WEAPON SYSTEMS ACQUISITION REFORM ACT (WSARA) AND THE ENHANCED SCENARIO-BASED METHOD (eSBM) FOR COST RISK ANALYSIS

Brian J. Flynn[1], Ph.D.
Paul R. Garvey[2], Ph.D.

Presented to the 44th Annual Department of Defense Cost Analysis Symposium
February, 2011, Williamsburg, Virginia

Published by the Naval Center for Cost Analysis, Washington, DC
APRIL 2011

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 a WSARA perspective. 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, Coefficient of Variation, Normal Distribution, Lognormal Distribution, Monte Carlo Simulation

[1] Special Assistant to the Deputy Assistant Secretary of the Navy (Cost and Economics), Naval Center for Cost Analysis (NCCA), brian.flynn@navy.mil.

[2] 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.

1.0 Background

This paper presents eSBM. eSBM is an enhancement to the Scenario-Based Method (SBM), which was developed as an alternative to advanced statistical methods for generating measures of cost risk. 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. The following presents a brief background on SBM and its enhanced formulation eSBM.

Created in 2006, SBM was formed in response to a question posed by a government agency. The question was *Can a valid cost risk analysis, 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.

SBM has appeared in a number of publications. These include the RAND book *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 has been enhanced in two ways. First, historical cost data is integrated into SBM's algorithms. Second, applying SBM from a WSARA perspective is considered. Together, these enhancements define eSBM.

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 driven by past program performance. With its simplified analytics, eSBM eases the mathematical burden on analysts. It focuses 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

Cost estimates of defense programs are inherently uncertain. Estimates are often required when little about a program 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 often 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. Furthermore, 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. Limited programmatic information may be available on such key parameters as schedule, quantity of units to be bought, performance, requirements, acquisition strategy, and future evolutionary increments. Key characteristics of the system may change as it proceeds through development and production. Increases in system weight, complexity, and lines of code are commonplace.

For these reasons, a cost estimate expressed as a single number is merely one outcome in a probability distribution of costs. An estimate is stochastic rather than deterministic, with uncertainty determining the shape and variance of the distribution of possible cost outcomes.

The terms "risk" and "uncertainty" are often used interchangeably. However, there is an important distinction between them. Risk is the chance of loss or injury. In a situation that includes favorable and unfavorable events, risk is the probability an unfavorable event occurs. Uncertainty is the indefiniteness about the outcome of a situation. We analyze uncertainty for measuring risk. In systems engineering the analysis might focus on measuring the risk of: failing to achieve performance objectives, overrunning the budgeted cost, or delivering the system too late to meet user needs.[3]

1.2 Techniques

In its highest form, defense cost analysis is an amalgam of scientific rigor and sound judgment. It requires knowledge, insight, and the application of statistical principles, as well as the critical interpretation of a wide variety of information imprecisely. Indeed, Maynard Keynes' observation on "the extreme precariousness of the basis of knowledge on which our estimates … have to be made" [4] often applies in defense cost analysis, especially for pre-Milestone 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 is 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."[5] This contention remains an open issue within the international defense cost analysis community. Some practitioners concur, some do not, and some are unsure.

Amidst this spectrum of opinion, best-available techniques for conducting risk and uncertainty analysis of cost estimates of defense acquisition programs include sensitivity analysis, Monte Carlo simulation, and eSBM.[6] Each technique, *if used properly*, can yield scientifically sound results. A best practice is to employ more than one technique and to compare findings. For example, detailed Monte Carlo simulation and eSBM both yield S-curves. Yet, the two techniques are fundamentally different in approach, with the former bottoms-up and the latter top-down. Divergence in results between the two procedures is a clarion call for explanation, while consistency inspires confidence in the validity of the estimates.

---

[3] Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective, Garvey, P. R., Chapman & Hall/CRC-Press, 2000, page 27.
[4] The General Theory of Employment, Interest, and Money; Keynes, John Maynard; Harcourt Brace Jovanovich; 1964, page 149.
[5] Economic Theory and Operations Analysis, Baumol, William; Prentice-Hall; 1977, page 621.
[6] 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.

1.3 Cost Estimate Confidence: A WSARA Perspective

In May 2009, the US Congress passed the Weapon Systems Acquisition Reform Act (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. They are the organizational structure of the DOD, its acquisition policies, and its congressional reporting requirements. The following 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.

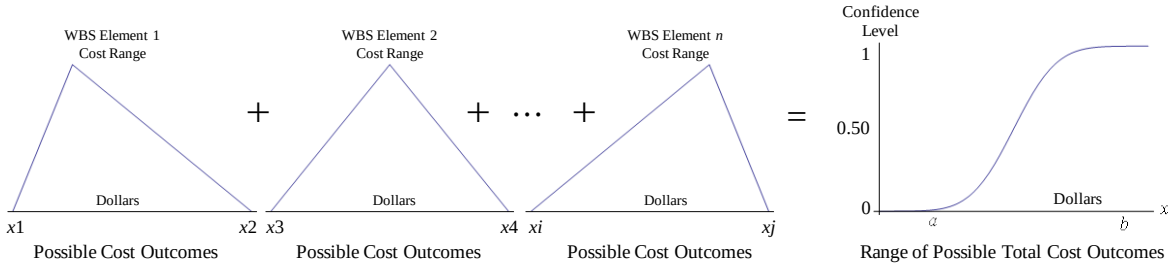

Figure 1. Cost Estimate Confidence: A Summation of Cost Element Cost Ranges

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. Each distribution is then combined by the probability calculus to generate an overall distribution of program total cost. This is 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 a cumulative probability distribution of a program's total cost. It is the output from a probability analysis of cost uncertainty, as described in Figure 1. Cost estimate confidence is read from this distribution. For example, there is a 25 percent chance the program will cost less than or equal to $100M, a 50 percent chance the program will cost less than or equal to $151M, and an 80 percent 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.

