 100%.

Please note that these steps are not relevant for every risk variable. Expert knowledge of economics and statistics is necessary to judge if and how much information can come from historical data for every risk variable.

ANNEX E: ESTIMATING CORRELATION

Risk variables can be correlated with each other or with other critical assumptions, such as design parameters. Excel and Google Sheets offer the CORREL() function, which accepts two data series and reports the coefficient of correlation. There is anecdotal evidence suggesting that fertilizer consumption in Rwanda fluctuates significantly due to its price. Since natural gas is the main input for urea fertilizer production, this exercise focuses on assessing the correlation between the price of natural gas and fertilizer consumption by farmers.

The accompanying spreadsheet contains the data for prices, consumption, and inflation. The inflation rate is used to convert current prices into real prices so that the correlation analysis is not diluted by variation in price due to inflation. The spreadsheet follows these steps in calculating the correlation.

1. Estimate the real price of natural gas: division by the price index can help convert nominal prices to real prices.
2. Find correlation: use Excel or Sheet CORREL() function to estimate the correlation (1: perfectly and positively correlated, -1: perfectly and negatively correlated).
3. Test if crude oil prices can help: oil prices have a long history compared to natural gas prices. Therefore, this final step checks whether oil prices can replace natural gas prices for estimating the volatility in consumption. The ability to use the oil price instead depends on its degree of correlation with natural gas prices.

ANNEX F: CASE STUDY

A water and sanitation project is the basis for this practice case. The CBA includes a financial analysis that focuses on a water vendor and an economic analysis that brings in the perspectives of consumers, labor, government, and USAID. The analysis of risk and uncertainty for this project follows the seven steps presented in the document.

STEP I: IDENTIFICATION OF CRITICAL ASSUMPTION

- The price and quantity of water sold by the vendor have an important role in determining the financial and economic feasibility of the investment. Therefore, the following assumptions will be considered for the first list of critical assignments.
 - Price of water per liter
 - Share of the targeted population that decides to switch to using potable water
 - Average consumption of water per person per day
- On the cost side, the investment cost of infrastructure projects is subject to overrun, and there is evidence suggesting that the material cost price can fluctuate from one year to another.
 - Investment cost overrun
 - Material cost
- The health benefits rely on expectations about how the investment will affect the prevalence of diarrhea and the mortality associated with child diarrhea
 - Child mortality associated with diarrhea

Based on preliminary analysis, we can place the critical assumptions in three buckets:

- **Decision parameters:**
 - Price of water per liter - given that the price is regulated, we can treat it as a decision parameter
- **Risk variables:**
 - Average consumption of water per person per day has been studied by many and is a function of the price. We can have a confidence interval for this variable.
 - Investment cost, in this case, is the price of the equipment. Historical data on equipment price can help us estimate a probability distribution for this variable.
 - Material cost is similar to equipment cost. Its historical data can help us build a probability distribution for it.
 - Child mortality associated with diarrhea comes from secondary literature along with a standard deviation.
- **Uncertain variables:**
 - The share of the targeted population that switches to potable water is a major source of uncertainty. Existing data from past programs are not available or relevant.

STEP 2: SENSITIVITY ANALYSIS

The model needs the analytical capacity to facilitate sensitivity analysis for each critical assumption. The main change applied to the model at this point is adding a cost overrun factor. Variation in investment cost is best described using percentages rather than absolute figures. The cost overrun factor is added to the inputs sheet, then integrated into the financial and economic analysis using the following formula.

$$\text{Investment cost} = \text{original investment cost} \times (1 + \text{cost overrun factor})$$

At this step, you need to ask similar questions about all critical assumptions. For instance, the model assumes that the costs remain the same irrespective of the number of people who switch and the expected water sales per person. This assumption is not realistic in the real world. However, we leave this as-is for this simplified case study.

The input and the sensitivity (data) tables must be on the same sheet. There are two options: 1) construct the data tables in the inputs sheet, or 2) relocate the critical parameters to a different sheet where the data tables will be located. Let's bring the critical assumptions to the "risk analysis" sheet. You can maintain the model's integrity by following the steps below for each input:

1. Copy and paste the critical assumption from the "inputs" sheet to the "risk analysis" sheet.
2. Replace the exogenous entry for the critical assumption in the "inputs" sheet with a link to its new location in the "risk analysis" sheet.
3. Adjust the formatting to clarify that the "inputs" sheet includes a linked cell.

