Testing S-Curves for Reasonableness: The NCCA S-Curve Tool

SCEA/ISPA Conference, Wednesday, June 8th, 2011

Richard L. Coleman (retired)

Peter J. Braxton, Richard C. Lee, Technomics, Inc.

Sponsored by Brian Flynn, Naval Center for Cost Analysis (NCCA)

Abstract

The requirement of the Weapon Systems Acquisition Reform Act of 2009 to disclose "the confidence level used in establishing a cost estimate for a major defense acquisition program or major automated information system program…" and the austerity of current Department of Defense (DoD) budgets has brought about an increased interest in risk analysis and widespread use of the S-curve (cumulative distribution function) of cost. As this interest intensified, experience in the use of S-curves widened.

As is becoming clear with experience so far gained, there is a natural evolution from naïve trust in S-curves; to the realization that not all S-curves are created equal; to the understanding that one of the quickest and surest ways to detect a suspicious S-curve is the coefficient of variation (CV); to the temptation for (if not the actuality of) gaming of CVs; to a mature practice of risk. The authors were centrally placed during the development and evolution of a company-wide risk analysis of the very sort that DoD is now undergoing. They sat on the Risk Working Group that determined policy and process from inception to completion; led implementation of accompanying tools in three different sectors of the company; and served on the committee that approved (and disapproved) the various candidate tools for use. In short, they were intimately involved in the life cycle of the risk evolution within the company, from the earliest stages until the practice had become mature. The briefing will outline the steps (and some missteps) that the Risk Working Group went through and apply the lessons learned to the current situation in the DoD.

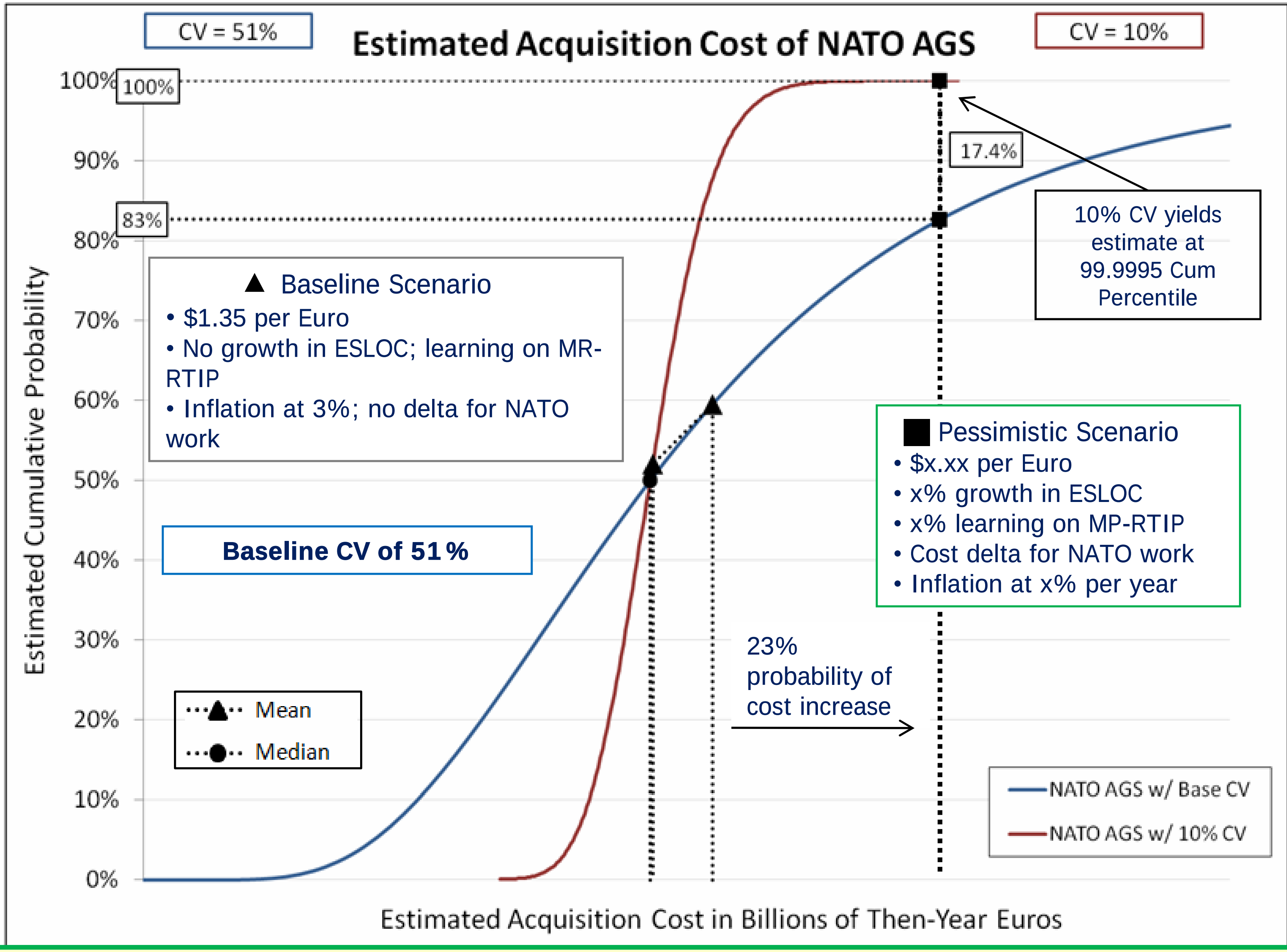
Finally, a robust Excel tool has been developed for practitioners to display the S-curve as developed by the estimator and to compare it to a historically-based, commodity-specific, phase-appropriate S-curve. It can be used to benchmark estimates, to compare current and prior estimates, and to reconcile between two estimates, with a variety of historically-based adjustments to either or both. This tool will guide the practitioner in judging the S-curve and will produce output intended for presentation to decision makers. This briefing will describe and demonstrate the tool.