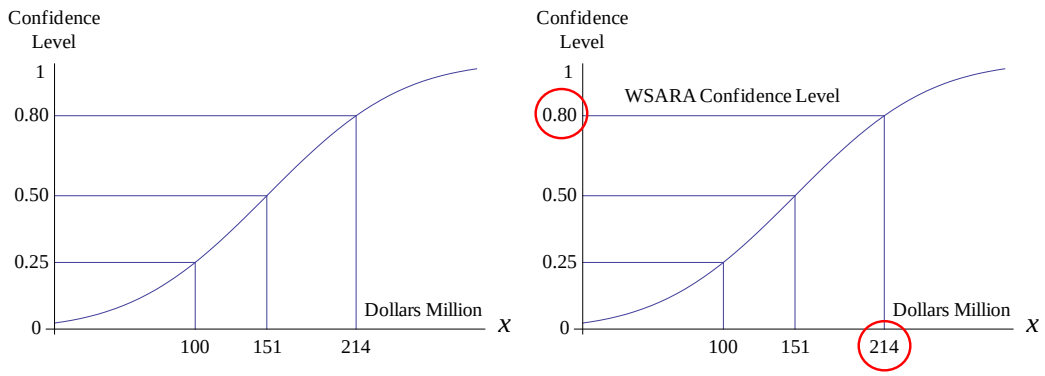

Figure 2. WSARA and Confidence Levels

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 cost analysts, which varies in the community. There are many subtleties in the underlying formalisms of Monte Carlo simulation. These must be understood to avoid errors in simulation design and in interpreting its outputs. For example, analysts must understand topics such as correlation and which of its many varieties is appropriate in cost uncertainty analysis. Analysts must understand 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 is an alternative to Monte Carlo simulation. Its straightforward algebraic equations ease the mathematical burden on analysts. 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 scenarios of events that, if realized, drive costs higher than planned.

Regardless of the approach, 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:

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, the analysis credibility is easily questioned. Details about the analysis methodology, including assumptions, are important to document. 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 by, not directed by, the methods discussed herein.

2.0 Scenario-Based Method (SBM)

The scenario-based method was developed along two implementation strategies. One is the non-statistical SBM. The other is the statistical SBM, which is the form needed for WSARA. The following 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 simulations if designed as arbitrary randomizations of possible program costs. This can lead to reserve recommendations absent of a clear program context for what these funds protect.

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

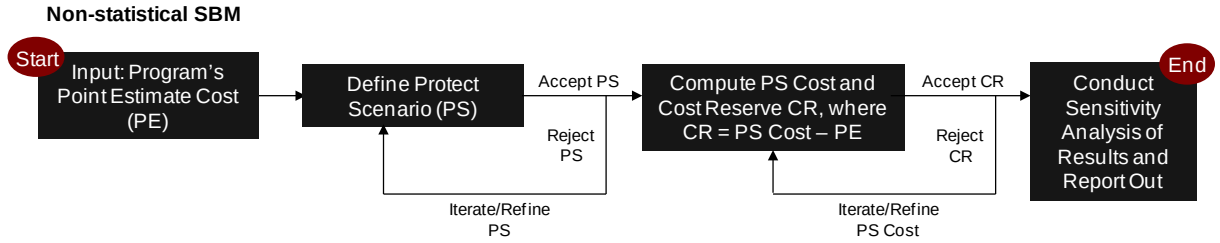

Figure 3. The Non-statistical SBM Process

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 summed 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 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 <u>not</u> being exceeded*. In practice, it is envisioned that management will converge on a protect scenario after deliberations on the one initially defined. This part of the process is to ensure 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 determined. 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. This may be an iterative process 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 with its case-basis not defined or carefully thought through. Scenario definition encourages a discourse on risks that otherwise might not be held. It allows risks to become fully visible, traceable, and estimative to program managers and decision-makers.

The non-statistical SBM does not produce confidence measures. The chance the protect scenario cost, or of any other defined risk scenario, 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 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.

There are many reasons to implement a statistical SBM. These include (1) it enables WSARA confidence measures to be determined (2) 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 (3) 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 of SBM. The upper part is the non-statistical SBM process steps. The lower part is the statistical SBM process steps. Thus, the statistical SBM is an augmentation to the non-statistical SBM.

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

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

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

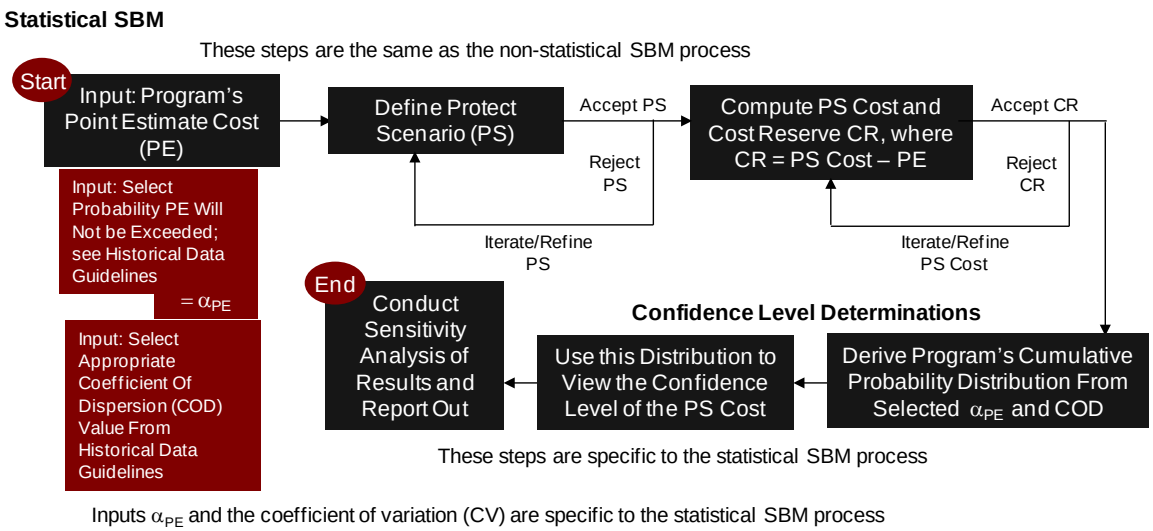

Figure 4. The Statistical SBM Process

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.

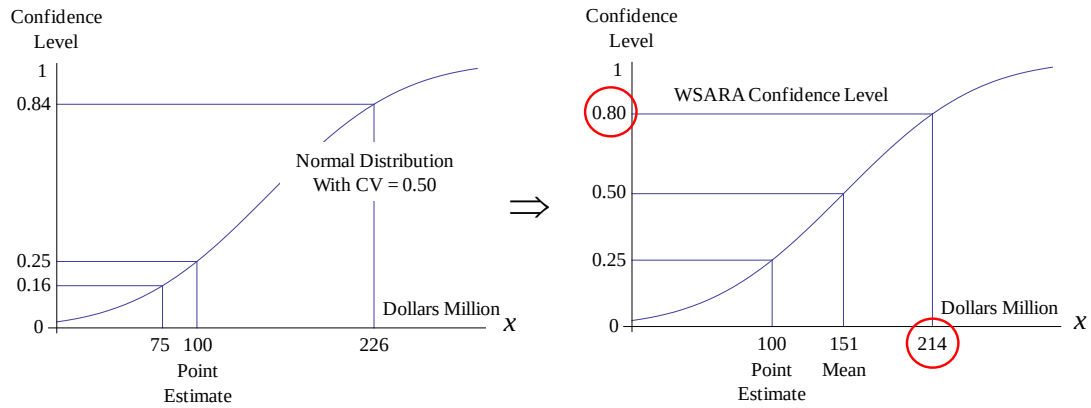

Figure 5. Statistical SBM Produces WSARA Confidence Levels

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].