With the right analytical capacity in place, the next step is setting a range for each variable:

TABLE 7: Ranges for Sensitivity Analysis		
Critical Assumption	Range	Notes
Price of the water per liter	0.1 - 0.5	The only limitation for this decision parameter is that the price cannot be negative.
# of people who switch to purchasing potable water	50,000 - 270,000	150,000 is the assumed number. This number is 50% of the population. The range is selected to go from 33% to 80%. However, there is no evidence to support any number in this range.
Expected sales of water per person per day	2.33 - 4.33	Data collection during the design stage suggests that potable water consumption can range from 2.33 to 4.33 liters per day per person
Investment cost overrun factor	-10% - +10%	Historical data on the price of equipment suggest that the cost can fluctuate by no more than 10% in each direction at 95% confidence level.
Annual material cost before import duty in base year	300,000 - 400,000	Historical data on the material cost suggest that they can fluctuate by no more than 50,000 USD per year in each direction at a 95% confidence level.
Diarrhea-related case fatality rate	0.08% - 0.24%	The evidence suggests that this estimate is between 0.08% and 0.24% at a 95% confidence level.

The main investment criteria in this CBA are the financial and economic NPVs. Therefore, one-way sensitivity tables only need to estimate alternative values for these two NPVs. The data tables include linked cells for these calculations.

Sensitivity analysis helps in many ways, including filtering out parameters that are not critical and locating sources for obtaining probability distributions for risk variables. Ultimately, sensitivity analysis informs the research questions and the analytical approach for the analysis of risk and uncertainty.

The sensitivity tables can help answer the following questions:

- Does the resulting change make theoretical sense? If not, fix the errors in the model.
- Is the resulting change negligible? If yes, drop the parameter but acknowledge the reason.

TABLE 8: Sensitivity Analysis Results						
Price of the water per liter						
	0.10	0.20	0.30	0.40	0.50	Notes: Decision parameter
Fin NPV	(10.30)	48.01	106.32	164.62	222.93	
Econ NPV	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)	
# of people who switch to purchasing potable water						
	50,000	105,000	160,000	215,000	270,000	Notes: Uncertain variable
Fin NPV	(29.73)	13.03	55.78	98.54	141.30	
Econ NPV	(8.42)	(4.01)	0.40	4.81	9.22	
Expected sales of water per person per day						
	2.33	2.83	3.33	3.83	4.33	Notes: Risk variable
Fin NPV	53.79	80.05	106.32	132.58	158.85	
Econ NPV	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)	
Investment cost overrun factor						
	-10%	-5%	0%	5%	10%	Notes: Drop from the list
Fin NPV	109.62	107.97	106.32	104.67	103.02	
Econ NPV	(0.20)	(0.30)	(0.40)	(0.50)	(0.60)	
Annual material cost before import duty in base year						
	300,000	325,000	350,000	375,000	400,000	Notes: Drop from the list
Fin NPV	110.02	108.17	106.32	104.47	102.61	
Econ NPV	(0.15)	(0.28)	(0.40)	(0.53)	(0.65)	
Diarrhea-related case fatality rate						
	0.08%	0.12%	0.16%	0.20%	0.24%	Notes: Risk variable
Fin NPV	106.32	106.32	106.32	106.32	106.32	
Econ NPV	(2.43)	(1.41)	(0.40)	0.61	1.62	

Some questions that require attention when reviewing the results of the sensitivity analysis:

- Why is the economic NPV not sensitive to the price of water?
- Why is the economic NPV not sensitive to water sales per person?

Both questions point to the analytical limitations of the model. For the first question, the model assumes that consumers do not react to price adjustments, which is not an accurate assumption. Whether or not the model requires an adjustment to accommodate the relationship between price and quantity demanded will depend on the research questions for the analysis of risk and uncertainty. We will come back to this question in step 5.

The answer to the second question is that the main benefits come from savings associated with reduced incidence of diarrhea and reduced time spent collecting water. While water consumption per person can affect time-saving, this model simplifies the assumption and removes any dependencies between time savings and water consumption.

Note that the sensitivity analysis helped drop investment cost overrun and material cost as critical assumptions since their variation leaves a negligible impact on the investment criteria.

STEP 3: ASSESSMENT OF CO-MOVEMENTS

The only relationship to consider is between price and consumption. Price can impact two critical variables: the number of people who switch to using potable water and the water consumption per person per day. Without data on the former variable, we can't find or estimate a correlation. However, an existing study suggests a weak correlation between the price of quantity consumed by a person equal to **-0.2**. This correlation reflects that at the current levels of potable water consumption, demand is relatively inelastic.

