ENHANCED SCENARIO-BASED METHOD FOR COST RISK ANALYSIS: THEORY, APPLICATION, AND IMPLEMENTATION

Mr. Peter Braxton,[1] Dr. Brian Flynn,[2] Dr. Paul Garvey[3], and Mr. Richard Lee[4]

In memory of Dr. Steve Book, nulli secundus, for his kindness and devotion, and for his invaluable comments and insights on an earlier draft.

ABSTRACT

In 2006, the Scenario-Based Method (SBM) was introduced as an alternative to advanced statistical methods for generating measures of cost risk. Since then, enhancements to SBM have been made. These include integrating historical cost performance data into SBM's algorithms and providing a context for applying SBM from the perspective of the 2009 Weapon Systems Acquisition Reform Act (WSARA). Together, these improvements define the enhanced SBM (eSBM) – an historical data-driven application of SBM. This paper presents eSBM and illustrates how it promotes realism in estimating future program costs, while offering decision-makers a traceable and defensible basis behind data-derived measures of risk and cost estimate confidence.

KEY WORDS: Scenario-Based Method (SBM), Enhanced Scenario-Based Method (eSBM), Weapon Systems Acquisition Reform Act (WSARA), Cost Estimate, Cost Risk, Historical Cost Data

[1] Technical Officer and Senior Cost Analyst, Technomics; PBraxton@Technomics.Net.
[2] Senior Cost Analyst, Technomics; BFlynn@Technomics.Net and Brian.Flynn.CTR@Navy.mil.
[3] Chief Scientist, Center for Acquisition and Systems Analysis, The MITRE Corporation, MITRE Paper MP100214, © 2010, All Rights Reserved, Approved for Public Release; Distribution Unlimited, 06-1299, pgarvey@mitre.org.
[4] Cost Analyst, Technomics; RLee@Technomics.Net.

1.0 Background

This paper presents eSBM, an enhancement to the Scenario-Based Method (SBM), which was originally developed as a "non-statistical" alternative to advanced statistical methods for generating measures of cost risk. Both SBM and eSBM emphasize the development of written risk scenarios as the foundation for deriving a range of possible program costs and assessing cost estimate confidence.

SBM was developed in 2006 in response to the following question posed by a government agency: *Can a valid cost risk analysis, one that is traceable and defensible, be conducted with minimal (to no) reliance on Monte Carlo simulation or other advanced statistical methods?* The question was motivated by the agency's unsatisfactory experiences in developing, implementing, and defending simulation-derived risk-adjusted program costs of their future systems.

Once developed, SBM has appeared in a number of publications, including the RAND monograph *Impossible Certainty* [Arena, 2006], the United States Air Force *Cost Risk and Uncertainty Analysis Handbook* (2007), and NASA's *Cost Estimating Handbook* (2008). SBM is also referenced in GAO's *Cost Estimating and Assessment Guide* (2009). It was formally published in the *Journal of Cost Analysis and Parametrics* [Garvey, 2008]. Since 2006, interest in SBM has continued to grow, and the method has been enhanced in two ways. First, historical cost data are now integrated into SBM's algorithms. Second, a framework for applying SBM from a WSARA perspective has been built into SBM. The acronym eSBM denotes SBM together with these two enhancements.

In short, eSBM is an historical data-driven application of SBM operating within WSARA. In support of WSARA, eSBM produces a range of possible costs and measures of cost estimate confidence that are driven by past program performance. With its simplified analytics, eSBM eases the mathematical burden on analysts, focusing instead on defining and analyzing risk scenarios as the basis for deliberations on the amount of cost reserve needed to protect a program from unwanted or unexpected cost increases. With eSBM, the cost community is further enabled to achieve cost realism – while offering decision-makers a traceable and defensible basis behind derived measures of risk and cost estimate confidence.

1.1 Requirement

Life-cycle cost estimates of defense programs are inherently uncertain. Estimates are sometimes required when little if any of a program's total definition is known. Years of system development and production and decades of operating and support costs, need to be estimated. Estimates, in turn, are based on historical samples of data that are almost always messy, of limited size, and difficult and costly to obtain. Herculean efforts are commonly required to squeeze usable information from a limited, inconsistent set of data. And no matter what estimating tool or method is used, historical observations never perfectly fit a smooth line or surface, but instead fall above and below an estimated value. To complicate matters, the weapon system or automated information system under study is often of sketchy design. Only limited programmatic information may be available on such key parameters as schedule, quantity, performance, requirements, acquisition strategy, and future evolutionary increments. Further, the historical record has shown that key characteristics of the system actually change as the system proceeds through development and even production. Increases in system weight, complexity, and lines of code are commonplace.

For all of these reasons, a life-cycle cost estimate, when expressed as a single number, is merely one outcome or observation in a probability distribution of costs. That is, the estimate is stochastic rather than deterministic, with uncertainty and risk determining the shape and variance of the distribution.

The terms “risk” and “uncertainty” are often used interchangeably, but they’re not the same.

- **Uncertainty** is the indefiniteness or variability of an event. It captures the phenomenon of observations, favorable or unfavorable, high or low, falling to the left or right of a mean or median.
- **Risk** is exposure to loss. In a defense acquisition context, it is “a measure of future uncertainties in achieving program performance goals within defined cost and schedule constraints. It has three components: a future root cause, a likelihood assessed at the present time of that future root cause occurring, and the consequence of that future occurrence.” [5]

Risk and uncertainty are related. Uncertainty is probability while risk is probability and consequence.

1.2 Techniques

Defense cost analysis, in its highest form, is an amalgam of scientific rigor and sound judgment. On the one hand, it requires knowledge, insight, and application of statistically-sound principles, and, on the other, critical interpretation of a wide variety of information that is often known with only limited precision. Indeed, Keynes’ observation on “the extreme precariousness of the basis of knowledge on which our estimates … have to be made” [6] often applies in defense cost analysis, especially for pre-Milestone (MS) B activities in the acquisition process and even more so for capability-based assessments in the requirements process. Since uncertainty and risk are always present in major defense acquisition programs and capability-based analyses, it’s essential to convey to senior leadership, in one fashion or another, the stochastic nature of the cost estimate. To do otherwise could lead to a false sense of security and a misallocation of resources.

Perhaps the ultimate expression of the randomness of a cost estimate is the S-curve, or cumulative probability distribution, employed frequently in both industry and government, often as a standard. Estimating these curves, *accurately and consistently* in a wide domain of applications, remains the Holy Grail in defense cost analysis. According to one school of thought, such distributions are “… rarely, if ever, known [within reasonable bounds of precision] ... for … investment projects.”[7] This contention remains an open issue within the international defense cost analysis community. Some practitioners concur, others don’t, and still others are unsure.

Amidst this spectrum of opinion, best-available techniques for conducting risk and uncertainty analysis of life-cycle cost estimates of defense acquisition programs include sensitivity analysis, Monte Carlo simulation, and eSBM.[8] Each technique, *if used properly*, can yield scientifically-sound results.
A best practice is to employ more than one technique and then compare findings. For example, detailed Monte Carlo simulation and eSBM both yield S-curves. Yet, the two techniques are fundamentally

[5] “Risk Management Guide for DoD Acquisition, Sixth Edition, August 2006; USD(AT&L), Systems and Software Engineering, Enterprise Development, page 33.
[6] The General Theory of Employment, Interest, and Money; Keynes, John Maynard; Harcourt Brace Jovanovich; 1964, page 149.
[7] Economic Theory and Operations Analysis, Baumol, William; Prentice-Hall; 1977, page 621.
[8] Interestingly, use of Monte Carlo simulation is more popular in the U.S. DoD than in the ministries of defense in other NATO countries where use of sensitivity analysis predominates.

different in approach, the former bottoms-up and the latter top-down. Divergence in results between the two procedures is a clarion call for explanation while consistency will inspire confidence in the validity of the estimates. Results of sensitivity analysis should be consistent with those from the other techniques in terms of impact on cost.

1.3 Cost Estimate Confidence: A WSARA Perspective

In May 2009, the US Congress passed the WSARA. This law aims to *improve the organization and procedures of the Department of Defense for the acquisition of weapon systems* [Public Law, 111-23]. WSARA addresses three areas: the organizational structure of the DOD, its acquisition policies, and its congressional reporting requirements. The following discussion offers a perspective on WSARA as it relates to reporting requirements for cost estimate confidence.

Public Law 111-23, Section 101 states the following:
The Director shall … issue guidance relating to the proper selection of confidence levels in cost estimates generally, and specifically, for the proper selection of confidence levels in cost estimates for major defense acquisition programs and major automated information system programs.

The Director of Cost Assessment and Program Evaluation, and the Secretary of the military department concerned or the head of the Defense Agency concerned (as applicable), shall each … disclose the confidence level used in establishing a cost estimate for a major defense acquisition program or major automated information system program, the rationale for selecting such confidence level, and, if such confidence level is less than 80 percent, justification for selecting a confidence level less than 80 percent.

What does cost estimate confidence mean? In general, it is a statement of the surety in an estimate along with a supporting rationale. The intent of WSARA's language suggests this statement is statistically derived; that is, expressing confidence as "*there is an 80 percent chance the program's cost will not exceed $250M*". How is cost estimate confidence measured?

Probability theory is the ideal formalism for deriving measures of confidence. With it, a program's cost can be treated as an uncertain quantity – one sensitive to many conditions and assumptions that change across its acquisition life cycle. Figure 1 illustrates the conceptual process for using probability theory to analyze cost uncertainty and producing confidence measures.

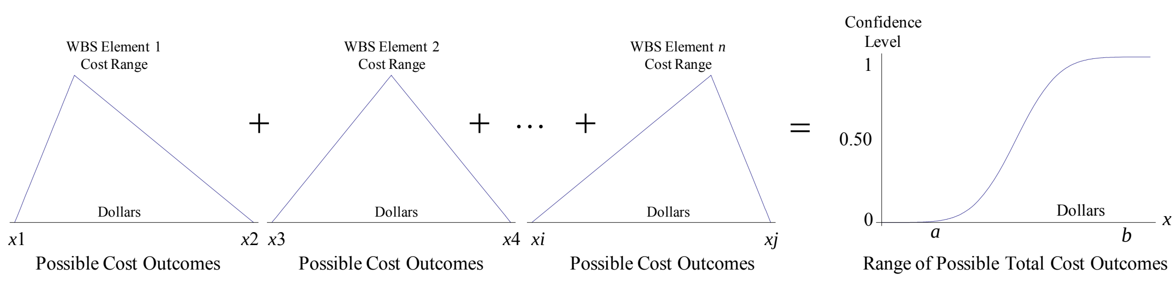

Figure 1. Cost Estimate Confidence: A Summation of Cost Element Cost Ranges

In Figure 1, the uncertainty in the cost of each work breakdown structure (WBS) element is expressed by a probability distribution. These distributions characterize each cost element's range of possible cost outcomes. All distributions are then combined by probability calculus to generate an overall distribution of program total cost. This distribution characterizes the range of total cost outcomes possible for the program. How does the output from this process enable confidence levels to be determined? Consider Figure 2.

Figure 2 illustrates the probability distribution of a program's total cost in cumulative form. It is another way to illustrate the output from a probability analysis of cost uncertainty, as described in Figure 1, specifically one that allows cost estimate confidence to be read from the distribution. For example, there is a 25% chance the program will cost less than or equal to \$100M, a 50% chance the program will cost less than or equal to \$151M, and an 80% chance the program will cost less than or equal to \$214M. These are confidence levels. The right side of Figure 2 shows the WSARA confidence level, as stated in Public Law 111-23, Section 101.

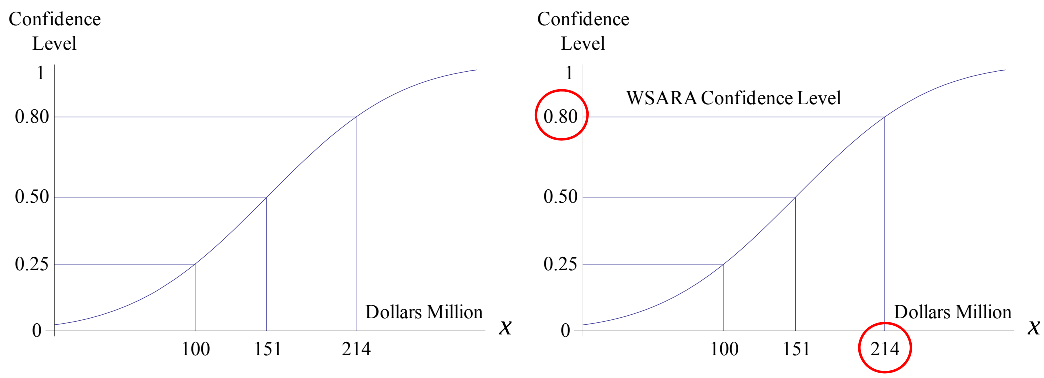

Figure 2. WSARA and Confidence Levels

A statistical technique known as Monte Carlo simulation is the most common approach for determining cost estimate confidence. This technique involves simulating the program cost impacts of all possible outcomes that might occur within a sample space of analyst-defined events. The output of a Monte Carlo simulation is a probability distribution of possible program costs. With this, analysts can present decision-makers a range of costs and a statistically derived measure of confidence the true or final program cost will remain in this range.

However, the soundness a Monte Carlo simulation is highly dependent on the mathematical skills and statistical training of the cost analysts conducting the analysis, traits that vary in the community. There are many subtleties in the underlying formalisms of Monte Carlo simulation, and these must be understood if errors in simulation design and in interpreting its outputs are to be avoided. For example, analysts must understand topics such as correlation and which of its many varieties is appropriate in cost uncertainty analysis. Analysts must understand that the sum of each cost element's most probable cost is not generally the most probable total program cost. In addition to understanding such subtleties, analysis must be skilled in explaining them to others.

SBM/eSBM, whose straightforward algebraic equations ease the mathematical burden on analysts, is an alternative to Monte Carlo simulation. . SBM/eSBM focuses on defining and analyzing risk scenarios as the basis for deliberations on the amount of cost reserve needed to protect a program from unwanted or unexpected cost increases. Such deliberations are a meaningful focus in cost reviews and in advancing cost realism. Defining, iterating, and converging on one or more risk scenarios is valuable for understanding elasticity in program costs, assessing cost estimate confidence, and identifying potential events a program must guard its costs against, if they occur. Scenarios build the necessary rationale for a traceable and defensible measure of cost risk. This discipline is often lacking in traditional Monte Carlo simulation approaches, where focus is often on its mathematical design instead of whether the design coherently models one or more scenarios of events that, if realized, drive costs higher than planned.

Regardless of the approach used, expressing cost estimate confidence by a range of possible cost outcomes has high information value to decision-makers. The breadth of the range itself is a measure of cost uncertainty, which varies across a program's life cycle. Identifying critical elements that drive a program's cost range offers opportunities for targeting risk mitigation actions early in its acquisition phases. Benefits of this analysis include the following three processes:

Establishing a Cost and Schedule Risk Baseline – Baseline probability distributions of program cost and schedule can be developed for a given system configuration, acquisition strategy, and cost-schedule estimation approach. The baseline provides decision-makers visibility into potentially high-payoff areas for risk reduction initiatives. Baseline distributions assist in determining a program's cost and schedule that simultaneously have a specified probability of not being exceeded. They can also provide decision-makers an assessment of the chance of achieving a budgeted (or proposed) cost and schedule, or cost for a given feasible schedule.