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 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 A-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 protect scenario of the program 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 A-1 in Appendix A. In this case, $P(Z \le z_{PE} = -0.6745) = 0.25$; therefore, with $z_{PE} = -0.6745$ we have

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

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

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).

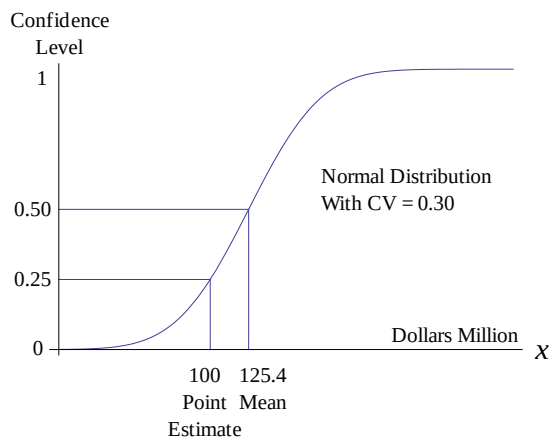

Figure 6. Probability Distribution Function of $Cost_{Pgm}$

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 A-1 in Appendix A we have $P(Z \le z_{PS} = 0.523) \approx 0.70$. 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 A-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.

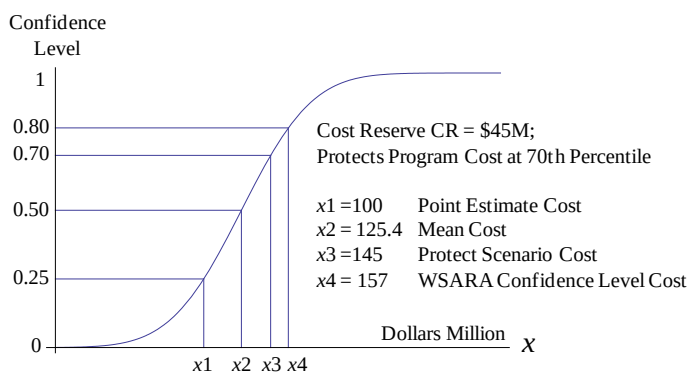

Figure 7. Example 1: Resultant Distribution Functions and Confidence Levels

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 protect scenario of the program 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.

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

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

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

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.

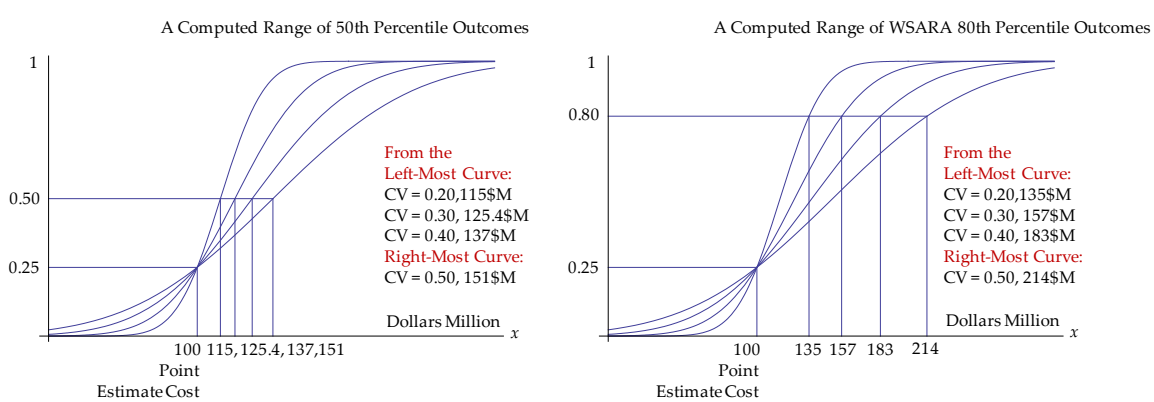

Figure 8. A Range of Confidence Level Cost Outcomes

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, 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, 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.

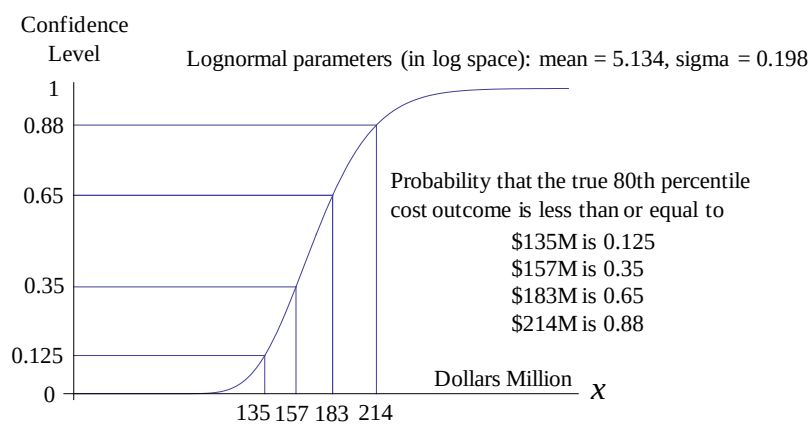

Figure 9. Measuring Confidence in
WSARA Confidence Levels: A Lognormal Statistical Fit

3.0 The Enhanced SBM (eSBM)

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). eSBM is an historical data-driven application of the statistical SBM, with heightened analytical features.

Two key inputs characterize the statistical SBM. They are the probability a program's point estimate cost will not be exceeded (α_{PE}) and 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. The following discusses and illustrates these features.

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 $P(Cost_{Pgm} \le x_{PE}) = \alpha_{PE}$. Anecdotal experience with α_{PE} is that a program's PE cost usually faces higher not lower probabilities of being exceeded. The interval $0.10 \le \alpha_{PE} \le 0.50$ expresses the view that a program's PE cost will very probably experience growth instead of reduction.

Recent research on coefficients of variation derived from historical cost data can provide insights into values for α_{PE}. From the 2010 Naval Center for Cost Analysis (NCCA) study, presented in Section 4, it can be derived that $\alpha_{PE} = 0.34$ for programs at Milestone B using the historical coefficient of variation equal to 0.51 (described in Table 2). From the 2007 RAND study, it can be derived that $\alpha_{PE} = 0.09$ for programs at Milestone B using their implied historical coefficient of variation equal to 0.26 [Younossi, Arena et al. pp. 16].