The Problem

- Growing realization in defense cost community that commonly estimated S-curves are sometimes too narrow and risk analysis is incomplete
  - OSD CAPE, and others, cite cases where actual acquisition costs fall at the 99^{th}+ percentile
    - For MDAPS
    - On S-curves estimated years previously
  - Anecdotal evidence that CV estimates greater than 10% difficult to achieve, in too many cases
  - Authors have seen values of under 10% as MS A, and values of ½ of 1% at roughly half way through production
- Lack of definition of CVs
  - Quantity and inflation as exogenous or random
- Inconsistency in CV estimation between and within organizations
- Guidelines on risk analysis
  - NCCA leading a DON cost-community effort
  - CV Tool and benchmark values will contribute to solution

Inaccurately steep S-curves can lead to an underestimation of the mean, misallocation of scare defense resources, and failure to understand program risk

NATO AGS Example



Previous case study graphic – We'll demo its generation in the S-Curve Tool shortly

NCCA S-Curve Tool Outline

- S-Curve Tool overview
  - SAR data analysis and NATO AGS case study presented by Dr. Flynn at DoDCAS
- Industry Risk Implementation Case Study
  - Focus on incorporation of Lessons Learned in S-Curve Tool
- Motivation for S-Curve Tool
- Evolution of S-Curve Tool
- Guided Tour of S-Curve Tool
  - Screen shots
- Demonstration of S-Curve Tool
  - NATO AGS, OEM examples

Additional detail on data analysis, risk methodologies, and properties of probability distributions at subsequent sessions tomorrow (Th 09 Jun)!

Warning: There is a lot of ground to cover, so please forgive us if we put the spur to the horse (including ourselves)!

Objective of S-Curve Tool

- Leverage Dr. Flynn's historical analysis of SARs for DoN programs to develop a tool that will allow practitioners to easily and clearly:
  - Compare their estimate (S-curve!) to history in coefficient of variation (CV) and cost growth factor (CGF) [Benchmarking]
  - Compare two different estimates [Reconciliation]
  - Generate graphics for decision briefs
- Compatible with both:
  - Empirical methods such as Monte Carlo risk analyses
  - Parametric methods such as enhanced Scenario-Based Method (eSBM)
- The development team leveraged experience in CV analysis at their last job in industry