Determining Cost Reserve – Cost uncertainty analysis provides a basis for determining cost reserve as a function of the uncertainties specific to a program. The analysis provides the direct link between the amount of cost reserve to recommend and cost confidence levels. An analysis should be conducted to verify the recommended cost reserve covers fortuitous events (e.g., unplanned code growth, unplanned schedule delays) deemed possible by the engineering team.

Conducting Risk Reduction Tradeoff Analyses – Cost uncertainty analyses can be conducted to study the payoff of implementing risk reduction initiatives on lessening a program's cost and schedule risks. Furthermore, families of probability distribution functions can be generated to compare the cost and cost risk impacts of alternative requirements, schedule uncertainties, and competing system configurations or acquisition strategies.

The strength of any cost uncertainty analysis relies on the engineering and cost team's experience, judgment, and knowledge of the program's uncertainties. Documenting the team's insights into these uncertainties is *a critical part of the process*. Without it, credibility of the analysis is easily questioned and difficult to defend. Details about the analysis methodology, including assumptions, are components of the documentation. The methodology *must be technically sound* and offer value-added problem structure and insights otherwise not visible. Decisions that successfully reduce or eliminate uncertainty ultimately rest on human judgment. This at best is aided, not directed, by the methods discussed herein.

2.0 Scenario-Based Method (SBM)

The scenario-based method was developed along two implementation strategies, the non-statistical SBM and the statistical SBM, the latter of which is the form needed for WSARA. The following discussion describes each implementation and their mutual relationship.

2.1 Non-Statistical SBM

The scenario-based method is centered on articulating and costing a program's risk scenarios. Risk scenarios are coherent stories about potential events that, if they occur, increase program cost beyond what was planned.

The process of defining risk scenarios is a good practice. It builds the rationale and case- arguments to justify the reserve needed to protect program cost from the realization of unwanted events. This is lacking in Monte Carlo simulation if designed as arbitrary randomizations of possible program costs, a practice which can lead to reserve recommendations absent clear program context for what these funds are to protect.

Figure 3 illustrates the process flow of the non-statistical implementation of SBM.

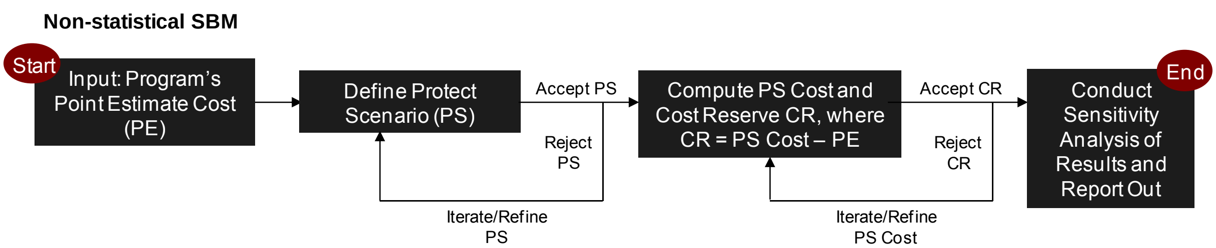

Figure 3. The Non-statistical SBM Process

The first step (Start) is input to the process. It is the program's point estimate (PE) cost. For purposes of this paper, the point estimate cost is the cost that does not include allowances for reserve. The PE cost is the sum of the cost-element costs across the program's work breakdown structure *without adjustments for uncertainty*. The PE cost is often developed from the program's cost analysis requirements description (CARD).

The next step in Figure 3 is defining a protect scenario (PS). A PS captures the cost impacts of major known risks to the program – those events the program must monitor and guard against occurring. The PS is not arbitrary, nor should it reflect extreme worst-case events. It should reflect a possible program cost that, in the judgment of the program, has *an acceptable chance of not being exceeded*. In practice, it is envisioned that management will converge on an "official" protect scenario after deliberations on the one initially defined. This part of the process ensures that all parties reach a consensus understanding of the program's risks and how they are best described by the protect scenario.

Once the protect scenario is established its cost is then estimated. The amount of cost reserve dollars (CR) needed to protect program cost can be computed as the difference between the PS cost and the PE cost. Shown in Figure 3, there may be additional refinements to the cost estimated for the protect scenario, based on management reviews and other considerations. The process may be iterated until the reasonableness of the magnitude of the cost reserve dollars is accepted by management.

The final step in Figure 3 is a sensitivity analysis to identify critical drivers associated with the protect scenario and the program's point estimate cost. It is recommended that the sensitivity of the amount of reserve dollars, computed in the preceding step, be assessed with respect to variations in the parameters associated with these drivers.

The non-statistical SBM, though simple in appearance, is a form of cost-risk analysis. The process of defining risk scenarios is a valuable exercise in identifying technical and cost estimation challenges inherent to the program. Without the need to define risk scenarios, cost risk analyses can be superficial, its case-basis not defined or carefully thought through. Scenario definition encourages a discourse on risks that otherwise might not be held, thereby allowing risks to become fully visible, traceable, and estimative to program managers and decision-makers.

The non-statistical SBM, in accordance with its non-statistical nature, does not produce confidence measures. The chance that the protect scenario cost, or of any other defined risk scenario's cost, will not be exceeded is not explicitly determined. The question is *Can this SBM implementation be modified to produce confidence measures while maintaining its simplicity and analytical features*? The answer is yes, and a way to approach this excursion is discussed next.

2.2 Statistical SBM

This section presents a statistical implementation of SBM. Instead of a Monte Carlo simulation, the statistical SBM is a closed-form analytic approach. It requires only a look-up table and a few algebraic equations.

Among the many reasons to implement a statistical track in SBM are the following: (1) it enables WSARA confidence measures to be determined, (2) it offers a way for management to examine changes in confidence measures as a function of how much reserve to buy to increase the chance of program success, and (3) it provides an ability to measure where the protect scenario cost falls on the probability distribution of the program's total cost.

Figure 4 illustrates the process flow of the statistical SBM. The upper part replicates the process steps of the non-statistical SBM, and the lower part appends the statistical SBM process steps. Thus, the statistical SBM is an augmentation of the non-statistical SBM.

To work the statistical SBM process, three inputs, as shown on the left in Figure 4, are required. These are the PE, the probability that the PE will not be exceeded, and the coefficient of variation, which will be explained below. The PE cost is the same as previously defined in the non-statistical SBM. The probability that PE cost x_{PE} will not be exceeded is the value α_{PE}, such that

$$P(Cost_{Pgm} \leq x_{PE}) = \alpha_{PE} \qquad (1)$$

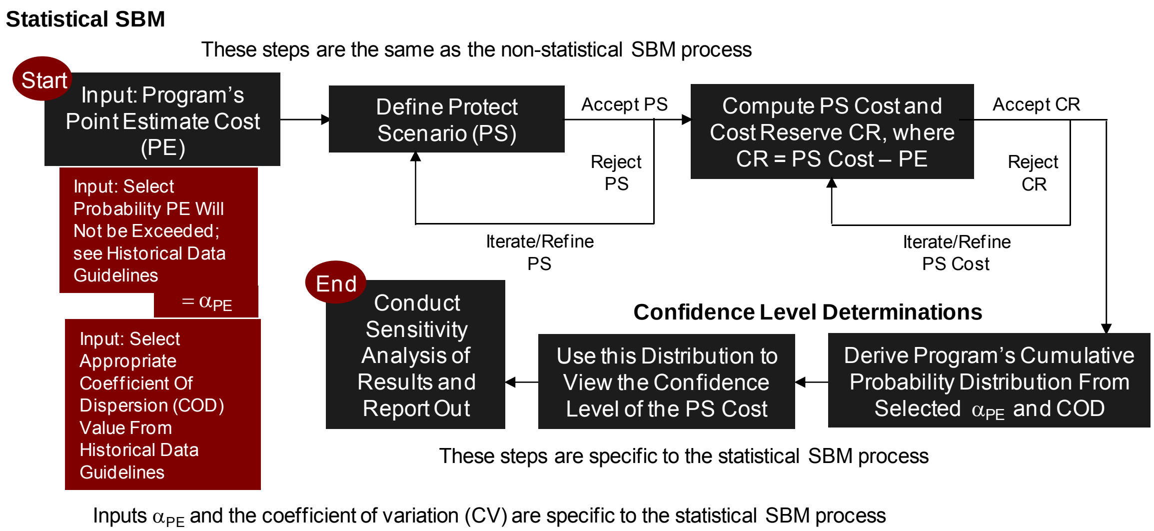

Figure 4. The Statistical SBM Process

In Equation 1, $Cost_{Pgm}$ is the true but uncertain total cost of the program and x_{PE} is the program's point estimate. The probability α_{PE} is a judged value guided by experience that it typically falls in the interval $0.10 \le \alpha_{PE} \le 0.50$. This interval reflects the understanding that a program's point estimate usually faces higher, not lower, probabilities of being exceeded.

The coefficient of variation (CV) is the ratio of a probability distribution's standard deviation to its mean. This ratio is given by Equation 2. The CV is a way to examine the variability of any distribution at plus or minus one standard deviation around its mean.

$$CV = D = \frac{\sigma}{\mu} \tag{2}$$

With values assessed for α_{PE} and CV, the program's cumulative cost probability distribution can then be derived. This distribution is used to view the confidence level associated with the PS cost, as well as confidence levels associated with any other cost outcome along this distribution.

The final step in Figure 4 is a sensitivity analysis. Here, we can examine the kinds of sensitivities previously described in the non-statistical SBM implementation, as well as uncertainties in values for α_{PE} and CV. This allows a broad assessment of confidence level variability, which includes determining a range of possible program cost outcomes for any specified confidence level.

Figure 5 illustrates an output from the statistical SBM process. The left picture is a normal probability distribution with point estimate PE equal to \$100M, α_{PE} set to 0.25, and CV set to 0.50. The range \$75M to \$226M is plus or minus one standard deviation around the mean of \$151M. From this, the WSARA confidence level and its associated cost can be derived. This is shown on the right in Figure 5.

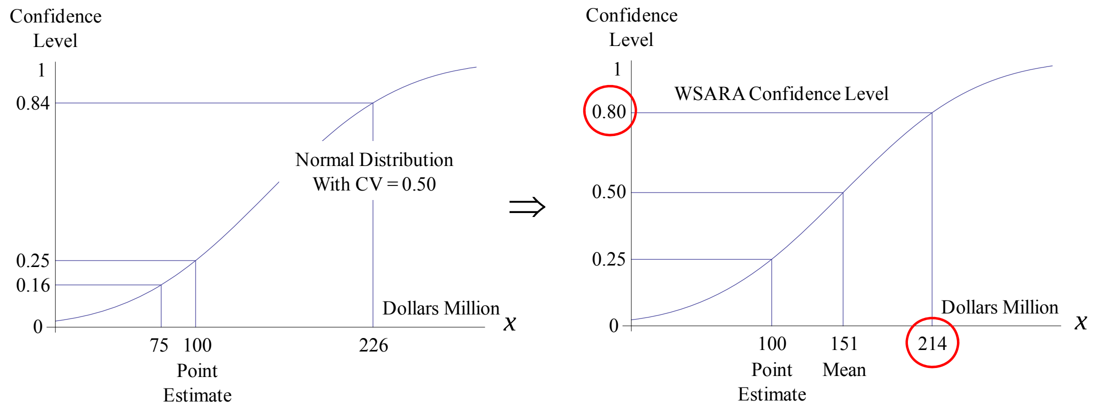

Figure 5. Statistical SBM Produces WSARA Confidence Levels

2.3 Statistical SBM Equations

This section presents the closed-form algebraic equations for the statistical SBM. Formulas to generate normal and lognormal probability distributions for program total cost are given.

Statistical SBM: An Assumed Underlying Normal for $Cost_{Pgm}$

The following equations derive from the assumption that $Cost_{Pgm}$ is normally distributed and the point (x_{PE}, α_{PE}) falls along this distribution. If we're given the point estimate cost PE, α_{PE}, and CV, then the mean and standard deviation of $Cost_{Pgm}$ are given by the following:

$$\mu_{Cost_{Pgm}} = x_{PE} - z_{PE}\frac{Dx_{PE}}{1 + Dz_{PE}} \tag{3}$$

$$\sigma_{Cost_{Pgm}} = \frac{Dx_{PE}}{1 + Dz_{PE}} \tag{4}$$

where D is the coefficient of variation (CV), x_{PE} is the program's point estimate cost, and z_{PE} is the value such that $P(Z \le z_{PE}) = \alpha_{PE}$ where Z is the standard (or unit) normal random variable. Values for z_{PE} are available in look-up tables for the standard normal, provided in Appendix A [Garvey, 2000] or by use of the built-in Excel function NORMSINV.

With the values computed from Equation 3 and Equation 4, the distribution function of $Cost_{Pgm}$ can be fully specified, along with the probability that $Cost_{Pgm}$ may take any particular outcome, such as the protect scenario cost. WSARA confidence levels such as the one in Figure 5 can then be determined.

Statistical SBM: An Assumed Underlying Lognormal for $Cost_{Pgm}$

The following equations derive from the assumption that $Cost_{Pgm}$ is lognormally distributed and the point (x_{PE}, α_{PE}) falls along this distribution. If we're given the point estimate cost PE, α_{PE}, and CV, then the mean and standard deviation of $Cost_{Pgm}$ are given by the following:

$$\mu_{\ln Cost_{Pgm}} = \ln x_{PE} - z_{PE}\sqrt{\ln(1+D^2)} \tag{5}$$

$$\sigma_{\ln Cost_{Pgm}} = \sqrt{\ln(1+D^2)} \tag{6}$$

where D is the coefficient of variation (CV), x_{PE} is the program's point estimate cost, and z_{PE} is the value such that $P(Z \le z_{PE}) = \alpha_{PE}$ where Z is the standard (or unit) normal random variable. Values for z_{PE} are available in Table B-1 in Appendix A.

However, values for $\mu_{\ln Cost_{Pgm}}$ and $\sigma_{\ln Cost_{Pgm}}$ are in log-dollar units. Equations 7 and 8 transform their values into dollar units.

$$\mu_{Cost_{Pgm}} = e^{\mu_{\ln Cost_{Pgm}} + \frac{1}{2}\sigma^2_{\ln Cost_{Pgm}}} \tag{7}$$

$$\sigma_{Cost_{Pgm}} = \sqrt{e^{2\mu_{\ln Cost_{Pgm}} + \sigma^2_{\ln Cost_{Pgm}}}\,(e^{\sigma^2_{\ln Cost_{Pgm}}} - 1)} \tag{8}$$

With the mean and standard deviation determined the distribution function of $Cost_{Pgm}$ can be fully specified, along with the probability that $Cost_{Pgm}$ may take any particular outcome such as the protect scenario cost. WSARA confidence levels such as the one in Figure 5 can be determined.

Example 1

Suppose the distribution function of $Cost_{Pgm}$ is normal. Suppose the program's point estimate cost is \$100M and this was assessed to fall at the 25th percentile. Suppose the type and life cycle phase of the program is such that 30 percent variability in cost around the mean has been historically seen. Suppose the program's protect scenario was defined and determined to cost \$145M.

a) Compute the mean and standard deviation of $Cost_{Pgm}$.
b) Plot the distribution function of $Cost_{Pgm}$.
c) Determine the confidence level of the protect scenario cost and its associated cost reserve.
d) Determine the program cost outcome associated with the WSARA confidence level.