These derivations of α_{PE} stem from lognormal distributions, which RAND demonstrated fit well to their collection of historical cost growth data. A similar observation is made in the 2010 NCCA study.

Unless there are special circumstances, the above indicates the reasonableness of choosing a value for α_{PE} from the interval $0.10 \le \alpha_{PE} \le 0.50$. The selected value should be accompanied by a written justification for the choice. Finally, it is good analytic practice to conduct a sensitivity analysis on other possible α_{PE} values with the results documented.

A variation on the above for assessing α_{PE} is to compute its value from two other judged probabilities; specifically, α_1 and α_2 shown in Figure 10.

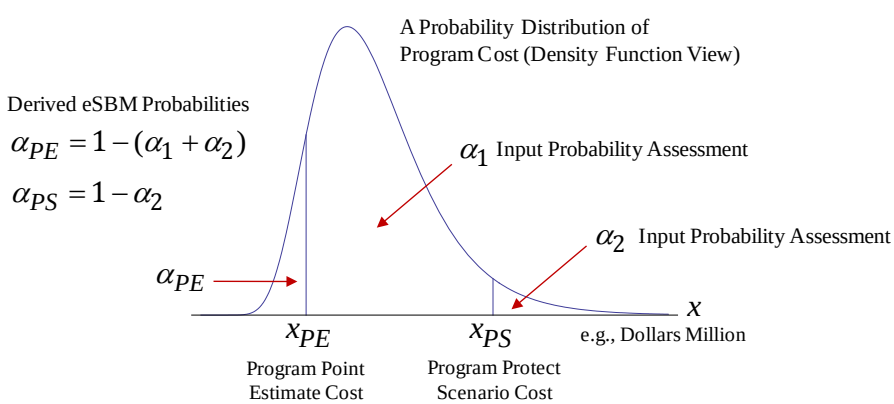

Figure 10. Determining eSBM Probabilities α_{PE} and α_{PS}

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.

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

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

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

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

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

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

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

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 A-1 in Appendix A

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

and

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

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.

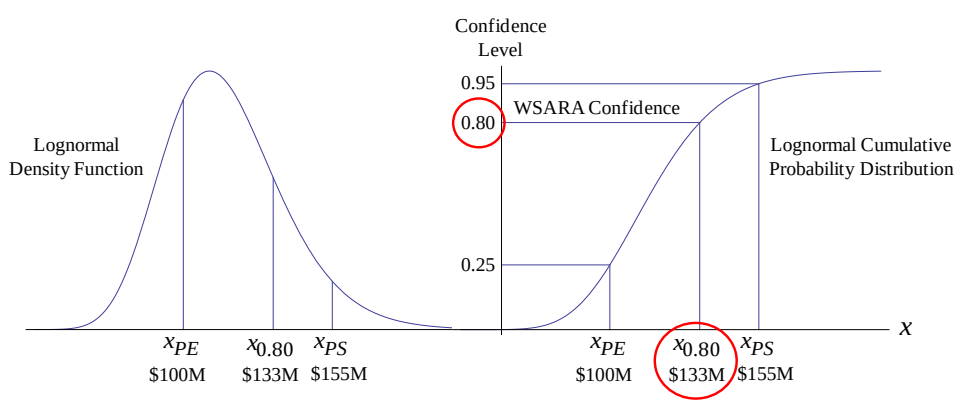

Figure 11. Example 3: Resultant Distribution Functions and Confidence Levels

4.0 NCCA Historical Cost Data Research

This section presents research findings on coefficients of variation derived from historical cost data collected across numerous Department of Defense programs. The study was conducted in 2010 by the Naval Center for Cost Analysis (NCCA). Its findings enable the integration of historical cost risk measures into the SBM algorithms. Doing so promotes realism in estimating future program costs, while offering a traceable and defensible basis behind data-derived measures of risk and cost estimate confidence. The NCCA research complements similar, and recently conducted, studies on defense system cost growth by RAND [Arena, Younossi, 2006, 2007].

To shed light on the behavior of cost distribution functions (S-curves) used in Department of Defense cost risk analyses, and to develop historical performance benchmarks, five conjectures on the behavior of coefficients of variation (CV) were proffered. These conjectures are as follows:

- **Conjecture 1**: Estimation Consistency
  - CVs in current cost estimates are consistent with those computed from acquisition histories.

- **Conjecture 2**: Decline During Acquisition
  - CVs decrease throughout the acquisition life cycle.

- **Conjecture 3**: Platform Homogeneity
  - CVs are equivalent for aircraft, ships, and other platform types.

- **Conjecture 4**: Adjustment Decline
  - CVs decrease when adjusted for changes in quantity and inflation.

- **Conjecture 5**: Secular Invariance
  - CVs are steady over the long run.

The first conjecture 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, system acquisitions. Note that consistency does not necessarily mean accuracy. Determining 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 recent OSD acquisition initiatives. However, we note that various acquisition reform initiatives have been present in the Department of Defense for the past many years with traits in the cost data collected for this study.

The second conjecture suggests that CVs should decrease monotonically throughout the acquisition life cycle as more information is acquired regarding the program in question. More information is known about a system's technical and performance characteristics at Milestone (MS) C than at MS A.

The third conjecture posits that CVs are neutral to platform type; that is, estimating costs for ships, aircraft, and space or ground systems all encounter similar challenges in cost growth variability. There is

no reason to believe, *a priori*, that CVs are platform-unique. All programs fall under the same acquisition management processes and policies. Furthermore, analysis tools and talent in the cost and acquisition management communities are likely distributed uniformly.

The fourth and fifth conjectures examine, respectively, hypotheses that CVs decrease when adjusted for changes in quantity and inflation and that CVs have not changed significantly over the long run.

The statistical acceptability of each conjecture was examined through a data collection and analysis effort. The following presents the foundations of the analysis and its findings.

4.1 Historical Cost Data

Selected Acquisition Reports (SARS) provided raw data on the cost outcomes of 100 mostly historical Department of the Navy (DON) major defense acquisition programs (MDAPs). These data also included a handful of on-going programs where cost growth has likely stabilized, such as LPD-17 program.

Over the year, numerous cost studies have indicated that SARs, while not perfect, are nevertheless a good and comprehensive official source of data on the cost, schedule, and technical performance of major defense acquisition programs. SARs are tied to milestones and present total program acquisition costs across multiple appropriations. For this research, data were culled from SAR Summary Sheets. These sheets present top-level numerical cost data.[7]

For a given program, the SAR provides two estimates of cost. The first is the baseline estimate (BE). This is usually made when the system nears a major milestone. The second is the current estimate (CE). This is based on best available information. It includes all known and anticipated revisions and changes to the program. For a completed acquisition program, the CE in its last SAR is regarded as the actual cost of the program. Costs in the SARs are shown in base-year and then-year dollars. This allows for dollar comparisons with and without the effects of inflation.