STEP 4: SELECTION OF PROBABILITY DISTRIBUTIONS

Table 9 summarizes the recommended probability distributions.

TABLE 9: Probability Distributions for Risk Variables	
Risk Variable	Distribution
Expected sales of water per person per day	Normal Mean: 3.33 Standard deviation: 0.30
Diarrhea-related case fatality rate	Normal Mean: 0.16% Standard deviation: 0.02%

STEP 5: IDENTIFICATION OF THE RESEARCH QUESTIONS AND METHODS

Key research questions:

- What is the chance of negative NPV from both financial and economic perspectives?
- How does the participation rate threaten the economic viability of the project?
- What is the breakeven price for the vendor?
- Would the project remain financially attractive at a lower price of water? Should USAID consider adding additional subsidies to ensure lower prices and greater consumption?

Recommendations on methods:

- Breakeven points:
 - Participation is an uncertain variable. Therefore, use “inverse” expected value analysis for participation rate to report the minimum participation rate at which the investment remains financially and economically viable.
 - Estimate a break-even price for vendors, the price at which the financial NPV equals zero.
- Scenario analysis: Instead of favorable and unfavorable scenarios, construct an alternative scenario with a lower price point. Then report investment criteria and carry the same scenarios to Monte Carlo simulations (report Monte Carlo results under two scenarios).
- Monte Carlo simulations: Conduct Monte Carlo simulations on both scenarios and then reassess the financial viability of the investment (i.e., does the lower bound of financial NPV stay at or greater than zero at a 95 percent confidence level?).

STEP 6: ANALYSIS OF RISK AND UNCERTAINTY

Calculate the breakeven points for the price (based on financial NPV) and participation rate (based on economic NPV) using Excel’s Solver. The analysis suggests that:

- Price has to drop from 0.30 to 0.12 Maloti before the project becomes financially unattractive
- Project becomes economically viable with a small increase in participation from 150,000 to 156,000 people (~3 percent increase)

Based on this analysis, one wonders if a lower price is justified that can encourage greater participation while maintaining financial attractiveness. The price selected for the low-price scenario is 0.20, above the breakeven price point. In this scenario, we assume that the participation increases from 150,000 to 160,000 (less than a 10 percent increase). A 10 percent increase due to a 33 percent drop in price is a defensible expectation despite the uncertainty surrounding the participation levels. Table 10 summarizes the assumptions and shows the financial and economic NPV for each scenario. Deterministic analysis suggests that the financial NPC remains positive under the lower price scenario. Probabilistic analysis can help understand whether the lower margin increases the financial risk to a level that might make the investment financially unattractive for the vendor.

Earlier, we learned that the demand is relatively inelastic. A 33 percent drop in price translates to a 6 percent change in consumption per person. Sensitivity analysis showed us that a 6 percent increase in consumption would not have a tangible impact on the financial or economic NPV (see table 8).

Therefore, we ignore this correlation in defining our scenarios to keep things simple for communication purposes.

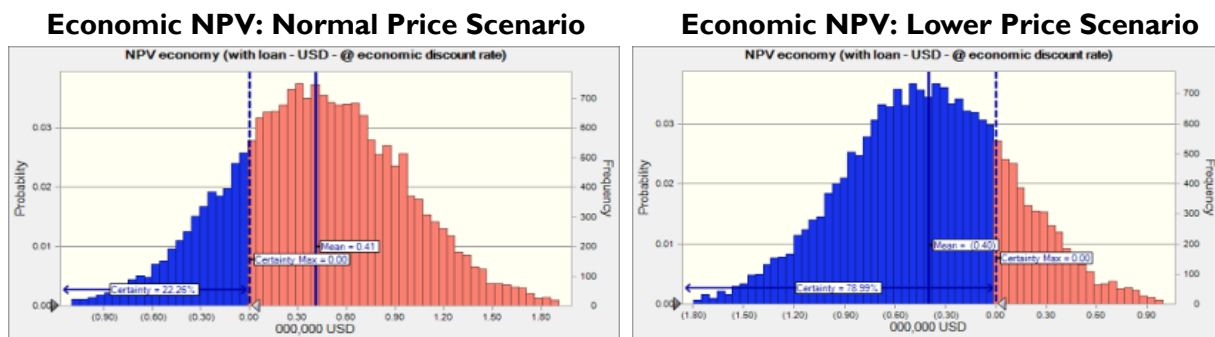
TABLE 10: The Scenarios and the Results of Deterministic Analysis		
Scenarios:	Normal Price	Lower Price
Assumptions:	Price: 0.30 Maloti Participation: 150,000 people	Price: 0.20 Maloti Participation: 160,000 people
Deterministic Results:	Financial NPV: 106.32 Mil Maloti Economic NPV: (0.40) Mil USD	Financial NPV: 55.78 Mil Maloti Economic NPV: 0.40 Mil USD

Next, we conduct Monte Carlo simulations under both scenarios and report not only the mean (expected value) for financial and economic NPV, but also the chance of having negative NPV. Table 11 summarizes the results of the probabilistic analysis.