Solution

a) From Equation 3 and Equation 4

$$\mu_{Cost_{Pgm}} = x_{PE} - z_{PE}\frac{Dx_{PE}}{1+Dz_{PE}} = 100 - z_{PE}\frac{(0.30)(100)}{1+(0.30)z_{PE}}$$

$$\sigma_{Cost_{Pgm}} = \frac{Dx_{PE}}{1+Dz_{PE}} = \frac{(0.30)(100)}{1+(0.30)z_{PE}}$$

We need z_{PE} to complete these computations. Since the distribution function of $Cost_{Pgm}$ is normal, it follows that $P(Cost_{Pgm} \le x_{PE}) = \alpha_{PE} = P(Z \le z_{PE})$, where Z is a standard normal random variable. Values for z_{PE} are available in Table B-1 in Appendix A. In this case, $P(Z \le z_{PE} = -0.6745) = 0.25$; therefore, with $z_{PE} = -0.6745$ we have

$$\mu_{Cost_{Pgm}} = x_{PE} - z_{PE}\frac{Dx_{PE}}{1+Dz_{PE}} = 100 - z_{PE}\frac{(0.30)(100)}{1+(0.30)z_{PE}} = 125.4 \text{ (\$M)}$$

$$\sigma_{Cost_{Pgm}} = \frac{Dx_{PE}}{1+Dz_{PE}} = \frac{(0.30)(100)}{1+(0.30)z_{PE}} = 37.6 \text{ (\$M)}$$

b) A plot of the probability distribution function of $Cost_{Pgm}$ is shown in Figure 6. This is a normal distribution with mean \$125.4M and standard deviation \$37.6M, as determined from a).

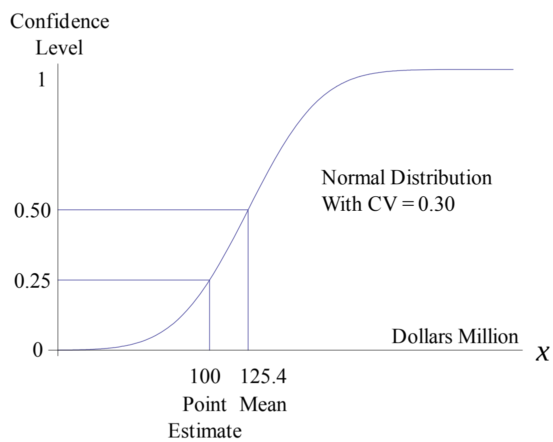

Figure 6. Probability Distribution Function of $Cost_{Pgm}$

c) To determine the confidence level of the protect scenario, find α_{PS} such that

$$P(Cost_{Pgm} \le x_{PS} = 145) = \alpha_{PS}$$

Finding α_{PS} is equivalent to solving for z_{PS} the expression

$$\mu_{Cost_{Pgm}} + z_{PS}(\sigma_{Cost_{Pgm}}) = x_{PS}$$

From this

$$z_{PS} = \frac{x_{PS} - \mu_{Cost_{Pgm}}}{\sigma_{Cost_{Pgm}}} = \frac{x_{PS}}{\sigma_{Cost_{Pgm}}} - \frac{1}{D}$$

Since $x_{PS} = 145$, $\mu_{Cost_{Pgm}} = 125.4$, and $\sigma_{Cost_{Pgm}} = 37.6$ it follows that

$$z_{PS} = \frac{x_{PS} - \mu_{Cost_{Pgm}}}{\sigma_{Cost_{Pgm}}} = \frac{x_{PS}}{\sigma_{Cost_{Pgm}}} - \frac{1}{D} = \frac{145}{37.6} - \frac{1}{(0.30)} = 0.523$$

From Table B-1, in Appendix A, $P(Z \le z_{PS} = 0.523) \approx 0.70$. Therefore, the \$145M protect scenario cost falls at the 70th percentile of the distribution. This implies a cost reserve CR equal to \$45M.
d) To determine the WSARA confidence level cost, from Table B-1 in Appendix A

$$P(Z \le z_{0.80} = 0.8416) = 0.80$$

From part c), we can write the expression

$$\mu_{Cost_{Pgm}} + z_{0.80}(\sigma_{Cost_{Pgm}}) = x_{0.80}$$

Substituting $\mu_{Cost_{Pgm}} = 125.4$ and $\sigma_{Cost_{Pgm}} = 37.6$ (determined in part a) yields the following:

$$\mu_{Cost_{Pgm}} + z_{0.80}(\sigma_{Cost_{Pgm}}) = 125.4 + 0.8416(37.6) = 157 = x_{0.80}$$

Therefore, the cost associated with the WSARA confidence level is \$157M. Figure 7 presents a summary of the results in this example.

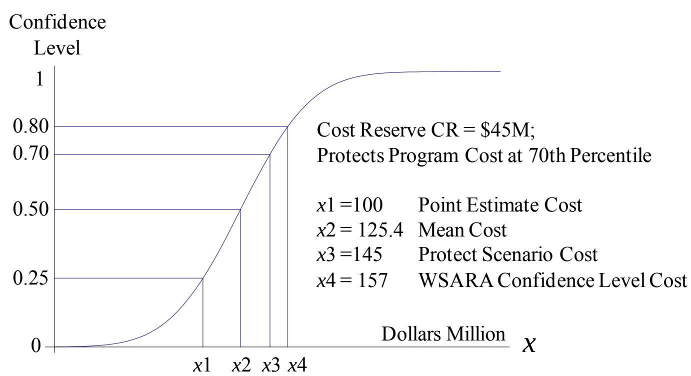

Figure 7. Example 1: Resultant Distribution Functions and Confidence Levels

Example 2
Suppose the distribution function of $Cost_{Pgm}$ is **lognormal**. Suppose the program's point estimate cost is \$100M and this was assessed to fall at the 25th percentile. Suppose the type and life cycle phase of the program is such that 30 percent variability in cost around the mean has been historically seen. Suppose the program's protect scenario was defined and determined to cost \$145M.

a) Compute $\mu_{Cost_{Pgm}}$ and $\sigma_{Cost_{Pgm}}$.

b) Determine the confidence level of the protect scenario cost and its associated cost reserve.

Solution

a) From Equations 5 and 6, and Example 1, it follows that

$$\mu_{\ln Cost_{Pgm}} = \ln x_{PE} - z_{PE}\sqrt{\ln(1+D^2)} = \ln(100) - (-0.6745)\sqrt{\ln(1+(0.30)^2)} = 4.80317$$

$$\sigma_{\ln Cost_{Pgm}} = \sqrt{\ln(1+D^2)} = \sqrt{\ln(1+(0.30)^2)} = 0.29356$$

From Equations 7 and 8 we translate the above mean and standard deviation into dollar units.

$$\mu_{Cost_{Pgm}} = e^{\mu_{\ln Cost_{Pgm}} + \frac{1}{2}\sigma^2_{\ln Cost_{Pgm}}} = e^{4.80317 + \frac{1}{2}(0.29356)^2} \approx 127.3 \ (\$M)$$

$$\sigma_{Cost_{Pgm}} = \sqrt{e^{2\mu_{\ln Cost_{Pgm}} + \sigma^2_{\ln Cost_{Pgm}}}(e^{\sigma^2_{\ln Cost_{Pgm}}} - 1)}$$

$$= \sqrt{e^{2(4.80317)+(0.29356)^2}(e^{(0.29356)^2} - 1)} \approx 38.2 \ (\$M)$$

b) To determine the confidence level of the protect scenario we need to find $\alpha_{x_{PS}}$ such that

$$P(Cost_{Pgm} \le x_{PS} = 145) = \alpha_{x_{PS}}$$

Finding $\alpha_{x_{PS}}$ is equivalent to solving

$$\mu_{\ln Cost_{Pgm}} + z_{x_{PS}}(\sigma_{\ln Cost_{Pgm}}) = \ln x_{PS}$$

for $z_{x_{PS}}$. From the above, we can write the expression

$$z_{x_{PS}} = \frac{\ln x_{PS} - \mu_{\ln Cost_{Pgm}}}{\sigma_{\ln Cost_{Pgm}}}$$

Since $x_{PS} = 145$, $\mu_{\ln Cost_{Pgm}} = 4.80317$, and $\sigma_{\ln Cost_{Pgm}} = 0.29356$ it follows that

$$z_{x_{PS}} = \frac{\ln x_{PS} - \mu_{\ln Cost_{Pgm}}}{\sigma_{\ln Cost_{Pgm}}} = \frac{\ln 145 - 4.80317}{0.29356} = 0.59123$$

From the look-up table in Appendix A we see that

$$P(Z \le z_{x_{PS}} = 0.59123) \approx 0.723$$

Therefore, the protect scenario cost of 145 ($M) falls at approximately the 72nd percentile of the distribution with a cost reserve (CR) of 45 ($M).

2.4 Measuring Confidence in WSARA Confidence

This section illustrates how SBM can examine the sensitivity in program cost at the 80th percentile to produce a measure of cost risk in the WSARA confidence level. Developing this measure carries benefits similar to doing so for a point cost estimate, except it is formed at the 80th percentile cost. Furthermore, a measure of cost risk can be developed at any confidence level along a probability distribution of program cost. The following uses Example 1 to illustrate these ideas.

In Example 1, single values for α_{PE} and CV were used. If a range of possible values is used then a range of possible program costs can be generated at any percentile along the distribution. For instance, suppose historical cost data for a particular program indicates its CV varies in the interval $0.20 \le CV \le 0.50$. Given the conditions in Example 1, variability in CV affects the mean and standard deviation of program cost. This is illustrated in Table 1, given a program's point estimate cost equal to $100M and its $\alpha_{PE} = 0.25$.

Coefficient of Variation (CV)	Standard Deviation ($M)	Mean ($M) 50th Percentile*	WSARA Confidence Level ($M) 80th Percentile
0.20	23.1	115	135
0.30	37.6	125	157
0.40	54.8	137	183
0.50	75.4	151	214

Table 1. Ranges of Cost Outcomes in Confidence Levels (Rounded)

**In a normal distribution, the mean is also the median (50th percentile)*

Table 1 shows a range of possible cost outcomes for the 50th and 80th percentiles. Selecting a particular outcome can be guided by the CV considered most representative of the program's uncertainty at its specific life cycle phase. This is guided by the scenario or scenarios developed at the start of the SBM process. Figure 1 graphically illustrates the results in Table 1.

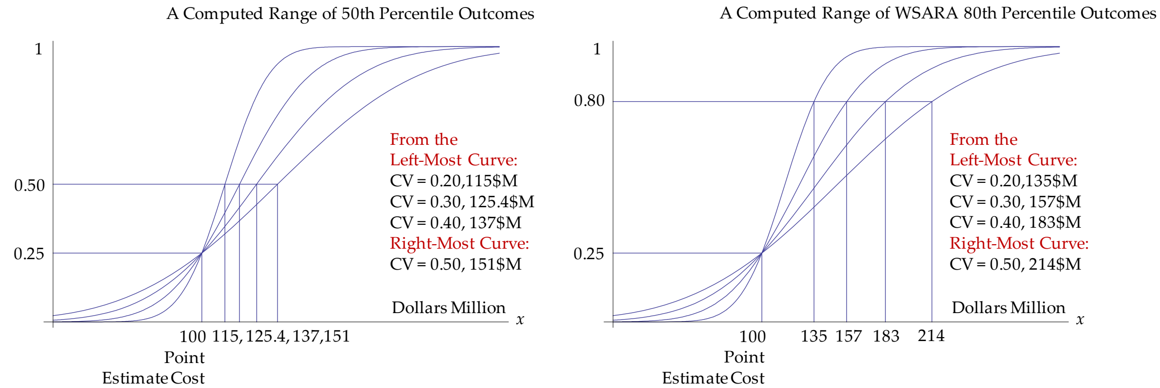

Figure 8. A Range of Confidence Level Cost Outcomes

Finally, one can use SBM outputs to generate a probability distribution of cost outcomes associated with any confidence level. In Figure 8, suppose we want a confidence level for each cost outcome in the WSARA range. To do this, we fit a distribution to the interval [135, 157, 183, and 214]. Suppose we hypothesize that values in this interval follow a lognormal distribution. The Kolmogorov-Smirnov (K-S) test [Garvey, 2000] can be used to accept or reject this hypothesis. When the K-S test was applied to these data, it indicated accepting the hypothesis.

Acceptance does not mean the lognormal is the unique distribution. It only means the lognormal is a statistically plausible distribution for the data in the WSARA interval [135, 157, 183, and 214]. Figure 9 shows the lognormal that best fits the WSARA interval. Confidence levels associated with each value in this interval are shown along its vertical axis. The 80th percentile cost outcome of $183M has a confidence level equal to 0.65. Thus, there is a 65 percent chance the 80th percentile cost will not be exceeded. *This statement is an expression of cost risk in the confidence of the 80th percentile cost outcome*. Confidence levels associated with the other cost outcomes, in the WSARA interval, are also shown in Figure 9. Expressions of cost risk associated with these outcomes can likewise be stated.

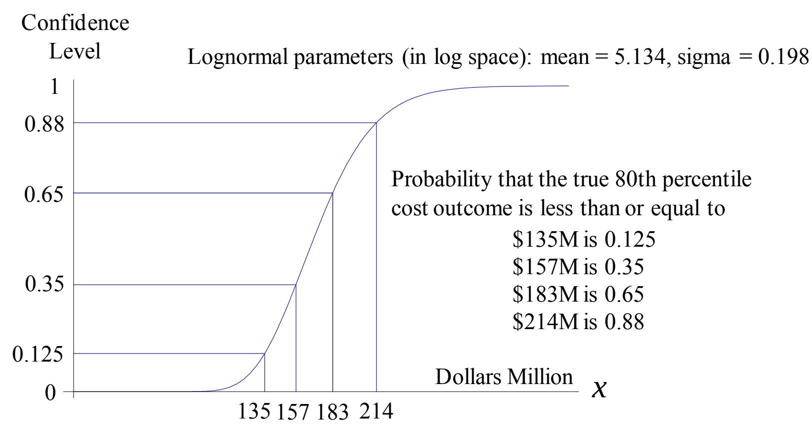

Figure 9. Measuring Confidence in
WSARA Confidence Levels: A Lognormal Statistical Fit

3.0 The Enhanced SBM (eSBM)

As mentioned earlier, the scenario-based method was introduced in 2006 as an alternative to Monte Carlo simulation for generating a range of possible program cost outcomes and associated confidence measures. This section presents the enhanced scenario-based method (eSBM), an historical- data-driven application of the statistical SBM, with heightened analytical features.

Two key inputs characterize the statistical SBM. They are (1) the probability that a program's point estimate cost will not be exceeded (α_{PE}) and (2) the coefficient of variation (CV). With these, risk analyses and confidence measures are easily produced. eSBM operates with these same inputs, while featuring additional ways to assess α_{PE} and CV.