The ratio of a program's CE to BE is defined as its Cost Growth Factor (CGF). The computed CGFs for large samples of completed programs are the basis for estimating the *standard deviation* σ and the *mean* μ of acquisition program cost outcomes. Defined by Equation 2, a program's coefficient of variation (CV) is the ratio of these two statistics (σ / μ).

In practice, there is very strong evidence to support the hypothesis that the SAR BE is a cost estimate for modern era programs. 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, and the Independent Cost Estimate (ICE) conducted either by the Naval Center for Cost Analysis (NCCA) or the Office of the Secretary of Defense (OSD). Unfortunately, independent cost estimates and program office cost estimates from the 1970s and 1980s were generally unavailable. SARs in that era were supposed to reflect cost estimates in accordance with a SECDEF Decision Memorandum, an output of the Defense System Acquisition Review Council (DSARC). The degree of compliance with this guidance is unknown.

[7] SAR Summary Sheets are produced annually by the Undersecretary of Defense (Acquisition, Technology and Logistics).

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. There are several ways to adjust raw then-year (or base-year) dollars in the SARs, to reflect changes in quantity that occurred. Three of these methods are shown below.

- Adjust baseline estimate to reflect current quantities: CGF = CE/(BE + QΔ)
- Adjust current estimate to reflect baseline quantities: CGF = (CE – QΔ)/BE
- Fisher index = square root of the product of the first two (the geometric mean).

The first two quantity adjustment rules are analogous to the Paasche and Laspeyres price indices, which are based on current and base-year quantities, respectively. The Fisher index is used to compute the GDP price index. Applying it 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. The deltas in CVs are negligible regardless of the adjustment method used.[8]

4.2 Milestone B Sample Data

Given the 100 programs in the collected sample, 50 were Milestone (MS) B estimates of total program acquisition cost. These costs included development, production, and (to a small extent) military construction. Platform types included aircraft, helicopters, missiles, ships and submarines, torpedoes, and other systems. From the SAR summary sheets, the following data elements were captured: base-year, baseline type, platform type, baseline and current costs and quantity estimates, changes to date, date of last SAR. All costs collected are in base-year and then-year dollars. Results were analyzed with CVs displayed in Table 2.

Cost Growth Factors & CVs for All DON MDAPs at MS B for 1969 & Later; n = 50				
	(Without Qty 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 2. Milestone B: Cost Growth Factors and Coefficients of Variation

In Table 2, four CVs were derived for Milestone B acquisition programs. When adjustments for quantity and inflation were made, the expected decrease in the magnitude of each CV is seen.

Figure 12 shows CGFs adjusted for quantity and represented in then-year dollars. The skew of histogram suggests a lognormal distribution, with the mean falling to the right of the median. Noted in the statistical literature, CVs computed using product-moment formulae are subject to the influence of outliers. This

[8] 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.

influence is seen in Figure 12 due to a program known as Harpoon. Harpoon is the right-most datum in Figure 12. It has a CGF of 3.96, indicating almost 300 percent cost growth. Eliminating this observation from the Milestone B data decreases the CV from 51 percent to 45 percent.

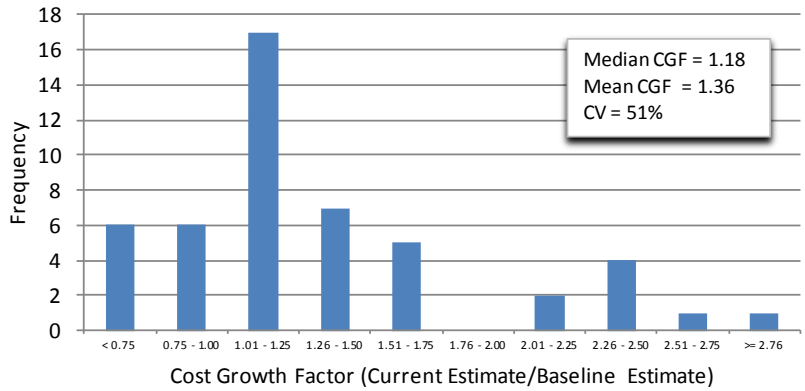

Figure 12. Milestone B: Acquisition Program Cost Growth Factors for All DON MDAPs

Figure 13 presents CVs derived by platform type. These findings are shown for the entire MS B dataset and by ships, aircraft, missiles, and electronics. The missiles group is also heavily influenced by the Harpoon program outlier. When this datum is eliminated, the quantity adjusted then-year dollar CV drops to 47 percent. This is remarkably close to the CV values for the other platform types.

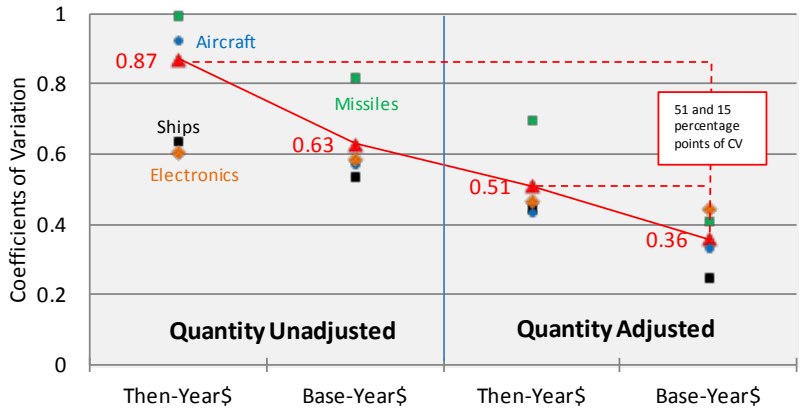

Figure 13. Milestone B: Quantity and Inflation Adjust CVs

Figure 14 shows the means and variances of CGFs from Milestone B data[9]. To shed light on the homogeneity of CVs (Conjecture 3) the null hypotheses of equal population means for platform type was

[9] 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 mean and standard deviation components of CV separately. 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 percent level of significance.

formulated versus the alternative of at least one pairwise difference. Similar hypotheses were formulated for the variance component of CVs. Statistical tests on the plausibility of these hypotheses strongly supports the conjecture of homogeneous CVs (quantity adjusted in then-year dollars) at Milestone B.