S-Curve Tool Status

- Beta version 1.0 internal to NCCA – February, 2011
- Beta version 2.0 general release – April, 2011
  - Macro-less Excel 2007 file (.xlsx)
  - Based on thorough vetting of breadboard / brassboard model
  - Accompanying detailed documentation
- Contains historical adjustment factors (CV and CGF) for [MS B & C], Acq, [w & w/o Qty and Inflation] for:
  - All DoN
  - Ships and Submarines
  - Aircraft
  - Missiles
  - Electronics / Other
- MS A factors and Development/Production splits developed by analogy

S-Curve Tool Access

- "How do I learn more?"
  - Discuss with one or more of the tool POCs:
    - Mr. Richard Lee, RLee@technomics.net, lead developer
    - Mr. Peter Braxton, PBraxton@technomics.net, algorithm development
    - Dr. Brian Flynn, Brian.Flynn@navy.mil, sponsor, historical data analysis
- "How do I get access to the tool?"
  - Contact one of the tool POCs
  - Visit the NCCA website http://www.ncca.navy.mil
    - S-Curve Tool and related materials in Tools section http://www.ncca.navy.mil/tools/tools.cfm
- "How can I make the tool better?"
  - Participate as a beta tester
  - Ongoing development to incorporate user feedback

Evolution of Risk at an Original Equipment Manufacturer (OEM)
A Case Study

"Progress, far from consisting in change, depends on retentiveness. When change is absolute there remains no being to improve and no direction is set for possible improvement: and when experience is not retained, as among savages, infancy is perpetual. **Those who cannot remember the past are condemned to repeat it.**"

-George Santayana, *The Life of Reason*

http://en.wikiquote.org/wiki/Santayana#Vol._I.2C_Reason_in_Common_Sense

Evolution of Risk at an OEM – First Stage (2003)

- Cost and Risk Evaluated by a single 3-person team from outside the sector
  - Short on cost people
  - Mostly from the company's original commodity, even if very different
- Only a few "major" risks were evaluated, tending to be more on the business side than the cost estimating side, but with some cost issues, e.g.:
  - RFP called for multiple FPI ships, but:
    - First hull EAC understated
    - Second hull more understated yet, causing an overly steep learning curve, compounding the first-EAC problem for follow ships
- S-curves rare for ICE Teams
  - Quantitative self evaluation for Proposal Teams almost non-existent

Evolution of Risk at an OEM – The Overhaul

- ICE/Risk Working Groups (I/RWGs) convened over 3 years
  - 1st I/RWG (2003-4)
    - Dictated increased cost involvement
    - Identified pockets of excellence in cost and risk
    - A sub-group investigated various practices and recommended Monte Carlo
  - 2nd I/RWG (2006)
    - Directed the use of S-curves
    - Recommend Monte Carlo
  - 3rd I/RWG (2007)
    - Ruled out the Method of Moments (as implemented, due to symmetry in results) and specification of the confidence level of inputs (abused by BOE authors)
    - Specified formats and processes
- I/RWGs progressively overcame the related inertias of legacy company and commodity (sometimes relating to customer)

Evolution of Risk at an OEM – Sixth Stage (2007)

- As cross-sector ICE teams became the norm, techniques matured through sharing of best practices
- Final major improvement came from observation that CVs were too narrow; cost estimating risk (statistical variability) was being omitted
  - ICE teams didn't have time for WBS-level statistical or expert-opinion methods for CE Risk
  - Most sectors used build-up and analogy, which don't lend themselves to CE Risk buildup
    - Two sectors did built-up WBS-level during in-stride support for proposals
  - Solution for 4 sectors was to inject a top-level CE Risk based on phase[1]
    - The top-level injection was symmetrical about the point estimate. Bias was captured by pluses and minuses derived from detailed BOE reading

1. "Analysis and Implementation of Cost Estimating Risk in the Ballistic Missile Defense Organization (BMDO) Risk Model, A Study of Distribution," J.R. Summerville, H.F. Chelson, R.L. Coleman, D.M. Snead, ISPA/SCEA, 1999.