Approaches for Assessing α_{PE}

Discussed earlier, the probability a program's point estimate PE cost will not be exceeded is the value α_{PE} such that $P(Cost_{Pgm} \le x_{PE}) = \alpha_{PE}$. Historical data on α_{PE} is poor. However, it is anecdotally well understood a program's PE usually faces higher, not lower, probabilities of being exceeded. The interval $0.10 \le \alpha_{PE} \le 0.50$ expresses this experience. It implies a program's PE will very probably experience growth instead of reduction. Unless there are special circumstances, a value for α_{PE} from this interval should be selected for eSBM and a justification written for the choice. A sensitivity analysis on other possible α_{PE} values should be conducted and the results documented.

Another approach for assessing α_{PE} is to compute its value from two other probabilities; specifically, α_1 and α_2 shown in Figure 10.

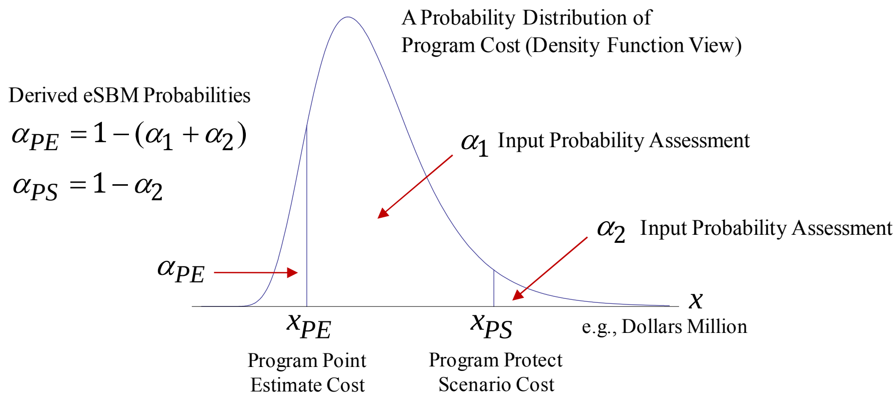

Figure 10. Determining eSBM Probabilities α_{PE} and α_{PS}

In Figure 10, probabilities α_1 and α_2 relate to x_{PE} and x_{PS} as follows:

$$\alpha_1 = P(x_{PE} \le Cost_{Pgm} \le x_{PS})$$

$$\alpha_2 = P(Cost_{Pgm} \ge x_{PS})$$

Values for α_1 and α_2 are judgmental. When they are assessed, probabilities α_{PE} and α_{PS} derive from Equation 9 and Equation 10, respectively.

$$\alpha_{PE} = P(Cost_{Pgm} \leq x_{PE}) = 1 - (\alpha_1 + \alpha_2) \quad (9)$$

$$\alpha_{PS} = P(Cost_{Pgm} \leq x_{PS}) = 1 - \alpha_2 \quad (10)$$

Given α_{PE} and α_{PS}, a normal or lognormal distribution for $Cost_{Pgm}$ can be fully specified. From either distribution, possible program cost outcomes at any confidence level (e.g., WSARA) can be determined.

Example 3

Suppose the distribution function of $Cost_{Pgm}$ is lognormal with $x_{PE} = \$100\text{M}$ and $x_{PS} = \$155\text{M}$. In Figure 10, if $\alpha_1 = 0.70$ and $\alpha_2 = 0.05$ then answer the following:

a) Derive probabilities α_{PE} and α_{PS}.
b) Determine the program cost outcome associated with the WSARA confidence level.

Solution

a) From Equations 9 and 10

$$\alpha_{PE} = P(Cost_{Pgm} \leq x_{PE}) = 1 - (\alpha_1 + \alpha_2) = 1 - (0.70 + 0.05) = 0.25$$
$$\alpha_{PS} = P(Cost_{Pgm} \leq x_{PS}) = 1 - \alpha_2 = 1 - 0.05 = 0.95$$

b) The probability distribution of $Cost_{Pgm}$ is given to be lognormal. From the properties of a lognormal distribution (Appendix B)

$$P(Cost_{Pgm} \leq x_{PE}) = P(Z \leq z_{PE} = \frac{\ln x_{PE} - \mu_{\ln Cost_{Pgm}}}{\sigma_{\ln Cost_{Pgm}}}) = \alpha_{PE}$$

$$P(Cost_{Pgm} \leq x_{PS}) = P(Z \leq z_{PS} = \frac{\ln x_{PS} - \mu_{\ln Cost_{Pgm}}}{\sigma_{\ln Cost_{Pgm}}}) = \alpha_{PS}$$

This implies

$$\mu_{\ln Cost_{Pgm}} + z_{PE}(\sigma_{\ln Cost_{Pgm}}) = \ln x_{PE}$$

$$\mu_{\ln Cost_{Pgm}} + z_{PS}(\sigma_{\ln Cost_{Pgm}}) = \ln x_{PS}$$

Since Z is a standard normal random variable, from Table B-1 in Appendix A

$$P(Z \leq z_{PE}) = \alpha_{PE} = 0.25 \text{ when } z_{PE} = -0.6745$$

and

$$P(Z \leq z_{PS}) = \alpha_{PS} = 0.95 \text{ when } z_{PE} = 1.645$$

Given $x_{PE} = \$100\text{M}$ and $x_{PS} = \$155\text{M}$ it follows that

$$\mu_{\ln Cost_{Pgm}} + (-0.6745)(\sigma_{\ln Cost_{Pgm}}) = \ln 100$$

$$\mu_{\ln Cost_{Pgm}} + (1.645)(\sigma_{\ln Cost_{Pgm}}) = \ln 155$$

Solving these equations yields $\mu_{\ln Cost_{Pgm}} = 4.73262$ and $\sigma_{\ln Cost_{Pgm}} = 0.188956$, which are in log-dollar units. Equations 7 and 8 transform their values into dollar units. The result is $\mu_{Cost_{Pgm}} = \$115.64\text{M}$ and $\sigma_{Cost_{Pgm}} = \$22.05\text{M}$.

To find the WSARA confidence level, from Example 1 recall that $P(Z \le z_{0.80} = 0.8416) = 0.80$. Since the distribution function of $Cost_{Pgm}$ is lognormal

$$\mu_{\ln Cost_{Pgm}} + (0.8416)(\sigma_{\ln Cost_{Pgm}}) = \ln x_{0.80}$$

In this case

$$4.73262 + (0.8416)(0.188956) = 4.89165 = \ln x_{0.80}$$

Thus, the program cost associated with the WSARA confidence level is

$$e^{4.89165} = x_{0.80} = \$133.2\text{M}$$

Figure 11 summarizes these results and illustrates other interesting percentiles. In this case, the WSARA confidence level cost is less than the protect scenario's confidence level cost. This highlights the importance of comparing these cost outcomes, their confidence levels, and the drivers behind their differences. Example 3 demonstrates a program's protect scenario cost is not guaranteed to be less than its WSARA confidence level cost.

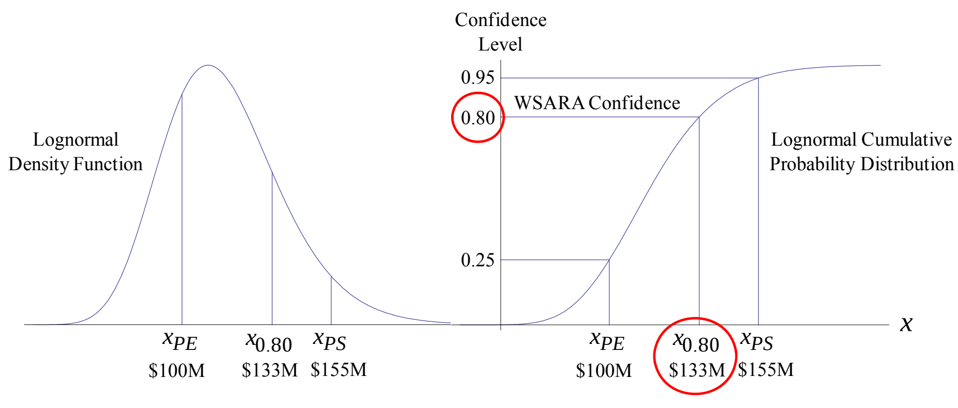

Figure 11. Example 3: Resultant Distribution Functions and Confidence Levels

4.0 Development of Benchmark Coefficient of Variation Measures

To shed light on the behavior of cost distribution functions (S-curves) employed in defense cost risk analyses and to develop historical performance benchmarks, five conjectures on coefficient of variation behavior (CV) are proffered:

- Consistency
  - CVs in current cost estimates are consistent with those computed from acquisition histories;
- Tendency to Decline During Acquisition Phase
  - CVs decrease throughout the acquisition lifecycle;
- Platform Homogeneity
  - CVs are equivalent for aircraft, ships, and other platform types;
- Tendency to Decrease after Normalization
  - CVs decrease when adjusted for changes in quantity and inflation; and
- Invariance of Secular Trend
  - CVs are steady long-term.

Assessment of the correctness of each of the above conjectures, through a data collection and analysis effort, follows.

The first conjecture, consistency, posits that CVs commonly estimated today in the defense cost-analysis community are consistent with values computed from the distribution of historical results on completed or nearly-completed weapon-system acquisition programs. Note that consistency does not necessarily mean accuracy. Accuracy is more problematic and requires evaluation of the *pedigree* of cost baselines upon which historical acquisition outcomes were computed. An additional issue is the degree to which historical results are applicable to today's programs and their CVs because of the possibility of structural change due to WSARA and other recent Office of the Secretary of Defense (OSD) acquisition initiatives.

The second conjecture, tendency to decline during the acquisition phase, suggests that CVs should decrease monotonically throughout the acquisition lifecycle as more information is acquired regarding the program in question. We certainly will know more about a system's technical and performance characteristics at MS C than we do at MS A

Regarding the third conjecture, platform homogeneity, there's no reason to believe, *a priori*, that CVs should differ by platform. All programs fall under basically the same acquisition management processes and policies. Further, tools and talent in the defense cost and acquisition-management communities are likely distributed uniformly, even though each of us thinks we have the best people and methods.

The fourth conjecture, tendency to decrease when data are normalized, suggests, logically, that CVs should decrease as components of variation in costs are eliminated.

And finally, the fifth conjecture, secular-trend invariance, hypothesizes that CVs have not changed (and therefore will not change) significantly over the long run.

4.1 Historical Cost Data

The degree to which these conjectures hold was examined through a data-collection and analysis effort based on 100 Selected Acquisition Reports (SARs) that contain raw data on cost outcomes of mostly historical Department of the Navy (DON) major defense acquisition programs (MDAPs) but also a handful of on-going programs where cost growth has likely stabilized, such as LPD-17. As enumerable studies elsewhere have indicated, the SARs, while not perfect, are nevertheless a good, convenient, comprehensive, official source of data on cost, schedule, and technical performance of MDAPs. More importantly, they're tied to milestones, as are independent cost estimates (ICEs), and they present total program acquisition costs across multiple appropriations categories and cycles. For convenience, data were culled from SAR Summary Sheets, which present top-level numerical cost data.[9]

For a given program, the SAR provides two estimates of cost. The first is a baseline estimate (BE), usually made when the system nears a major milestone. The second is the current estimate (CE), which is based on best-available information and includes all known and anticipated revisions and changes to the program. For completed acquisitions, the CE in the last SAR reported is regarded as the actual cost of the program.

SAR costs are reported in both base-year and then-year dollars, allowing for comparisons both with and without the effects of inflation.

The ratio of the CE to the BE is a cost growth factor, (CGF), reported as a metric in most SAR-based cost-growth studies. Computation of CGFs for large samples of completed programs serves as the basis upon which to estimate the standard deviation and the mean of acquisition cost outcomes, and hence the CV. An outcome, as measured by the CGF, is a percent deviation, in index form, from an expected value or the BE. For current acquisition programs, the BE is *supposed* to reflect the costs of an Acquisition Program Baseline (APB) and to be "consistent with" an ICE.[10]

In practice, for modern-era programs, there's very strong evidence to support the hypothesis that the SAR BE is, in fact, a cost estimate. Based on an analysis of 10 programs in our database dating from the 1990s, there is little difference between the SAR BE, the program office estimate (POE) of acquisition costs, and the ICE conducted either by the Naval Center for Cost Analysis (NCCA) or OSD.[11] The outstanding fact is rather the degree of conformity of the values, with the POEs averaging 2% less and the ICEs 3% more than the SAR BE in then-year dollars.

---

[9] SAR Summary Sheets are produced annually by the Office of the Undersecretary of Defense (Acquisition, Technology and Logistics); Acquisition Resources and Analysis.

[10] "APBs as we know them today did not start until 1988, and were not incorporated into the Consolidated Acquisition Reporting System (CARS), [now replaced by the] Defense Acquisition Management Information Retrieval DAMIR [system] until 1990;" e-mail from the late Ms. Chris Knoche, USD(AT&L); 25 Feb 2011. Historically (1970s and 1980s), DoD's SAR Instruction 7000.3 indicated that "A Secretary of Defense Decision Memorandum will normally be the source of … cost estimates" in the SAR.

[11] Special thanks are due to Mr. John McCrillis of NCCA who enabled the comparison by assembling files of OSD ICE memos and program-office cost estimates for many ACAT IC and ID programs in the 1990s and 2000s. There was an intersection of 10 acquisition programs between Mr. McCrillis' files and the ones in this study in terms of compatibility of the estimates, i.e., same program, same milestone, same quantities, and inclusion of required cost data.

Unfortunately, ICE memos and program-office estimates from the 1970s and 1980s are generally unavailable. SARs in that era were supposed to reflect cost estimates in a SECDEF Decision Memorandum, an output of the Defense System Acquisition Review Council (DSARC), predecessor of today's Defense Acquisition Board. Degree of compliance with this guidance is unknown to us.

Prospective changes in acquisition quantity from a program baseline are generally regarded as beyond the purview of the cost analyst in terms of generating S-curves.[12] There are several ways for adjusting raw then-year or base-year dollars in the SARs to reflect the changes in quantity that did occur, including but not limited to the ones shown here below. The estimated cost change corresponding to the quantity change is denoted QΔE.

- Adjust baseline estimate to reflect current quantities
    - CGF = CE/(BE + QΔE)
    - Used in SARs
- Adjust current estimate to reflect baseline quantities
    - CGF = (CE – QΔE)/BE
- "Fischer" index = Square root of the product of the first two.

The first two formulae are analogous to the Paasche and Laspeyres price indices, which are based on current and base year quantities, respectively. The third we dub "Fischer's" index which, in the context of price indices, is the square root of the product of the other two. The Fischer index, used to compute the GDP Price Index but not previously employed in SAR cost-growth studies, takes into consideration the reality that changes in quantity are typically implemented *between* the base year and current year rather than at either extreme. In any event, the deltas in CVs are typically negligible no matter which method of adjustment is used.[13]

4.2 Sample Data at MS B

Of the 100 programs in the sample, 50 were MS B estimates of total program acquisition cost (development, production, and, less frequently, military construction). Platform types included aircraft, helicopters, missiles, ships and submarines, torpedoes, and a few other systems. From the SAR summary sheets, these data elements were captured: base year, baseline type, platform type, baseline and current cost and quantity estimates, changes to date, date of last SAR, and with all costs in both base-year and then-year dollars. Results were analyzed, and the means, standard deviations, and CVs are displayed in Table 4.

[12] Performing what-if drills for alternative development and production quantities and schedules, however, is a legitimate and necessary undertaking.