TABLE 11: The Results of Probabilistic Analysis		
Scenarios:	Normal Price	Lower Price
Probabilistic Results:	Financial NPV: <ul style="list-style-type: none"> • Mean: 106.31 Mil Maloti • Chance of neg. NPV: 0% Economic NPV: <ul style="list-style-type: none"> • Mean: (0.40) Mil USD • Chance of neg. NPV: 78.99% 	Financial NPV: <ul style="list-style-type: none"> • Mean: 55.90 Mil Maloti • Chance of neg. NPV: 0% Economic NPV: <ul style="list-style-type: none"> • Mean: 0.40 Mil USD • Chance of neg. NPV: 22.26%

Figure 22 shows the probability distribution for economic NPV under each scenario.

FIGURE 22. Probability Distribution of Economic NPV



STEP 7: COMMUNICATION (ASSUMPTIONS, METHODS, FINDINGS)

Below are the suggested communication elements for the results of analyzing the risk and uncertainty.

- **What are the results of the deterministic analysis?**
 - Financial NPV (vendor): 106.31 Million Maloti
 - Economic NPV: (0.4) Million USD
- **What are the critical assumptions?** Variations in the following assumptions can threaten the viability of the investment:
 - Price of water (regulated through contracts)
 - # of people who switch to purchasing potable water (a major source of uncertainty)
 - Expected sales of water per person per day (source of risk)
 - Diarrhea-related case fatality rate (source of risk)
- **What assumptions did you drop from the list of critical assumptions? Why?** Historical data suggests that the following costs are reasonably stable, and their expected variation does NOT have a tangible impact on the viability of the investment
 - Investment cost overrun factor
 - Annual material cost before import duty in base year
- **What are the breakeven points for the price and participation rate?**
 - The price charged by the vendor has to drop from 0.30 to 0.12 Maloti before the project becomes financially unattractive
 - Project becomes economically viable with a small increase in participation from 150,000 to 156,000 people (~3 percent increase)
- **Would a lower price make the project economically more attractive while maintaining its financial attractiveness?**
 - Yes, based on the analysis, a 33% drop in the price
 - Increases the expected economic NPV from (0.4) to 0.4 million USD,
 - Reduces the chance of getting a negative economic NPV from 80% to 22%,
 - Reduces the financial NPV from 106 to 56 Million Maloti, while
 - The chance of getting a negative financial NPV stays at zero.

REFERENCES

- Belt, Juan, Nicolas Allien, Jay Mackinnon, and Bahman Kashi. 2017. *Cost Benefit Analysis of Power Sector Reform in Haiti*. Haiti Priorise.
- Boardman, Anthony, David Greenberg, Aidan Vining, and David Weimer. 2018. *Cost-Benefit Analysis: Concepts and Practice*. (5th ed.). Cambridge University Press
- Florio, Massimo. 2014. *Applied Welfare Economics: Cost-benefit of Projects and Policies*. Routledge.
- Jenkins, Glenn, Chun-Yan Kuo, and Arnold Harberger. 2011. *Cost-Benefit Analysis for Investment Decisions*. Kingston, Ontario: JDI Executive Programs.
- Marchau, Vincent AWJ, Warren E. Walker, Pieter JTM Bloemen, and Steven W. Popper. 2019. *Decision Making Under Deep Uncertainty: From Theory to Practice*. Springer Nature.
- Sartori, Davide, Gelsomina Catalano, Mario Genco, Chiara Pancotti, Emanuela Sirtori, Silvia Vignetti, and Chiara Bo. 2014. *Guide to Cost-Benefit Analysis of Investment Projects. Economic Appraisal Tool for Cohesion Policy 2014-2020*. European Commission
- USAID. 2022a. *USAID Risk Appetite Statement—A Mandatory Reference for ADS Chapter 596*. Washington, DC: USAID. https://www.usaid.gov/sites/default/files/documents/596mad_0.pdf
- USAID. 2022b. *ADS Chapter 201—Program Cycle Operational Policy*. Washington, DC: USAID. <https://www.usaid.gov/sites/default/files/documents/201.pdf>