Evolution of Risk at an OEM – The End Stage (2009)

- Methods and reporting formats standardized
- Training was developed and conducted on S-curves and risk for BOE authors, Pricing, Contracts and decision makers
  - OJT was used heavily for the Proposal Team and ICE risk analysts
- As procedures, methods, policies, and displays standardized, belief emerged that processes and standards had eliminated errors, but considerable variances between deltas found by ICEs persisted
- In the end, it came down to the rigor and independence of the ICE teams, which should not have been anything new, but seemed to be a surprise
  - Many parties *never did* believe this
  - Although everyone knows that architects and engineers have a profound effect on buildings and bridges, nobody seems to believe that the cost estimators and risk analysts have much effect on cost and risk estimates … *“it’s all just pull-down menus and radio buttons”*

Evolution of Risk at an OEM – Methods

- Cost methods:
  - Four sectors used a composite of build-up, analogy, and parametric
  - One sector used a composite of build-up and analogy
  - Two sectors used analogy
- Risk Methods:
  - Schedule & Technical Risk: All sectors used the Risk Cube method with a mix of expert-opinion-based and analogy estimates of consequences
    - Probabilities mostly opinion-based, but some were historically derived
    - Probabilities were mostly Bernoulli or categorical with a smattering of triangulars and normals (one sector used predominantly triangulars)
  - Cost Estimating Risk:
    - Two sectors used built-up (WBS-level) CE Risk during in-stride support
    - Four sectors used top-level cost estimating risk by phase
      - Three of these used BOE examination to account for non-zero bias in CE Risk and to locate specific weaknesses