[13] The high-to-low spread in CVs computed using the three methods of quantity adjustment for a sample size of 50 ship and submarine acquisition programs at MS B is only 0.02 and 0.04 in base-year and then-year dollars, respectively.

Cost Growth Factors and CVs for DON MDAPs at MS B

	Without Quantity Adjustment		Quantity Adjusted	
Statistics	Base-Year$	Then-Year$	Base-Year$	Then-Year$
Mean	1.48	1.84	1.23	1.36
Standard Deviation	0.94	1.60	0.44	0.69
CV	0.63	0.87	0.36	0.51

Table 4

Four CVs were tallied, corresponding to the four types of CGFs estimated. As adjustments for quantity and inflation were made, the CVs decreased, as expected.

Figure 12 shows CGFs adjusted for changes in quantity but not inflation.[14] The histogram's skewness suggests a lognormal distribution, with the mean falling to the right of the median. As has been noted in the statistical literature, CVs, as they are computed in the cost community using traditional product-moment formulae, are subject to the influence of outliers. The CV numerator, after all, is the sum of *squared* differences of observations from the mean. That's certainly the case here because of Harpoon, the right-most datum, with a CGF of 3.96, indicating almost 300% cost growth. Eliminating this observation from the sample decreases the CV from 51% to 45%.

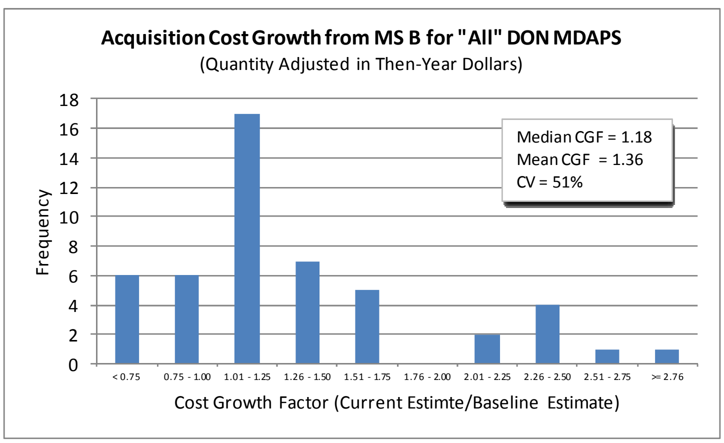

Figure 12. MS B CGFs

CVs were then analyzed by type of platform, with results illustrated in Figure 13 first for the entire data set and then separately for ships and submarines, aircraft, missiles, and electronics. The missiles group is

[14] We believe that CVs and S-curves should be estimated and employed with quantity regarded as exogenous but inflation as random. Quantity changes are typically the results of changes in requirements and departmental or congressional funding decisions. Their impact on cost, we think, is handled best through what-if drills. For the second term, inflation, more and more acquisition programs are using non-OSD rates in generating then-year dollar cost estimates. The treatment of outyear inflation rates as a stochastic variable therefore seems appropriate. We recognize that cost-analysis organizations may logically proffer different guidance on this issue. It's essential, in any event, to make crystal clear the type of CV employed.

heavily influenced by the aforementioned Harpoon outlier; eliminating it drops the quantity-adjusted then-year dollar CV for that group to 47%, remarkably close to the values for the other types of platforms.

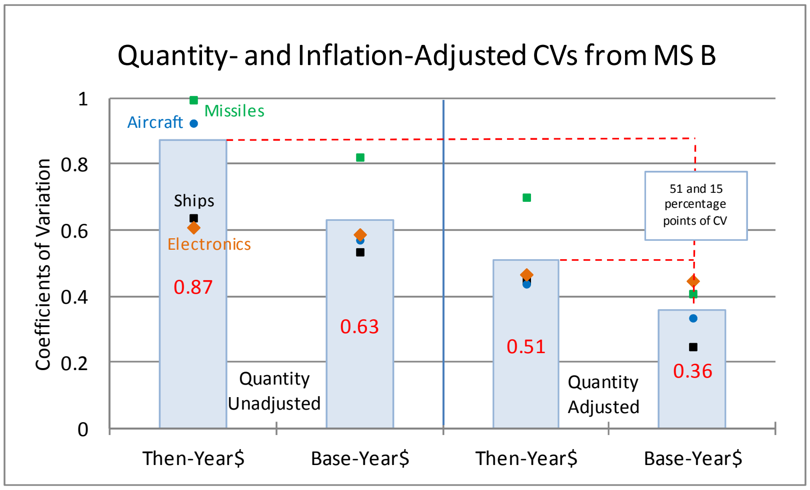

Figure 13. MS B CVs

To shed light on the homogeneity of CVs, the null hypotheses of equal population means for platform type was formulated versus the alternative of at least one pairwise difference. [15]

- H_o: $\mu_1 = \mu_2 = \ldots = \mu_k$, where μ_i is a platform population mean CGF
- H_a: $\mu_i \neq \mu_j$, for at least one (i,j) pair.

The appropriate test statistic in this case is the F, or the ratio of between sample variance to within sample variance, with sample data shown in Figure 14.

[15] To our knowledge, a test for the equality of *k* coefficients of variation from lognormal distributions in small samples has not been developed. Hence, we examined the behavior of the two components of a CV separately, the mean and standard deviation. In the case of normal distributions, see "Confidence Bounds and Hypothesis Tests for Normal Distribution Coefficients of Variation," Verrill and Johnson, U.S. Department of Agriculture, 2007. In the *hypothetical* case of a *normal* distribution for MS B data, the Verrill and Johnson sample-sample procedure does not reject the null hypothesis of equal CVs, at the 10% level of significance.

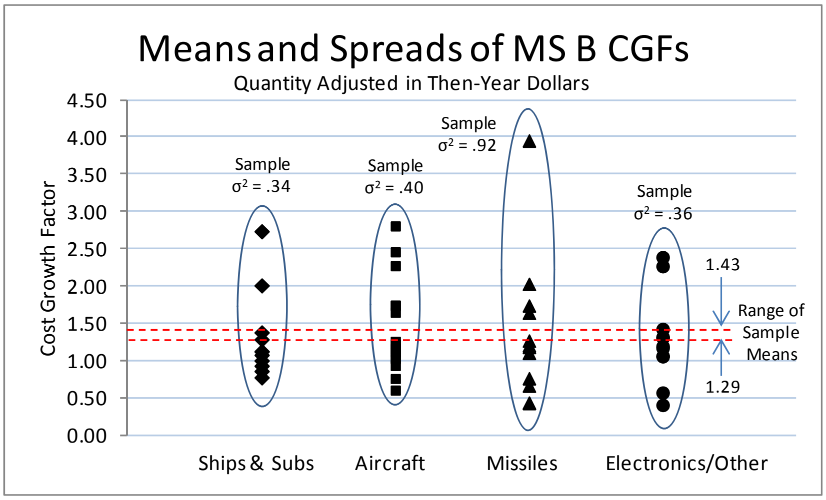

Figure 14. Means and Spreads of MS B CGFs

Intuitively, a high ratio of between sample variance to within sample variance, for different platform types, is suggestive of different population means. The low value of the computed test statistic [$F_{(3,45)}$ = 0.12] suggests insignificance; the data, in other words, provide no evidence that the population means are different.

Similar hypotheses were formulated for the other component of CVs, platform variances.

- H_o: $\sigma^2_1 = \sigma^2_2 = \ldots = \sigma^2_k$, where σ^2_i is a platform population variance
- H_a: $\sigma^2_i \neq \sigma^2_j$, for at least one (i,j) pair.

Two statistical tests were employed, pairwise comparisons and Levene's test for *k* samples for skewed distributions, with the null hypothesis, in all cases, not rejected at the 5% level of significance.[16]

The combination of statistical evidence for the dual hypotheses of homogeneous means and variances, therefore, strongly supports the conjecture of homogeneous CVs, quantity-adjusted in then-year dollars, at MS B.

4.3 Additional Findings at MS B

As Figure 14 shows, CVs do in fact decrease significantly as components of the variation in costs are explained. The data set of 50 observations, it's important to note, contains two programs with BEs in the late 1960s and more for the 1970s. Notice the adjustments for inflation. The total delta in CVs from unadjusted in then-year dollars to quantity-adjusted in base-year dollars is 51 percentage points. Of this amount, *after* adjusting for changes in quantity, inflation represents a full *15 percentage points*.

That's a significant contribution. Perhaps it's due to the volatility in average annual rates during the Nixon/Ford (6.5%), Carter (10.7%), Reagan (4.0%), G.H.W. Bush (3.9%), and Clinton (2.7%)

[16] Levene, Howard (1960). "Robust Tests for Equality of Variances," Ingram Olkin, Harold Hotelling, et alia. Stanford University Press. pp. 278–292. Details of the test results are available from the authors.

administrations.[17] During the mid 1970s, OSD Comptroller (Plans and Systems) was promulgating inflation forecasts of 3 to 4% per annum – received of course from the Office of Management and Budget (OMB), but with inflation in the general economy rising to over 10% per annum during the peak inflation period of 1978 to 1981.

That disconnect caused tremendous churn in defense acquisition programs. No one in the early or even mid 1970s was predicting double digit inflation and interest rates. For the most part, defense acquisition programs used OSD rates in estimating then-year dollar total obligational authority. Double-digit inflation reality simply did not jibe with values that had been used years previously to create the defense budget.[18] To complicate matters, OMB eventually recognized their rates were too low and began promulgating higher rates only to see inflation fall significantly in the early 1980s. The existence and size of a DoD inflation "dividend," resulting from prescribed rates *exceeding* actual values, was hotly debated but could have caused additional perturbations.

Turning now to the conjecture of constant CVs over lengthy periods, Figure 15 shows a pronounced decline in values.

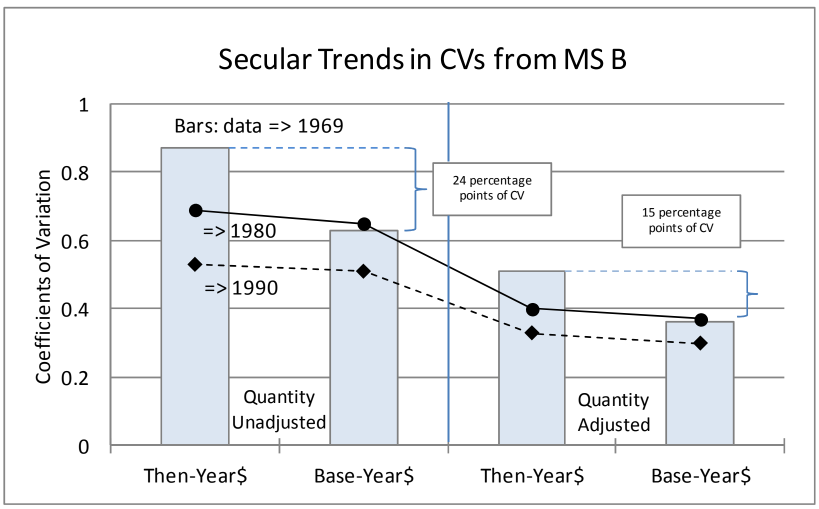

Figure 15. Secular Trend

Inflation had much less impact on the magnitude of CVs in the 1980s and 1990s than in the 1970s, likely due to less volatility in rates and a secular decline in their values. But, it's unclear if the current trend of price stability will continue over the next 20 or 30 years for our current acquisition programs. With $15+ trillion in direct national debt, we can envision at least one unpleasant yet plausible scenario for the general level of prices in the U.S. economy. The big econometric models, by the way, simply cannot predict turning points in any economic aggregate such as the rate of inflation. Nevertheless, the view of future price stability, or lack thereof, will influence the choice of CV values to be used as benchmarks for supporting eSBM.

[17] Average annual rate of inflation during a presidency, as measured by the Consumer Price Index. Arguably, inflation for defense was higher.
[18] Because of long profiles for expenditures of TOA (seven years for ships, for example), a blip upward in inflation in one year perturbed not only an acquisition program's budget in that year but in many years previously, too, thus amplifying the problem.

4.4 Sample Data at MS C

Turning to MS C, the SAR Production Estimate (PdE) is of total program acquisition costs, including the sunk cost of development. Out of the 100 programs in the database, 43 were for MS C estimates, with Table 6 showing overall results.

Cost Growth Factors and CVs for DON MDAPs at MS C

	Without Quantity Adjustment		Quantity Adjusted	
Statistics	Base-Year$	Then-Year$	Base-Year$	Then-Year$
Mean	1.11	1.08	1.11	1.10
Standard Deviation	0.50	0.58	0.21	0.28
CV	0.45	0.53	0.19	0.26

Table 6

The values exhibit an across-the-board drop from MS B estimates. This results not only from the inclusion of sunk development costs in the calculations, but probably also from increased program knowledge and program stability moving from MS B to MS C.

As before, CVs were analyzed by type of platform, i.e., ships and submarines, aircraft, and other.[19] As was the case for Milestone B programs, CGFs at Milestone C were remarkably close, with Figure 16 showing means and ranges.

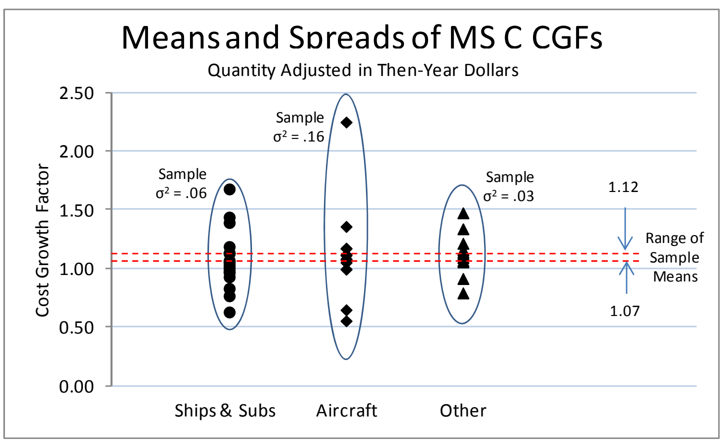

Figure 16. Means and Spreads of MS C CGFs

The relatively wide span for aircraft CGFs is driven entirely by the EA-6B outlier, with a CGF of 2.25, indicating 125% cost growth. Eliminating this datum reduces the aircraft CV (quantity-adjusted in then-year dollars) from 36% to 22%, a value in line with that of ships and submarines (22%) and “other” (16%). Even in the presence of the outlier, the null hypothesis of constant CGF population means is not

[19] A paucity of data did not allow use of the same platform categories as for Milestone B.

rejected at the 5% level of significance. For the null hypothesis of constant population variances, on the other hand, results are mixed. Levene's test supports the null hypothesis whereas pairwise F-tests reject it in cases involving the outlier. On balance, then, there's moderately strong support for the conjecture of homogeneous CVs at Milestone C.[20]

As was the case for Milestone B, Figure 17 shows a pronounced drop in CVs from the 1980s to the 1990s at Milestone C. Reasons might include better cost estimating, an increase in program stability, better linkage of the SAR BE to an ICE, decreased inflation volatility, or the results of previous efforts in acquisition reform.