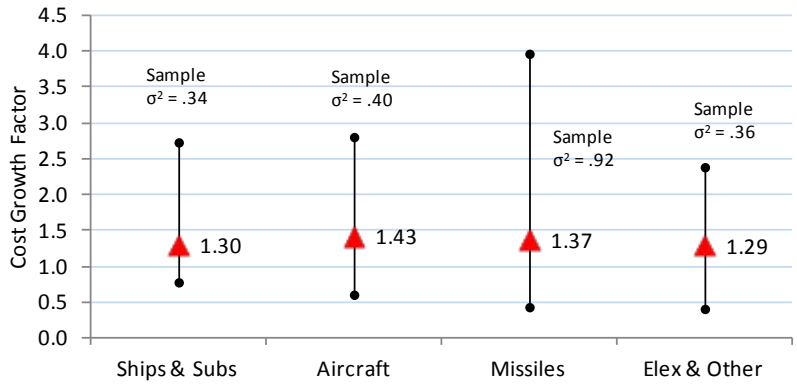

Figure 14. Milestone B: CGF Means and Variances (Quantity Adjusted in Then-Year Dollars)

4.3 Milestone B: Additional Findings

Shown in Figure 15, CVs significantly fall as components of the variation in costs are explained. The dataset of 50 observations 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 percent. Of this amount, *after* adjusting for changes in quantity, inflation represents a full 15 percentage points. That is a significant contribution.

This effect might be due to the volatility in average annual inflation rates during the administrations of Nixon/Ford (6.5 percent), Carter (10.7 percent), Reagan (4.0 percent), GHW Bush (3.9 percent), and Clinton (2.7 percent).[10] During the mid-1970s, OSD Comptroller (Plans and Systems) was promulgating inflation forecasts of 3 to 4 percent per annum. This was received from the Office of Management and Budget (OMB); however, inflation in the general economy was rising to over 10 percent per annum during the peak inflation period of 1978 to 1981.

This disconnect caused tremendous churn in defense acquisition programs. Defense acquisition programs used OSD rates in estimating then-year dollar total obligation authority (TOA). Double-digit inflation reality was not expected when creating the defense budget.[11] 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 and may have caused additional perturbations.

Figure 15 shows a pronounced decline in Milestone B CVs over lengthy time-periods. These findings reflect secular trends in CVs from Milestone B acquisition programs.

[10] Average annual rate of inflation during a US presidency, measured by the consumer price index.
[11] Because of long profiles for expenditures of TOA (seven years for ships, for example), an upward tick 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.

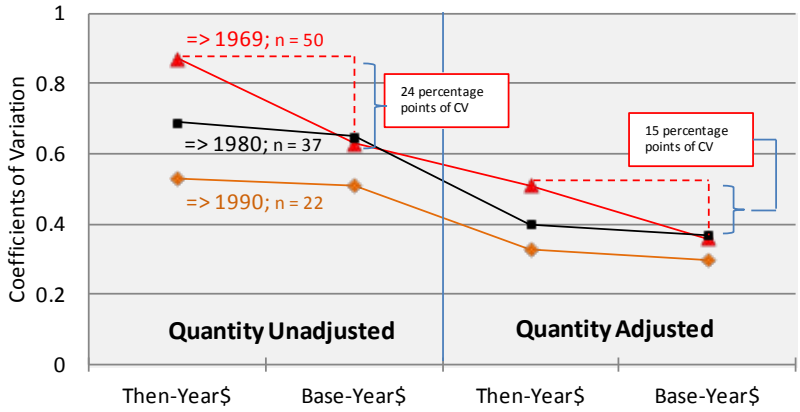

Figure 15. Milestone B: Declines in CV over Long Time-Periods

The historical data revealed inflation affected the magnitude of CVs much less in the 1980s and 1990s than in the 1970s. This is likely due to less volatility in rates and a secular decline in their values. It is unclear if the current trend of price stability will continue over the years ahead for our current acquisition programs. With $14 trillion in direct national debt, there is legitimate cause for concern.

4.4 Milestone C Sample Data

Given the 100 programs in the collected sample, 43 were Milestone (MS) C estimates of total program acquisition cost. Table 3 summarizes the four CVs derived for Milestone C acquisition programs, with adjustments for quantity and inflation.

Cost Growth Factors & CVs for All DON MDAPs at MS C for 1969 & Later; n = 43				
	(Without Qty 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 3. Milestone C: Cost Growth Factors and Coefficients of Variation

Figure 16 shows Milestone C CGFs adjusted for quantity and represented in then-year dollars. The skew of the histogram in Figure 16 is noticeably less than the histogram shown in Figure 12 for Milestone B, with the mean CGF value only slightly higher than the median CGF value.

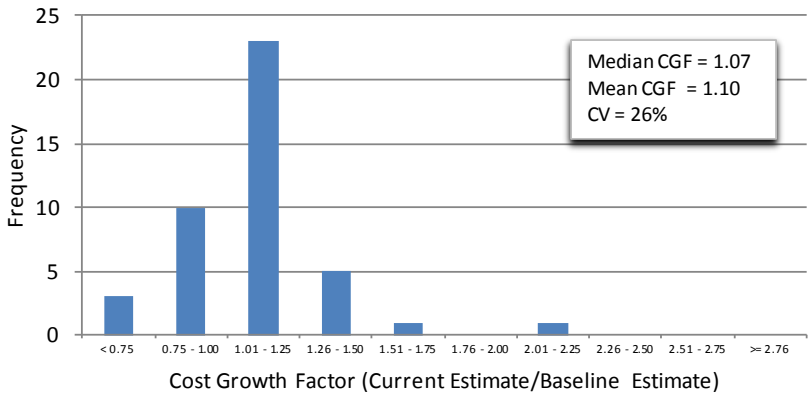

Figure 16. Milestone C: Acquisition Program Cost Growth Factors for All DON MDAPs

Figure 17 shows a significant decline in the magnitude of CVs as programs move from acquisition Milestone B to acquisition Milestone C. This reflects not only the inclusion of sunken development costs but the likely benefits of reduced technical and programmatic uncertainties as programs mature.