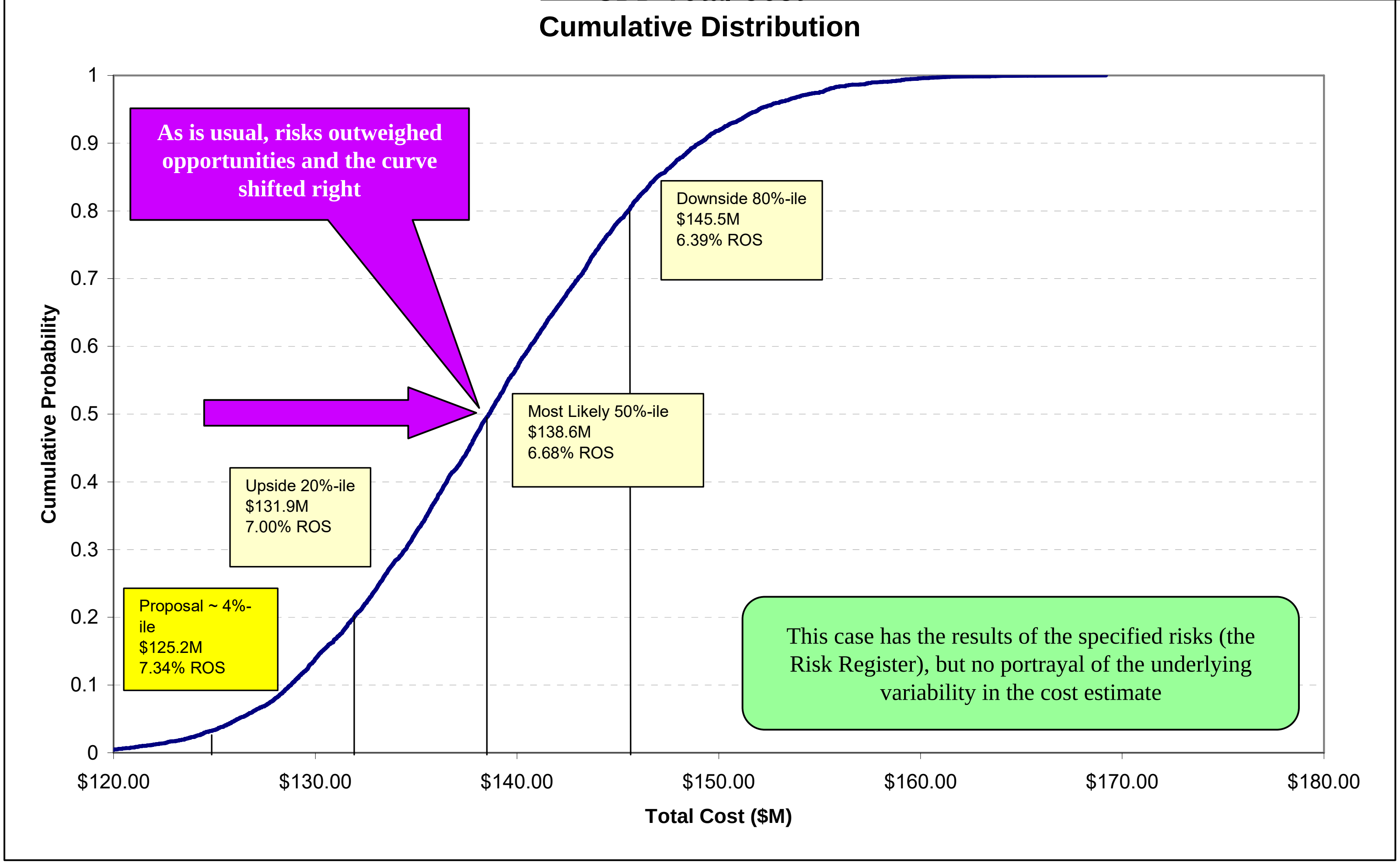
Insufficient CV

An ICE with and without Cost Estimating Variance
(a.k.a. Cost Estimating Uncertainty or Cost Estimating Risk)

SDD (Risks & Opportunities Only)

± 4.9%

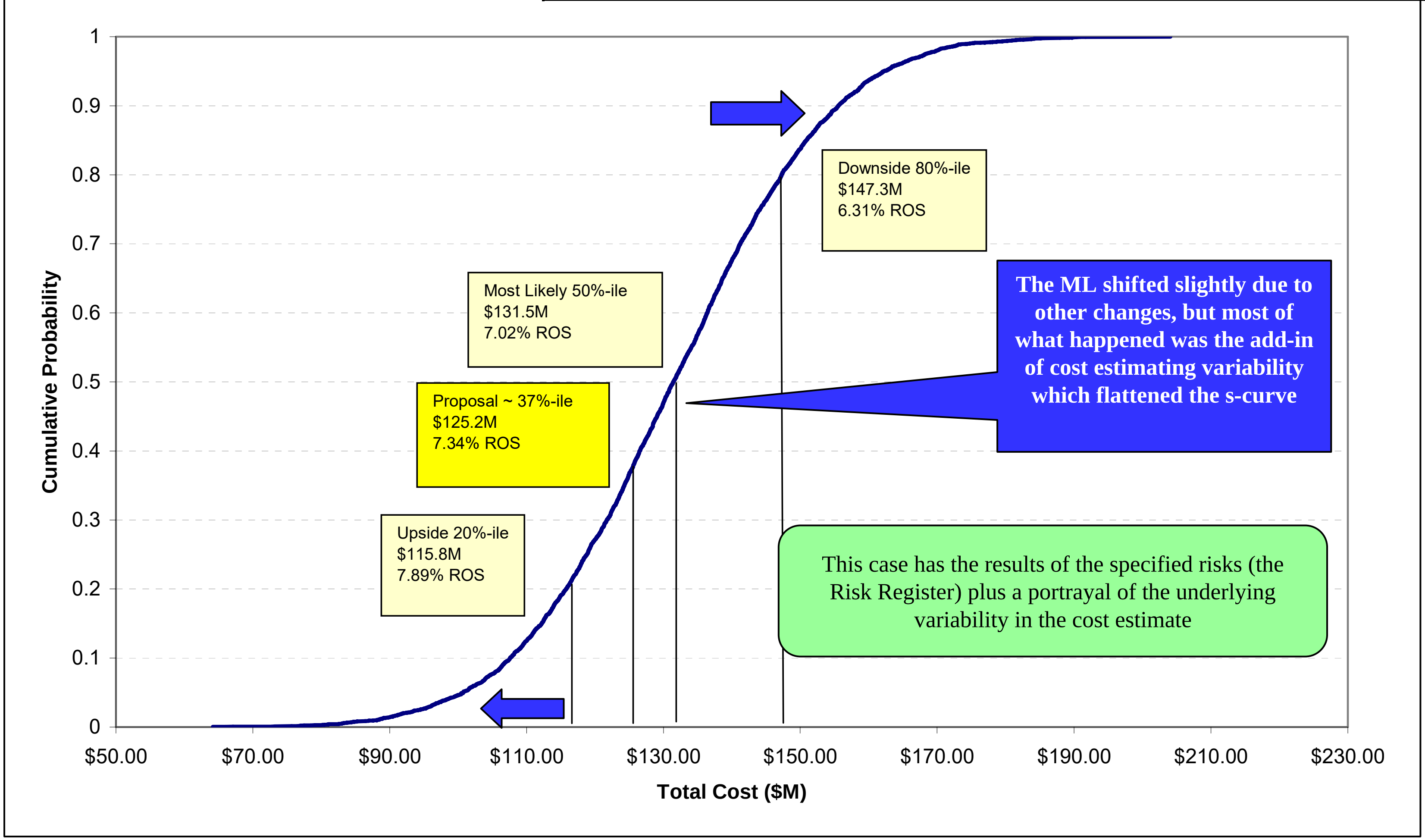
	Proposal	Upside	Most Likely	Downside
Cost	$ 125,219,843	$ 131,924,597	$ 138,603,127	$ 145,505,394
Fee	$ 9,924,967	$ 9,924,967	$ 9,924,967	$ 9,924,967
Price	$ 135,144,810	$ 141,849,565	$ 148,528,094	$ 155,430,362
ROS	**7.34%**	**7.00%**	**6.68%**	**6.39%**
		CPFF Contracts		