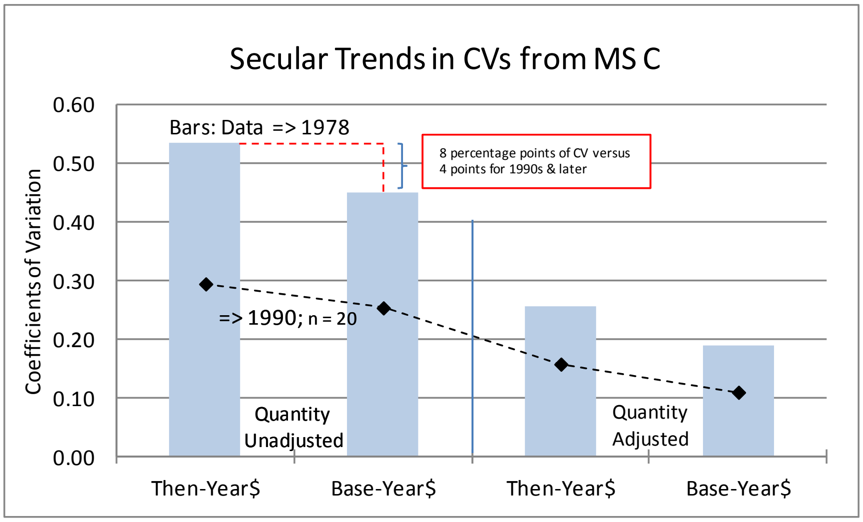

Figure 17. Secular Trend from MS C

4.5 Sample Data at MS A

For Milestone A, the sample size of seven was insufficient for making any statistically sound inferences. Estimation by analogy seems a logical alternative. Assuming that the degree of risk and uncertainty is the same between MS A and MS B as it is between MS B and MS C, then the application of roughly 15 percentage points of additional CV seems appropriate at MS A.

[20] These results are "not surprising. The F-test is closely tied to the assumption of data normality. It's not reliable if the distribution of data is significantly non-normal. On the other hand, Levene's test makes no distributional assumptions, but tends to favor the null unless the counterevidence is strong." Comment from Dr. Steve Book, Jan 2012.

4.6 Operational Construct

Figure 18 and Appendix C show benchmark CVs by milestone. The choice of which values to use for eSBM or as benchmarks for Monte Carlo simulation will likely depend upon the unique circumstances of a given acquisition program as well as organizational views on issues such as the likelihood of significant volatility in outyear rates of inflation and the effects on costs of current acquisition initiatives. Keep in mind that low rather than high estimates of CVs have been the norm in the defense cost community.

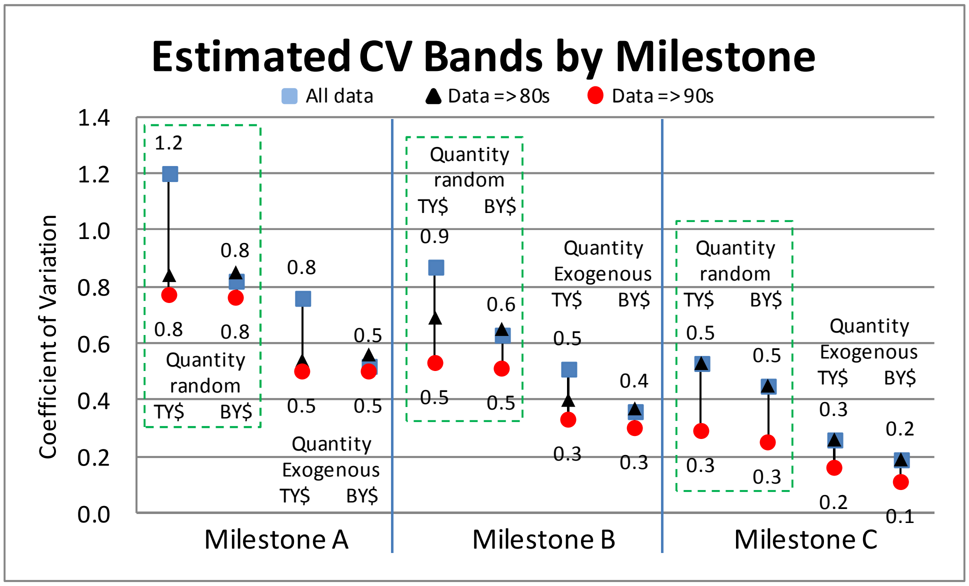

Figure 18. Operational Contruct

4.7 Summary of Findings

We offer these observations regarding the accuracy of conjectured CV behavior:

- Consistency
  - Conjecture: CVs from ICEs and cost assessments jibe with acquisition experience
  - Finding: Ad hoc observation suggests a pervasive underestimation of CVs in the international defense community
- Tendency to Decline During Acquisition Phase
  - Conjecture: CVs decrease throughout acquisition lifecycle
  - Finding: Strongly supported
- Platform Homogeneity
  - CVs are equivalent for aircraft, ships, and other platform types
  - Finding: Strongly supported, especially for MS B
- Tendency to Decrease after Normalization
  - CVs decrease when adjusted for changes in quantity and inflation
  - Finding: Strongly supported
- Invariance of Secular Trend
  - CVs steady long-term

    - Finding: Strongly rejected

4.8 Recommendations

Based on the forgoing analysis, we offer these recommendations:

- Define the type of CV employed or under discussion
    - The spreads of max-to-min values of the four types of CVs presented here (unadjusted and adjusted for quantity and inflation) are simply too large to do otherwise.
- Use a quantity-adjusted, then-year dollar CV for most acquisition programs
    - That is, regard quantity as exogenous but inflation as random in generating S-curves.
- Define CV benchmark values in terms of bands or ranges at each milestone
    - Use of single values presumes a level of knowledge and degree of certainty that simply doesn't exist.
    - A view of future price stability would argue for the use of lower CVs and instability for higher.
    - A belief in the positive effect of structural change due to recent acquisition initiatives would argue for lower CVs.
- Exercise prudence in choosing CV benchmarks.
    - Better to err on the side of caution and choose high-end benchmark values until costs of completed acquisition programs clearly demonstrate lower CGFs and CVs.
- Choose the high-end of benchmark CV bounds established at Milestone A to support AoAs and Materiel Development Decisions.
- Define a “trigger point” or floor for CV estimates, for each milestone, below which a call-for-explanation will be required
    - Employ trigger points for both Monte Carlo simulation and eSBM.
    - Base trigger points on confidence intervals for the CVs.

5.0 The S-Curve Tool

To support the development of better probabilistic cost estimates, the Naval Center for Cost Analysis (NCCA) has championed the development of the S-Curve Tool, which was well received at the 44th Annual Department of Defense Cost Analysis Symposium (DODCAS) in February, 2011, the joint Society of Cost Estimating and Analysis (SCEA) / International Society of Parametric Analysts (ISPA) conference in June, 2011, and the 45th ADoDCAS in February, 2012.

The purposes of the S-Curve Tool are to allow practitioners to easily and clearly

(1) compare their s-curve to another s-curve;
(2) compare their results to historical coefficients of variations (CVs) and/or cost growth factors (CGFs);
(3) generate graphics for decision briefs.

Figure 19 shows a flowchart diagram of the S-Curve Tool beta v2.0. For the estimate(s), the user chooses either Empirical (i.e., a set of outcomes from a Monte Carlo risk run), Parametric (e.g., enhanced Scenario-Based Method (eSBM) or parameters from an external risk analysis), or a Point Estimate (i.e., risk analysis not yet done).

If the estimate type is Empirical, the user inputs (1) the number of trials, (2) the cost units for the empirical data, and (3) all of the values for the trial runs. There is an optional feature to assess the empirical data by overlaying a parametric curve created by the empirical parameters on the raw data. For the optional feature, the user selects either the normal or lognormal distribution.

If the estimate type is Parametric, the user defines the type of distribution (either normal or lognormal) and the type of parameters. There are three options for parametric inputs in the tool, (1) Mean and CV, (2) Mean and Specified Cost (Xp) with corresponding %tile (p), and (3) CV and Specified Cost (Xp) with corresponding %tile (p). There are other ways to define a parametric curve (e.g., two percents (p) with two specified costs (Xp)), but they are not implemented in beta v2.0.

If the estimate type is a Point Estimate, the user defines the type of distribution (either normal or lognormal) and whether the point estimate is a Mean or a Median. If the point estimate is a median, the historical adjustment "pivots" on the median. All other cases (including the Parametric and Empirical cases) "pivot" on the mean when the estimate is historically adjusted.

Historical adjustments are based on the analysis of Selected Acquisition Reports (SARs), which is mentioned in Section 4 of this paper. These adjustments in the S-Curve Tool are dependent on five different inputs: (1) commodity, (2) life cycle phase, (3) milestone, (4) inflation, and (5) quantity. After the user selects these inputs, there are three options to apply the historical adjustment to the estimate: (1) apply CV only ("shaping" the s-curve); (2) apply CGF only ("shifting" the s-curve); and (3) apply CV & CGF ("shaping" and "shifting" the s-curve). If users decide not to apply historical adjustments to the estimate, they can proceed with the base s-curve that was generated. The historical benchmarks that are stored in the NCCA S-Curve Tool beta v2.0 are derived from the 100 programs mentioned in Section 4.2 of this paper.

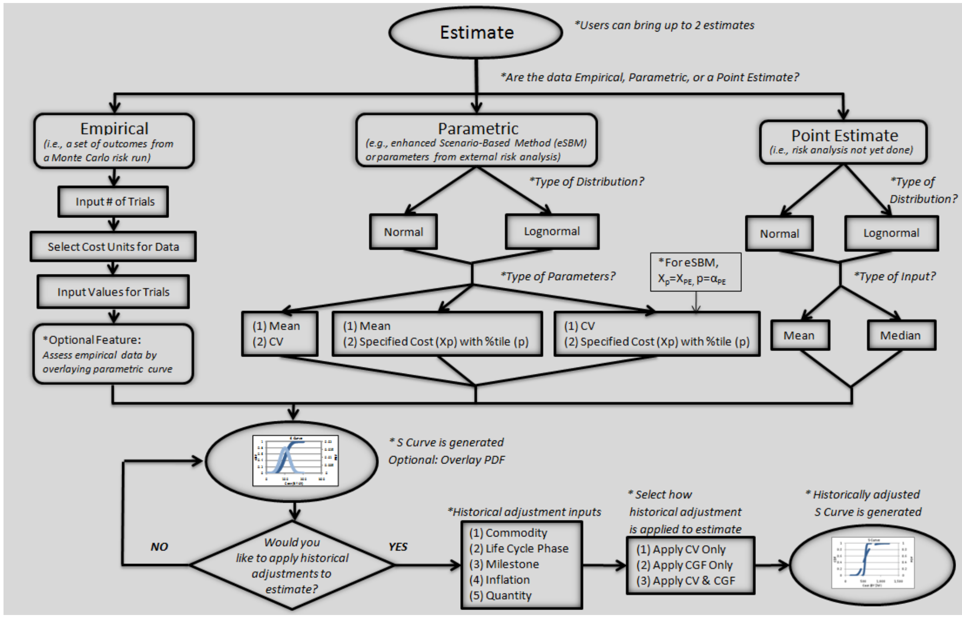

Figure 19. Flowchart Diagram of NCCA S-Curve Tool beta v2.0

There are ongoing research efforts to support both continued improvement of the S-Curve Tool and greater understanding of the nature of cost growth for major acquisition programs; its mean value (risk) and variability (uncertainty); and the components thereof. There have been tremendous results of extensive data collection, validation, normalization, and analysis using cost variance data from SARs across all Services DoD components. By shifting from the SAR Summaries to the SARs themselves, the authors were able to decompose the data set used in the S-Curve Tool, which were at the level of total Acquisition cost with Quantity and Economic adjustments only, into appropriation types – Research, Development, Test, and Evaluation (RDT&E), Procurement, Military Construction (MILCON), and (Acquisition-phase) Operating and Support (O&S) – and all seven SAR Cost Variance categories.

Two additional categories were identified and quantified: (1) Baseline Adjustments (identified elsewhere in the SAR) and (2) Inter-Phase growth, which occurs when the initial Baseline Estimate of one phase does not match the final Current Estimate of the previous phase. The current data set comprises of more than 400 milestone estimates from more than 300 programs. The expanded and refined data set will be used to update the historical benchmarks in the NCCA S-Curve Tool beta v2.0. Updates are planned to be posted on NCCA's website in the near future.

To obtain a copy of the NCCA S-Curve Tool beta v2.0 and any related documentation (e.g., User Guide and Technical Manual), please visit https://www.ncca.navy.mil/tools/tools.cfm.

6.0 eSBM Case Study

The North Atlantic Treaty Organization (NATO) is acquiring an Alliance Ground Surveillance (AGS) system with the capability of performing persistent surveillance and reconnaissance over wide areas from high-altitude, long-endurance, Global Hawk unmanned aerial vehicles (UAVs). The Multi-Platform Radar Technology Insertion Program (MP-RTIP) payload will enable the AGS UAVs to look at what is happening on the earth's surface, providing situational awareness before, during, and after NATO operations, through interfaces with a to-be-designed-and-developed ground segment. The ultimate objective of AGS is "… to make soldiers from all NATO countries safer and more effective when they are deployed on operations."[21] At NATO's *Heads of State and Government* summit in Lisbon in 2010, President Obama and his 27 national counterparts collectively reaffirmed the acquisition of AGS as "a top priority" for the Alliance.[22]

NATO's SAS-076 Task Group used eSBM to perform risk and uncertainty cost analysis for the AGS acquisition. This was part of a larger effort to generate an ICE on the program, the first ever for a NATO weapon system acquisition program conducted by NATO.[23]

To conduct eSBM, the Task Group needed to

- Generate a point estimate for acquisition costs,
- Identify the position of the point estimate on the S-Curve,
- Identify and analyze major elements of risk and uncertainty,
- Select an appropriate CV,
- Develop scenarios, and
- Combine these components into an integrated whole.

6.1 Point Estimate and Position on S-Curve

The Task Group employed a number of techniques to estimate the costs of AGS. These included learning curves, averages of historical data, CERs, and analogies. In many cases, cross checks were developed based on German experience with Eurohawk. Since it's necessary in eSBM to anchor a cost estimate to a point on a cumulative probability distribution, baseline costs were generated, by design, at the *median*, or the 50th percentile. Another choice could have been the *mean*. Generally, there's flexibility in choosing either, or perhaps a point in between.

In the case of AGS, many cost elements were estimated using unit learning curves or power-function CERs with a multiplicative random error term (e.g., $Y = \alpha Q^{\beta} e^{\varepsilon}$, where Y = unit cost, Q = lot-midpoint quantity, α and β are parameters (T1 and elasticity), and $\varepsilon \sim N(\mu,\sigma^2)$). Examples include the wing, fuselage, and empennage of the UAV, and final assembly, integration, test and check out. In these cases, plugging a value of an explanatory variable into the equation yields an estimated *median* rather than *mean*

[21] Mr. Anders Fogh Rasmussen, Secretary General of NATO; 22 June 2010.
[22] Lisbon, Portugal, 19 and 20 November, 2010.
[23] The Systems Analysis and Studies (SAS) Task Group 076 is chaired by Dr. Flynn, and consists of representatives from Sweden, a dozen NATO countries, and OCCAR (Organisation conjointe de coopération en matière d'armement), a European armaments agency. Members of the NATO AGS ICE team are Dr. Brian Flynn, Dr. Paul Desmier, Mr. Walter Boos, Mr. Joachim Schotten, Mr. Phiippe Lacheret, Mr. Murat Cakmak, Mr. Renzo Chiomento, Dr. Dagfinn Vatne, and Ms. Solveig Krey.

value.[24] In other cases, such as for software development where representatives from several participating nations each generated a cost estimate independently, a middle value (*median*) was selected as the baseline. Moreover, the CERs employed in producing the middle value were themselves *median*-yielding power-function equations. In other cases, where costs appear normally distributed, the choice or median or mean is a moot point since the values are equal. Examples include systems engineering and program management, initial spares, and support equipment.