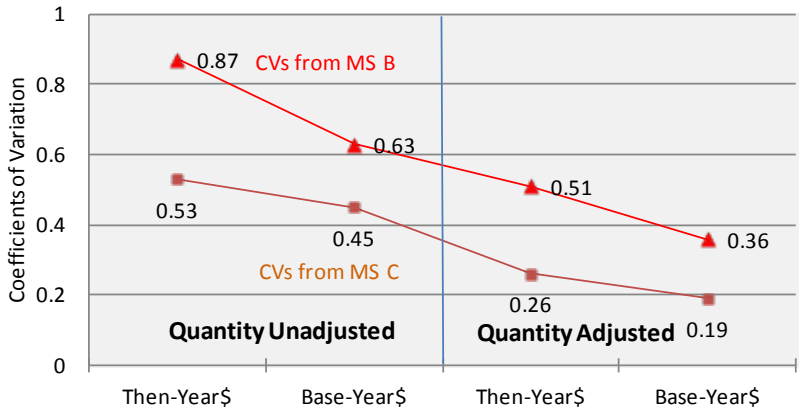

Figure 17. Reduction in Uncertainties across Milestones Expressed by Declining CVs

Figure 18 shows the means and variances of CGFs from Milestone C data.[12] Similar to the Milestone B findings in Figure 14, Milestone C CGFs were remarkably close across the platform types.

In Figure 18, the wide variance in CGFs for aircraft platforms was entirely driven by the EA-6B program outlier. This program has a CGF of 2.25, indicating 125 percent cost growth. Eliminating this datum reduces the aircraft CV (quantity adjusted in then-year dollars) from 36 percent to 22 percent. This value is in line with that of ships and submarines (22 percent) and "other" (16 percent). Even in the presence of the outlier, the null hypothesis of constant CGF population means is not rejected at the 5 percent level of significance. The results were mixed for the null hypothesis of constant population variances. Levene's

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

test supports the null hypothesis, whereas pairwise F-tests rejected it in cases involving the outlier. From this, moderately strong support is indicated for the conjecture of homogeneous CVs at Milestone C.

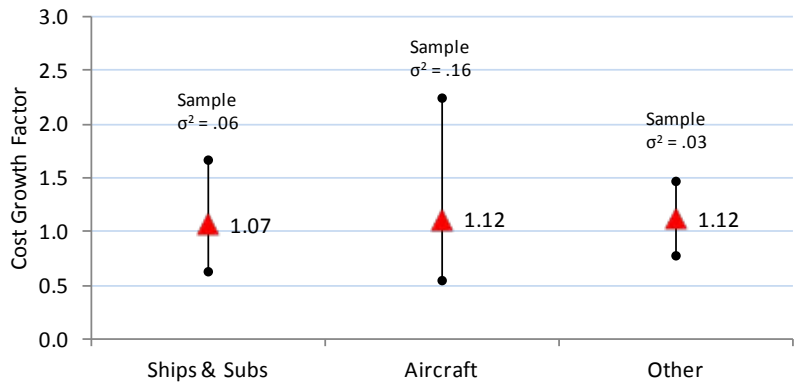

Figure 18. Milestone C: CGF Means and Variances (Quantity Adjusted in Then-Year Dollars)

Figure 19 shows a pronounced decline in Milestone C CVs over lengthy time-periods similar to the trend seen in Milestone B. 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.

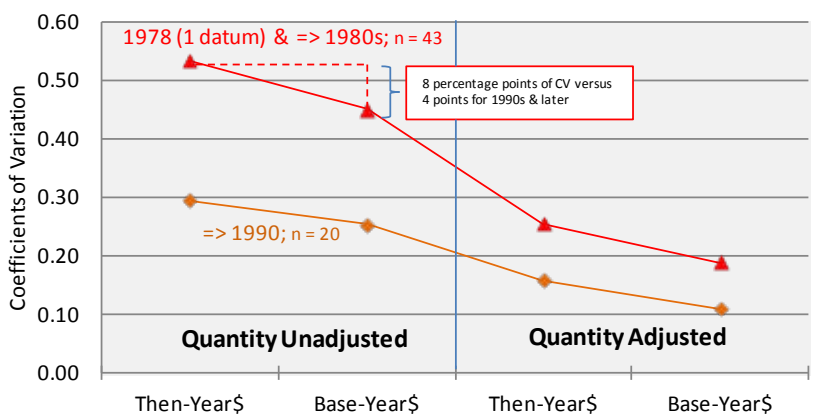

Figure 19. Milestone C: Declines in CV over Long Time-Periods

4.5 Milestone A Sample Data

Given the 100 programs in the collected sample, only 7 were Milestone (MS) A estimates of total program acquisition cost. This sample is too small to make statistically plausible inferences. Estimation by analogy seems a logical alternative. Suppose the degree of risk and uncertainty is the same between Milestones A and B as it is between Milestones B and C. If so, then the application of roughly 15 percentage points of additional CV seems appropriate at Milestone A. This is shown in Figure 20.

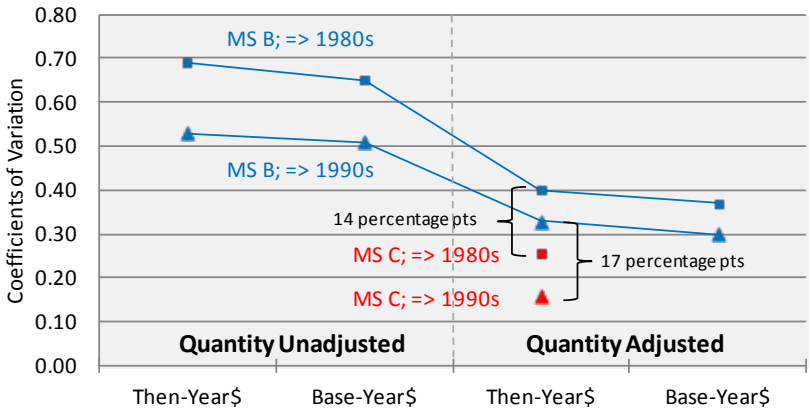

Figure 20. Posited Deltas in CVs from Milestone B to Milestone C

4.6 CV Selection for Cost Uncertainty Analysis

Figure 21 shows benchmark CVs by acquisition milestone. The upper-band reflects findings from the complete dataset of acquisition programs. The lower-band reflects findings from the subset of programs from the 1990s and later.

Selecting a CV value within the band depends on circumstances specific to a given acquisition. It also depends on organizational views on issues such as risk tolerance, the chances of significant volatility in out-year inflation rates, and the effects of current acquisition reform initiatives. When choosing a value, note that lower rather than higher values of CVs have been the norm in the defense cost community.