SDD (Cost Estimating Variance Added)

± 12.0%

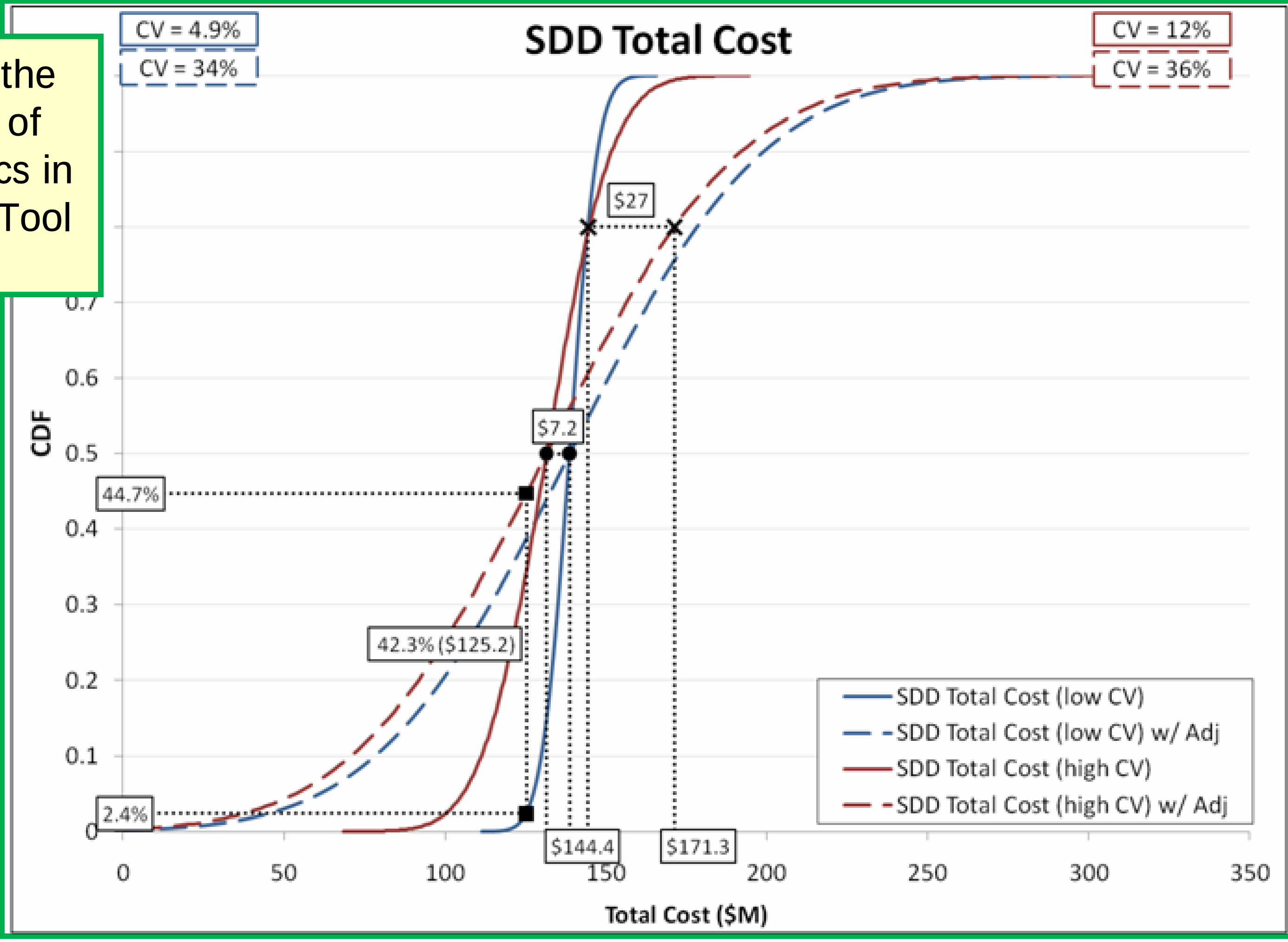
	Proposal	Upside	Most Likely	Downside
Cost	$ 125,219,843	$ 115,792,733	$ 131,455,455	$ 147,279,388
Fee	$ 9,924,967	$ 9,924,967	$ 9,924,967	$ 9,924,967
Price	$ 135,144,810	$ 125,717,700	$ 141,380,422	$ 157,204,355
ROS	**7.34%**	**7.89%**	**7.02%**	**6.31%**
		CPFF Contracts		