6.2 Risk Elements

Next, the Task Group identified these major areas of cost risk and uncertainty:[25]

- **Exchange Rate**

  The AGS contract will be a firm-fixed price direct commercial sale to Northrop Grumman, with a ceiling price denominated in 2007 base-year euros, but with much of the work done in the United States. Converting from dollars to euros, then, is a major issue. Unfortunately, currency exchange rates are notoriously difficult if not impossible to predict accurately and consistently. The projections of Figure 20, using random walk theory, don't exactly inspire confidence in anyone's ability to hone in on the future value of the $/€ exchange rate.[26]

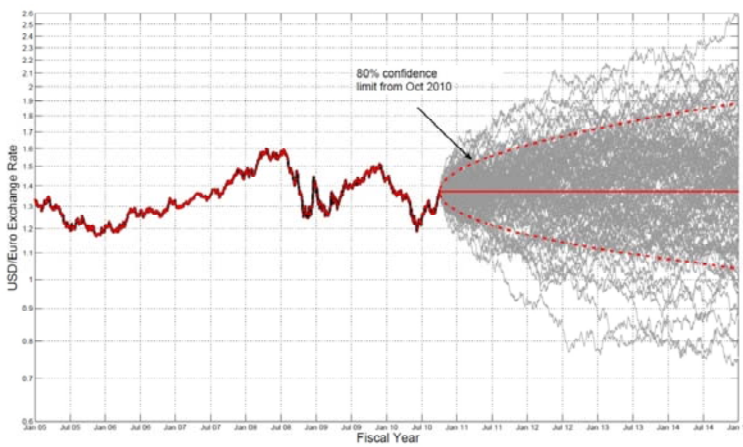

Figure 20. Projections of Euro/Dollar Exchange Rate

  Since its introduction roughly a decade ago, the value of the euro has varied from a low of $0.83/€ in 2000 to a peak of $1.60/€ in 2008, a swing of 93%. More recently, during the height of the Greek credit crisis in early 2010, the euro fell to $1.18/€. It then returned to pre-crisis levels only to fall once again with the Irish debt crisis.

---

[24] Goldberger, Arthur S., "The Interpretation and Estimation of Cobb-Douglas Functions," Econometrica, Vol. 35, July-October, 1968, pp. 464-472. "… the standard specification and approach to estimation shift attention, apparently unwittingly, from the mean to the median as a measure of central tendency."

[25] Areas of risk were identified based on general knowledge of defense acquisition programs and the discipline of international economics, site visits to Northrop Grumman's International Program Office in Florida and to NATO's AGS Management Agency in Brussels, and on meetings with other key AGS acquisition officials, including the Chairman of the Board of Directors of the NATO AGS Management Organization.

[26] Projections are from Dr. Paul Desmier of the Department of National Defense, Canada.

- **Inflation**
  The ICE team is using a baseline value of 3% inflation per annum for outyear projections, weighted according to the relative contributions of the 13 NATO countries participating in the program. However, as Figure 21 shows, inflation in Europe, as measured by the consumer price index, seems to be related to the growth rate in GDP. EU inflation rates fell precipitously during the global financial crisis of 2008 but then resumed their secular trend with economic recovery. If the current recession in Europe is short-lived, an uptick in inflation during the procurement of NATO AGS could occur.

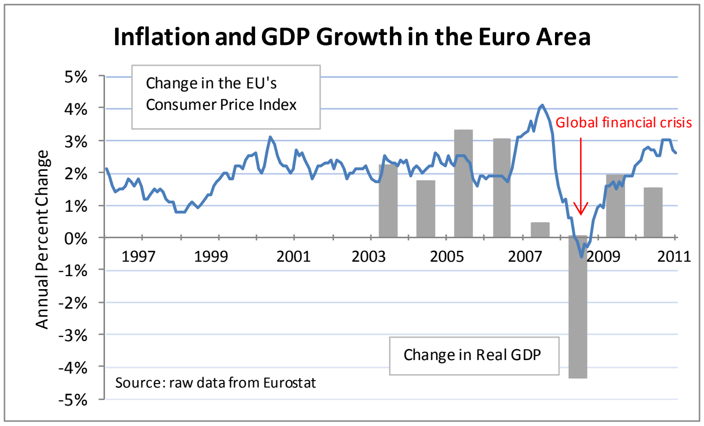

Figure 21. Inflation and GDP

- **Schedule**
  The acquisition of NATO AGS has continually slipped. Further delays will increase then-year dollar and euro costs due to inflation.

- **Software Development**
  European participants in the AGS program, and Canada, are responsible for ground segment design, development, and build. Elements of the ground segment include several types of ground vehicles, command and control units, training equipment, and an extensive software development effort. The *baseline* count of equivalent source lines of code (ESLOC) is unusually large from a U.S. perspective, [27] and includes no factor for *growth*.

  In the case of AGS, software will be developed in many different countries by many different companies, possibly using different computer languages and even operating systems. Levels of Capability Maturity Model Integration (CMMI) certification vary among vendors. Integration of software modules, hardware with software, and AGS with other ISR assets such as ASTOR from the U.K. and Global Hawk from the U.S. will all be required. Configuration management and

[27] Unlike software for Global Hawk in the U.S., AGS includes software for not only the sensor data collection element (the UAV, sensor and control stations), but also for the user exploitation elements (mobile and transportable ground stations) and a support segment (trainers, maintenance equipment, etc). In the U.S. environment, the operational exploitation and use of the collected ISR data is performed outside of the Global Hawk system (through the Distributed Common Ground Systems), while for NATO the AGS system will need to include this essential part as well.

software integration will be major issues. The AGS ICE team uses the *average* historical growth in ESLOC count as a proxy variable for risk for the entire software development effort.

- **Radar**
  The MP-RTIP payload uses advanced electronically scanned array (AESA) technology currently employed on the F/A-18, among other platforms. However, the MP-RTIP development program has experienced significant cost and schedule growth which might translate into higher unit production costs.

- **International Participation**
  Thirteen of NATO's 28 members are funding the acquisition of AGS. Each, in return, demands a fair share of the work. NATO intends to award a contract to Northrop Grumman Integrated Systems Sector International, Inc., NGISSI, who will have total system performance management responsibility and will subcontract work in various countries, as shown in Figure 22. Developing and producing hardware and software under this work-share constraint in such a multi-cultural, geographically-dispersed environment runs the risk of introducing inefficiencies into the program.

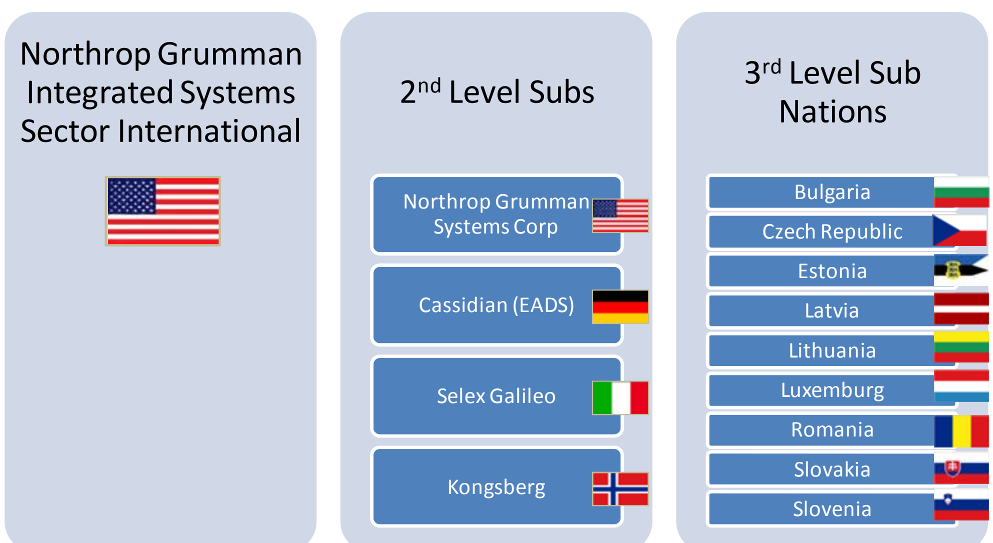

Figure 22. AGS International Contracting

- **Affordability**
  Fitting the desired numbers and capabilities of UAVs and ground-segment elements under NATO's ceiling price remains a challenge. A few years ago NATO and industry envisioned a mixed fleet of manned and unmanned air vehicles to perform the ground surveillance mission. However, that option was scrapped because of cost. Affordability continues to be an issue as many European nations struggle to reduce budget deficits and national debt. Denmark, for example, announced its intention in 2010 to withdraw from the program because of budget tightening.

6.3 Selection of CV

With a point estimate generated and major elements of risk identified, the Task Group needed a CV to implement eSBM. NATO, it's important to note, uses the Phased Armaments Programming System (PAPS) as its acquisition framework. In PAPS, AGS is near what we call a Milestone B decision in the United States. Further, although AGS is an Alliance rather than a U.S. DoD acquisition program, the prime contractor will be Northrop Grumman, and close to two-thirds of the costs will be incurred in the U.S. The use of benchmark CV data from the U.S., then, seems appropriate, especially since none is available for NATO acquisitions.

The Task Group selected a MS B CV of 51% based on the full sample of data for quantity-adjusted then-year dollar acquisition outcomes. That is, the Task Group regarded the quantity of UAVs and ground vehicles as exogenous but the rate of inflation as random, for purposes of generating an S-curve. Given affordability and exchange rate issues, the massive software development effort, and extensive international participation with each country demanding "noble work," the Task Group judged the higher-end of MS B CV values to be appropriate.

6.4 Scenarios

The Task Group bundled risks and uncertainties into these scenarios:[28]

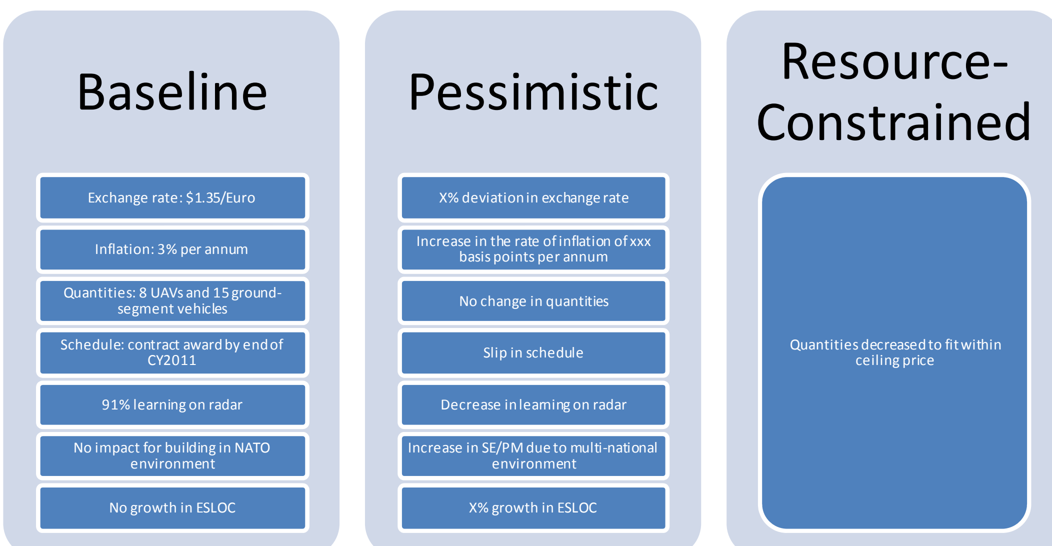

Figure 23. eSBM Scenarios for AGS

In keeping with the tenets of eSBM, the non-baseline scenarios did not represent *extreme* cases but rather a set of conditions that could easily occur in the future. For example, the *worst imaginable* case for the U.S. dollar or the Euro might be a severe devaluation of either.[29] On the other hand, modest appreciation or depreciation of the dollar or Euro is certainly plausible, depending upon circumstances.

---

[28] Data that might be construed as business sensitive are omitted from the display.

[29] Pundits in Europe and the U.S. speculated in February 2010, for example, that the European Union, and the Euro, might not survive. Other pundits project financial Armageddon for the U.S. because of our enormous direct national debt ($14 trillion) and unfunded liabilities for Social Security, Medicare, and other entitlement programs (another $50 trillion).

6.5 S-Curve

With a point estimate, CV, and anchored position on the S-curve in hand, eSBM results were calculated, with Figure 24 showing results. Costs are not displayed on the X-axis due to the business-sensitive nature of the information.

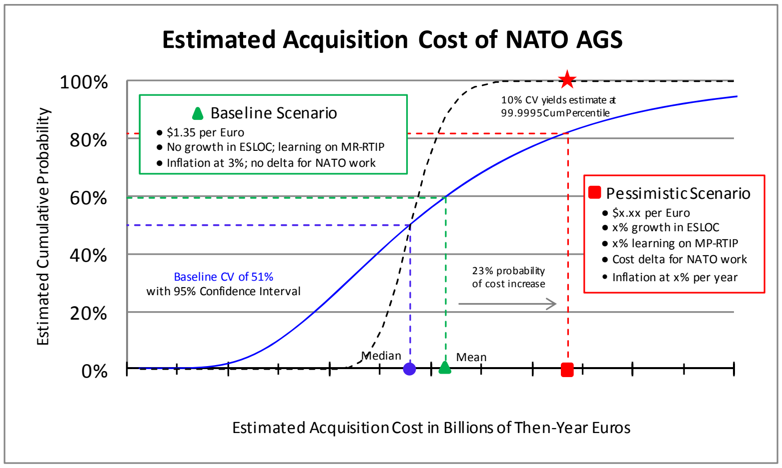

Figure 24. S-Curve for NATO AGS Cost

Nevertheless, several points of the analysis merit emphasis:

- Not all areas of risk and uncertainty are captured in the scenarios, nor do they need to be. That's the great advantage of employing an historically-based CV in computing an S-curve. The distribution implicitly captures the totality of events that will influence the cost of the acquisition. A maintained hypothesis, of course, is that historical variances will continue to hold.

- The scenarios reflect cost estimates for the events or assumptions they portray. In supporting the NATO AGS Management Agency (NAGSMA) in Brussels, the Task Group estimated and presented costs of each element of risk (e.g., the cost impact of a 50% increase in ESLOC count) within a scenario. The associated probabilities for the scenarios were simply *read* from the S-curve.

- The resource-constrained scenario was regarded as a distinct, separate what-if drill. Estimating the risk and uncertainty of changes in acquisition quantity is generally beyond the purview of cost analysts. Such changes often result from Congressional action, which can be hard to predict. The historical CVs were all adjusted for quantity variation; if they weren't, they'd be far higher than they are.

- The probability of NATO AGS acquisition costs increasing from the baseline to the pessimistic scenario is roughly one in four or five. Identification and quantification of this possibility can form the basis of risk-management planning by both the program office (NAGSMA) and the prime contractor.