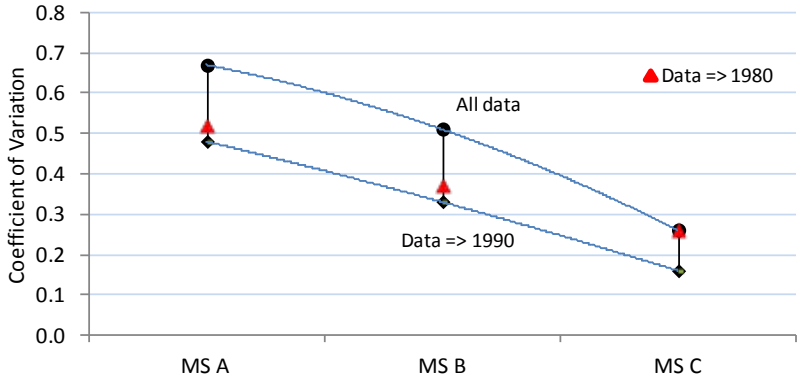

Figure 21 Benchmark CVs by Acquisition Milestone – Quantity Adjusted in Then-Year Dollars

4.7 Historical Analysis Summary

Section 4 began with statements of five conjectures. The following summarizes the statistical findings associated with each.

- **Conjecture 1**: Estimation Consistency
  - CVs in current cost estimates are consistent with those computed from acquisition histories.

  - **Finding**: Ad hoc observation suggests a pervasive underestimation of CVs in the international defense community.

- **Conjecture 2**: Decline During Acquisition
  - CVs decrease throughout the acquisition life cycle.
  - **Finding**: Strongly supported.

- **Conjecture 3**: Platform Homogeneity
  - CVs are equivalent for aircraft, ships, and other platform types.
  - **Finding**: Strongly supported, especially for Milestone B

- **Conjecture 4**: Adjustment Decline
  - CVs decrease when adjusted for changes in quantity and inflation.
    Finding: Strongly supported.

- **Conjecture 5**: Secular Invariance
  - CVs are steady over the long run.
    Finding: Strongly rejected.

In addition, the following is recommended.

- 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. The use of single values presumes a level of knowledge and degree of certainty that simply does not exist. A view of future price stability would argue for the use of lower CVs and instability for higher CVs. A belief in the positive effect of structural change due to recent acquisition reform initiatives would argue for lower CVs.

- Exercise prudence in choosing CV benchmarks. Better to err on the side of caution until costs on completed acquisition programs clearly demonstrate lower CGFs and CVs.

- Choose the high-end of benchmark CV bounds established at Milestone A to support Analysis of Alternatives 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 these for both Monte Carlo simulation and eSBM. Base these on confidence intervals for the CVs.

To support the use of these results in eSBM or to validate a Monte Carlo simulation, Figure 22 and Table 4 show benchmark CV values and bands for estimates of total program acquisition costs.

Each milestone contains four bands corresponding to adjustments for changes in quantity and inflation. Employ the benchmark values according to whether these two terms are regarded as exogenous or random in generating an S-curve

- Quantity. If stochastic, use the boxed values; otherwise do not.
- Inflation. If stochastic, use the bands labeled TY$ (then-year dollars); otherwise use the bands labeled BY$ (base-year or constant dollars).

Since CVs change over time, each band displays three values. These correspond to the sample used to generate them. Generally, higher values reflect findings from the complete dataset of acquisition programs with acquisition baselines dating from the late 1960s. Lower values reflect findings from the subset of programs from the 1990s and later.

The connecting lines in Figure 22 are visual aids. Any estimated CV that falls within the high-low range should be regarded as reasonable. The choice of a specific value within a band to use for eSBM will depend upon an organization's view of structural change and perhaps on the unique characteristics of the acquisition program. The more stochastic and undefined the program, the higher should be the CV.

Finally, the benchmarks are an *aid* to analysis but not a substitute for an in-depth understanding of a program's risk and uncertainty. Solid application of theory, for both eSBM and Monte Carlo simulations requires thorough investigations of underlying acquisition assumptions, difficulties, and challenges if strong results are to be achieved and successfully defended to decision-makers.

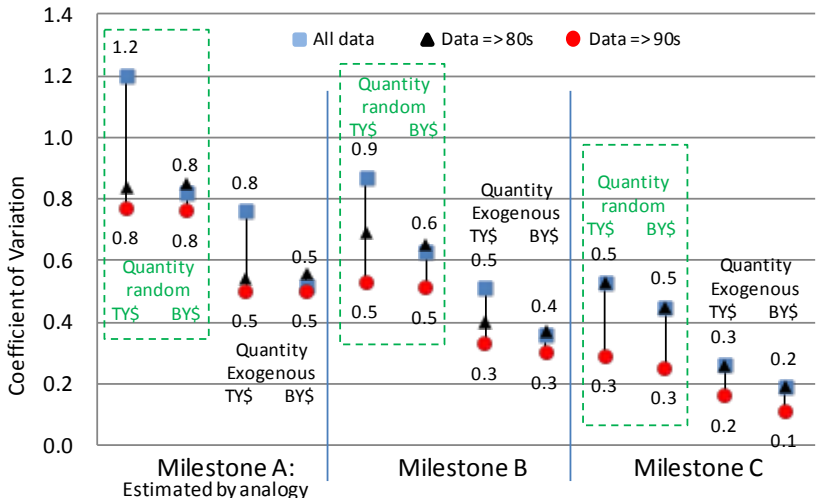

Figure 22. Estimated CV Bands by Acquisition Milestone

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

Table 4. Tabular Data for Figure 22

5.0 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. This paper presented 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.

Features of eSBM include the following:

- The ability to establish a range of possible program costs and measures of cost estimate confidence, driven by past program performance;
- A focus 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;
- 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 decision-makers; this is a more meaningful topic to focus on instead of statistical abstractions simulation approaches can sometimes create;
- The ability to generate measures of cost estimate confidence in a spreadsheet environment;
- Does not require analysts develop probability distribution functions for all the uncertain variables in a WBS, which can be time-consuming and hard to justify;
- Correlation is captured implicitly in the analysis by the chosen coefficient of variation – thereby eliminating the need to address this measure explicitly.
- The approach fully supports traceability and focuses attention on key program risk events, identified 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 based on current concerns and past program cost histories. It requires the development of written scenarios that represent the program's "risk story" rather than debating 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 analysis and discussions 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 ~ 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 ~ 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

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

REFERENCES AND RELEVANT LITERATURE

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

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

Garvey, P. R., 2000. *Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective*, Chapman-Hall/CRC Press, Taylor & Francis Group, Boca Raton, New York, London.

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

Garvey, P. R., 2008. *Analytical Methods for Risk Management: A Systems Engineering Perspective*, Chapman-Hall/CRC Press, Taylor & Francis Group, Boca Raton, New York, London.

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

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

NASA *Cost Estimating Handbook*, 2008.

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

Younossi, O., et al. 2007. *Is Weapon System Cost Growth Increasing*?: *A Quantitative Assessment of Completed and Ongoing Programs*, The RAND Corporation.

USAF Air Force Cost Analysis Agency, *Cost Risk and Uncertainty Analysis Handbook*, 2007.