The Graphic Nobody Saw (or Wanted to See!)

ESTIMATE 1		ESTIMATE 2 (Historical Adjustment)	
ESTIMATE 1 TITLE	SDD Total Cost (low CV)	ESTIMATE 2 TITLE	SDD Total Cost (high CV)
Mean	$138.7	Mean	$131.5
20th Percentile	$132.95,20%	20th Percentile	$118.22,20%
80th Percentile	$144.39,80%	80th Percentile	$144.78,80%
Proposal Estimate	$125.2,2%	Proposal Estimate	$125.2,34%

We'll demo the generation of these graphics in the S-Curve Tool shortly



The Seductive (Sibilant!) S-Curve

- **Smoothness**:
  - Empirical CDFs are **smooth**, even for 100 trials of a Monte Carlo, conveying a false sense of precision
  - Corresponding empirical PDFs show noise and thus convey our uncertainty
- **Scale, Steepness, and Spread**:
  - Because the y-axis of an S-curve is always 0 to 1 (cumulative probability) and the x-axis usually auto-scales in Excel, it's hard to get a sense of the **scale** and corresponding **steepness** (or conversely **spread**) of an S-curve
  - A practiced reviewer will look at the x-axis and do some quick mental math, but S-curves should *always* be labeled with their CV to easily convey scale and to lift the burden of caution from the reviewer
- **Similarity**: Because of the previous two bullets, S-curves tend to look the **same**, which means we need to be cautious in viewing them
- **Comparisons (Serial vs. Side-by-Side)**:
  - Meaningful comparisons between S-curves need to be done on the same graph; it is too hard to detect differences flipping from one chart to the next, especially given the previous three bullets
  - By contrast, if a *change in a series, or variability*, is being illustrated, then serial display[1] is fine
- **Basis**: Along with the above concerns, there is the question of the underpinnings of risk analysis that the S-curve conveys: the S-curve is the **sausage**, and the risk analysis is the **sausage factory**; we cannot be confident in the former without fully delving into the latter

1. Serial displays, called "small multiples", are strongly advocated by Edward Tufte, an expert in the presentation of informational graphics

Comparison of Metrics for Variability

Metric/Requirement	Shows Asymmetry?	Good for Comparisons?	Difficulty of Computation	Utility for Percentiles
Standard Deviation (SD)	N	N	L	L
CV (SD/mean)	N	Y	L	L
20th /50th /80th Percentiles	Y	N	M	M
Three-Point Estimate (L/ML/H)	Y	N	L	L
PDF	Y	N	H	M
S-curve (CDF)	Y/N	N	H	H