6.6 S-Curve Excursion

To demonstrate the consequences of using an inaccurate CV, Figure 23 shows a second S-curve based on a hypothetical value of 10%.

If the 10% value had been used in the actual analysis, the probability of costs reaching or exceeding the value of the pessimistic scenario would be calculated as only *five in one million*! Yet, to repeat, the parameters used in this scenario are realistic, modest deltas to the baseline. They include what the Task Group regarded as a very slight change in the exchange rate, a modest increase in the rate of inflation, use of a *mean* growth factor for number of equivalent source lines of code, and a percentage increase in the cost of system engineering/program management due to international participation that at least one member of the team thought was appropriate only if two rather than14 nations were involved in AGS design, development, and build. In short, the steep S-curve would deceive stakeholders into believing that risk was slight and the probability of significant cost growth minimal.

Specification and adoption of a pessimistic scenario in eSBM, which can easily be adopted in Monte Carlo methodology, too, has the virtue of providing the means with which to perform a sanity check on the value of the CV. If the resulting S-curve is too steep to be plausible, then underlying risk and uncertainty distributions for major cost elements must not have been calculated correctly. Chances are that at least one of them in the actual analysis was too narrow.

Finally, since the mean of a lognormal distribution involves a variance term, use of a steeper S-curve would have resulted in a lower value of expected cost.

7.0 Summary

In 2006, the Scenario-Based Method (SBM) was introduced as an alternative to advanced statistical methods for generating measures of cost risk. Since then, enhancements to SBM have been made. These include integrating historical cost performance data into SBM's algorithms and providing a context for applying SBM from a WSARA perspective. Together, these improvements define the enhanced SBM (eSBM) – an historical data-driven application of SBM.

Among other things, eSBM does the following:

- Provides an analytic argument for deriving the amount of cost reserve needed to guard against well-defined "scenarios";
- Brings the discussion of "scenarios" and their credibility to the decision-makers; this is a more meaningful topic to focus on, instead of statistical abstractions that simulation approaches can sometimes create;
- Does not require the use of statistical methods to develop a valid measure of cost risk reserve; this is the non-statistical eSBM;
- Allows percentiles (confidence measures) to be designed into the approach with a minimum set of statistical assumptions;
- Allows percentiles (as well as the mean, median (50th%), variance, etc.) to be calculated algebraically in near-real time within a simple spreadsheet environment;

- Avoids the requirement to develop probability distribution functions for all the uncertain variables in a WBS, a task that can be time-consuming and hard to justify;
- Captures correlation indirectly in the analysis by the magnitude of the coefficient of variation applied in the statistical eSBM;
- Supports traceability and focuses attention on key risk events in the written scenarios that have the potential to drive cost higher than expected.

In summary, eSBM encourages and emphasizes a careful and deliberative approach to cost-risk analysis. It requires the development of scenarios that represent the program's "risk story" rather than a debate about what percentile to select. Time is best spent building the case arguments for how a confluence of risk events that form a risk scenario might drive the program to a particular percentile. This is where the debate and the analysis should center.

APPENDIX A
Percentiles of the Standard Normal Random Variable

The tables below are values of the cumulative distribution function of the standard normal random variable Z, where $Z \sim N(0, 1)$. The columns with three-digits represent values for Z denoted by the small letter z. The columns with the eight-digits are equal to the probability given by the integral below.

$$P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

Since $Z \sim N(0, 1)$ the following is true; $P(Z \le -z) = P(Z > z) = 1 - P(Z \le z)$.

0.00	0.5000000	0.21	0.5831661	0.42	0.6627572	0.63	0.7356528
0.01	0.5039894	0.22	0.5870644	0.43	0.6664021	0.64	0.7389138
0.02	0.5079784	0.23	0.5909541	0.44	0.6700314	0.65	0.7421540
0.03	0.5119665	0.24	0.5948348	0.45	0.6736448	0.66	0.7453732
0.04	0.5159535	0.25	0.5987063	0.46	0.6772419	0.67	0.7485712
0.05	0.5199389	0.26	0.6025681	0.47	0.6808225	0.68	0.7517478
0.06	0.5239223	0.27	0.6064198	0.48	0.6843863	0.69	0.7549030
0.07	0.5279032	0.28	0.6102612	0.49	0.6879331	0.70	0.7580364
0.08	0.5318814	0.29	0.6140918	0.50	0.6914625	0.71	0.7611480
0.09	0.5358565	0.30	0.6179114	0.51	0.6949743	0.72	0.7642376
0.10	0.5398279	0.31	0.6217195	0.52	0.6984682	0.73	0.7673050
0.11	0.5437954	0.32	0.6255158	0.53	0.7019441	0.74	0.7703501
0.12	0.5477585	0.33	0.6293000	0.54	0.7054015	0.75	0.7733727
0.13	0.5517168	0.34	0.6330717	0.55	0.7088403	0.76	0.7763728
0.14	0.5556700	0.35	0.6368306	0.56	0.7122603	0.77	0.7793501
0.15	0.5596177	0.36	0.6405764	0.57	0.7156612	0.78	0.7823046
0.16	0.5635595	0.37	0.6443087	0.58	0.7190427	0.79	0.7852362
0.17	0.5674949	0.38	0.6480272	0.59	0.7224047	0.80	0.7881447
0.18	0.5714237	0.39	0.6517317	0.60	0.7257469	0.81	0.7910300
0.19	0.5753454	0.40	0.6554217	0.61	0.7290692	0.82	0.7938920
0.20	0.5792597	0.41	0.6590970	0.62	0.7323712	0.83	0.7967307

Table A-1. Percentiles of the Standard Normal Values (continued)

Example Computations

1. $P(Z \le z = -0.525) = P(Z > z = 0.525) = 1 - P(Z \le z = 0.525) = 1 - 0.70 = 0.30$
2. $P(Z \le z = -0.675) = P(Z > z = 0.675) = 1 - P(Z \le z = 0.675) = 1 - 0.75 = 0.25$
3. $P(Z \le z = 0.525) = 0.70$

0.84	0.7995459	1.05	0.8531409	1.26	0.8961653	1.47	0.9292191
0.85	0.8023375	1.06	0.8554277	1.27	0.8979576	1.48	0.9305633
0.86	0.8051055	1.07	0.8576903	1.28	0.8997274	1.49	0.9318879
0.87	0.8078498	1.08	0.8599289	1.29	0.9014746	1.50	0.9331928
0.88	0.8105704	1.09	0.8621434	1.30	0.9031995	1.51	0.9344783
0.89	0.8132671	1.10	0.8643339	1.31	0.9049020	1.52	0.9357445
0.90	0.8159399	1.11	0.8665004	1.32	0.9065824	1.53	0.9369916
0.91	0.8185888	1.12	0.8686431	1.33	0.9082408	1.54	0.9382198
0.92	0.8212136	1.13	0.8707618	1.34	0.9098773	1.55	0.9394292
0.93	0.8238145	1.14	0.8728568	1.35	0.9114919	1.56	0.9406200
0.94	0.8263912	1.15	0.8749280	1.36	0.9130850	1.57	0.9417924
0.95	0.8289439	1.16	0.8769755	1.37	0.9146565	1.58	0.9429466
0.96	0.8314724	1.17	0.8789995	1.38	0.9162066	1.59	0.9440826
0.97	0.8339768	1.18	0.8809998	1.39	0.9177355	1.60	0.9452007
0.98	0.8364569	1.19	0.8829767	1.40	0.9192433	1.61	0.9463011
0.99	0.8389129	1.20	0.8849303	1.41	0.9207301	1.62	0.9473839
1.00	0.8413447	1.21	0.8868605	1.42	0.9221961	1.63	0.9484493
1.01	0.8437523	1.22	0.8887675	1.43	0.9236414	1.64	0.9494974
1.02	0.8461358	1.23	0.8906514	1.44	0.9250663	1.65	0.9505285
1.03	0.8484950	1.24	0.8925122	1.45	0.9264707	1.66	0.9515428
1.04	0.8508300	1.25	0.8943502	1.46	0.9278549	1.67	0.9525403

1.68	0.9535214	1.89	0.9706211	2.10	0.9821356	2.31	0.9895559
1.69	0.9544861	1.90	0.9712835	2.11	0.9825709	2.32	0.9898296
1.70	0.9554346	1.91	0.9719335	2.12	0.9829970	2.33	0.9900969
1.71	0.9563671	1.92	0.9725711	2.13	0.9834143	2.40	0.9918025
1.72	0.9572838	1.93	0.9731967	2.14	0.9838227	2.50	0.9937903
1.73	0.9581849	1.94	0.9738102	2.15	0.9842224	2.60	0.9953388
1.74	0.9590705	1.95	0.9744120	2.16	0.9846137	2.70	0.9965330
1.75	0.9599409	1.96	0.9750022	2.17	0.9849966	2.80	0.9974448
1.76	0.9607961	1.97	0.9755809	2.18	0.9853713	2.90	0.9981341
1.77	0.9616365	1.98	0.9761483	2.19	0.9857379	3.00	0.9986500
1.78	0.9624621	1.99	0.9767046	2.20	0.9860966	3.10	0.9990323
1.79	0.9632731	2.00	0.9772499	2.21	0.9864475	3.20	0.9993128
1.80	0.9640697	2.01	0.9777845	2.22	0.9867907	3.30	0.9995165
1.81	0.9648522	2.02	0.9783084	2.23	0.9871263	3.40	0.9996630
1.82	0.9656206	2.03	0.9788218	2.24	0.9874546	3.50	0.9997673
1.83	0.9663751	2.04	0.9793249	2.25	0.9877756	3.60	0.9998409
1.84	0.9671159	2.05	0.9798179	2.26	0.9880894	3.70	0.9998922
1.85	0.9678433	2.06	0.9803008	2.27	0.9883962	3.80	0.9999276
1.86	0.9685573	2.07	0.9807739	2.28	0.9886962	3.90	0.9999519
1.87	0.9692582	2.08	0.9812373	2.29	0.9889894	4.00	0.9999683
1.88	0.9699460	2.09	0.9816912	2.30	0.9892759	5.00	0.9999997

Table A-1. Percentiles of the Standard Normal Values (concluded)

APPENDIX B
The Lognormal Distribution

The lognormal probability distribution has broad applicability in engineering, economics, and cost analysis. In engineering, the failure rates of mechanical or electrical components often follow a lognormal distribution. In economics, the random variation between the production cost of goods to capital and labor costs is frequently modeled after the lognormal distribution. In cost analysis, the lognormal often approximates the probability distribution of a program's total cost [Young, 1995; Garvey, 2000].

If X is a nonnegative random variable and $\ln X$ follows the normal distribution, then X is said to have a lognormal distribution. The lognormal probability density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_Y} \frac{1}{x} e^{-(\ln X - \mu_Y)^2 / 2\sigma_Y^2}$$

where $0 < x < \infty$, $\sigma_Y > 0$, $\mu_Y = E(\ln X)$, and $\sigma_Y^2 = Var(\ln X)$. The parameters μ_Y and σ_Y^2 are the mean and variance of the normally distributed random variable $Y = \ln X$, which is the logarithmic representation of X. The lognormal is positively skewed and values for x are always nonnegative.

If X is a lognormal random variable with mean μ_Y and variance σ_Y^2, then

$$P(X \le x) = P\left(Z \le z = \frac{\ln x - \mu_Y}{\sigma_Y} \right)$$

Values for z are available in Appendix A of this paper.

Theorem C-1: If X is a lognormal random variable, then X has mean and variance

$$E(X) = \mu_X = e^{\mu_Y + \frac{1}{2}\sigma_Y^2}$$

$$Var(X) = \sigma_X^2 = e^{2\mu_Y + \sigma_Y^2}(e^{\sigma_Y^2} - 1).$$

Theorem C-2: If X is a lognormal random variable, with mean $E(X) = \mu_X$ and $Var(X) = \sigma_X^2$, then

$$\mu_Y = E(\ln X) = \frac{1}{2} \ln \left[\frac{(\mu_X)^4}{(\mu_X)^2 + \sigma_X^2} \right]$$

and

$$\sigma_Y^2 = Var(\ln X) = \ln \left[\frac{(\mu_X)^2 + \sigma_X^2}{(\mu_X)^2} \right]$$

APPENDIX C

CVs for Estimates of Total Program Acquisition Cost

CV Calculation and Definition Using Complete Sample of Data	Milestone A	Milestone B	Milestone C
Quantity Random (Unadjusted for Qty Δ)			
Then-year dollars	1.20	0.87	0.53
Base-year dollars	0.82	0.63	0.45
Quantity Exogenous (Adjusted for Qty Δ)			
Then-year dollars	0.76	0.51	0.26
Base-year dollars	0.52	0.36	0.19

CV Calculation and Definition Using Data from 1980 and Later	Milestone A	Milestone B	Milestone C
Quantity Random (Unadjusted for Qty Δ)			
Then-year dollars	0.84	0.69	0.53
Base-year dollars	0.85	0.65	0.45
Quantity Exogenous (Adjusted for Qty Δ)			
Then-year dollars	0.54	0.40	0.26
Base-year dollars	0.56	0.37	0.19

CV Calculation and Definition Using Data from 1990 and Later	Milestone A	Milestone B	Milestone C
Quantity Random (Unadjusted for Qty Δ)			
Then-year dollars	0.77	0.53	0.29
Base-year dollars	0.76	0.51	0.25
Quantity Exogenous (Adjusted for Qty Δ)			
Then-year dollars	0.50	0.33	0.16
Base-year dollars	0.50	0.30	0.11

REFERENCES AND RELEVANT LITERATURE

Arena, M. V., et al. 2006. *Impossible Certainty: Cost Risk Analysis for Air Force Systems*, Project Air Force, published by The RAND Corporation.

Arena, M. V., et al. 2006. "Historical Cost Growth of Completed Weapon System Programs", Project Air Force, TR-343-AF, ©2006, The RAND Corporation,

Garvey, P. R. 2000. *Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective*, 2000, Marcel Dekker, Inc., Taylor & Francis Group, Chapman-Hall/CRC Press, New York, N.Y., 10016.

Garvey, P. R. 2008. "A Scenario-Based Method for Cost Risk Analysis", *Journal of Cost Analysis and Parametrics*, Vol. 1, Spring, 2008.

General Accountability Office (GAO) 2009. *Cost Estimating and Assessment Guide*.

Hitch, C. J. 1955. *An Appreciation of Systems Analysis*, P-699. Santa Monica, California: The RAND Corporation.

NASA *Cost Estimating Handbook*, 2008.

Public Law 111-23, 2009. "Weapon Systems Acquisition Reform Act of 2009", 22 May 2009.

Young, P. H. 1995. A Generalized Probability Distribution for Cost/Schedule Uncertainty in Risk Assessment. *Proceedings of the 1995 Western MultiConference on Business/MIS Simulation Education*, The Society for Computer Simulation, San Diego, California.

Younossi, O., et al. 2007. *Is Weapon System Cost Growth Increasing?: A Quantitative Assessment of Completed and Ongoing Programs*, ©2007, The RAND Corporation, ISBN 978-0-8330-4135-7.

United States Air Force, Air Force Cost Analysis Agency, 2007. *Cost Risk and Uncertainty Analysis Handbook (CRUH)*.