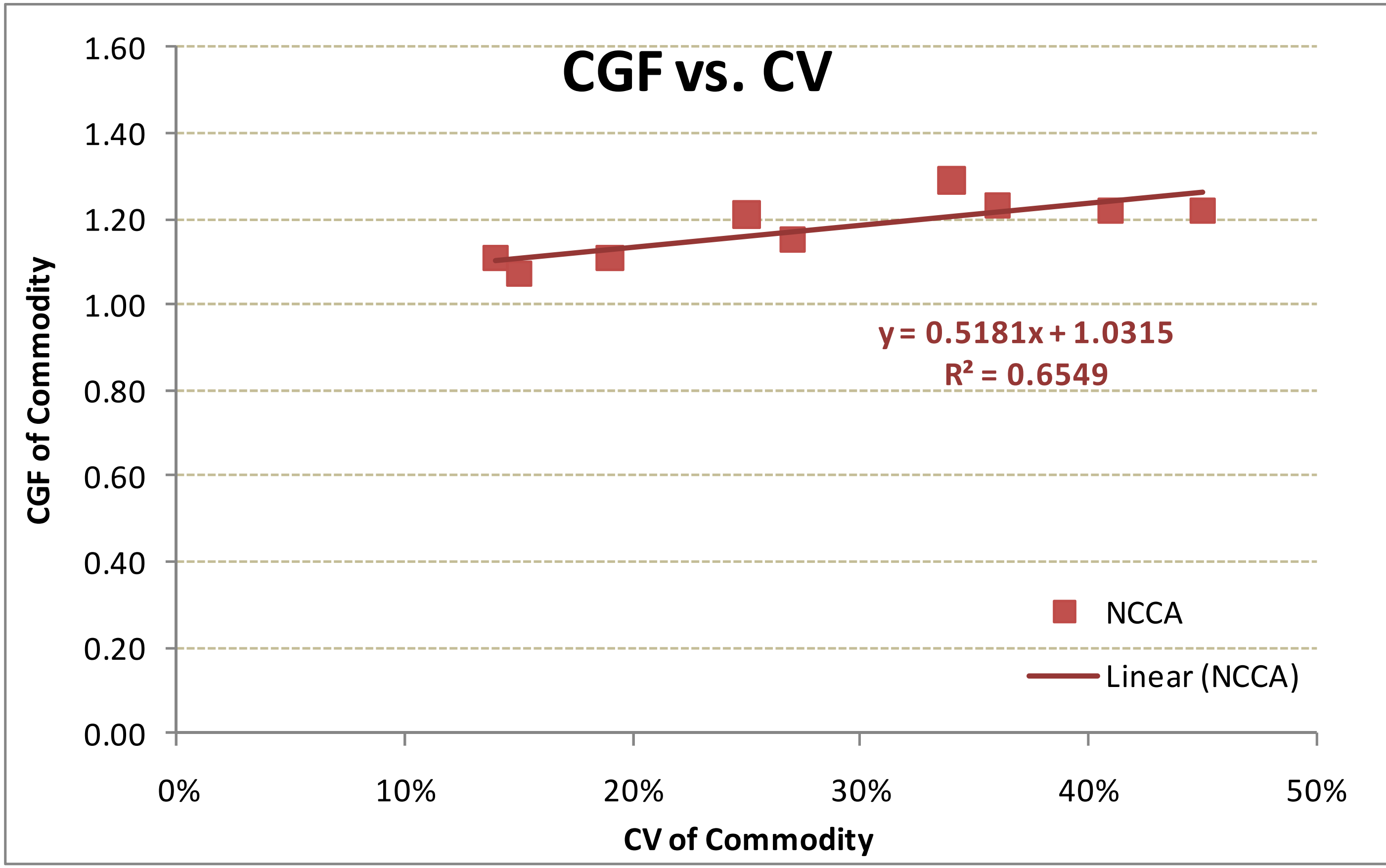
Y = Yes, N = No | H = High, M = Medium, L = Low

green = good, yellow = OK, red = poor

- No single metric does everything
- Minimum requirements can be met with the S-curve and the CV
  - With the PDF superimposed, all requirements are met

More Motivation!

- Overly tight CVs indicate overstatement of our certainty, that's a given
- Even more compelling is that the understatement of CV is tied to understatement of cost growth
  - Caveat: this is growth by commodity, and may not relate to cohorted estimate pairs
- This relationship is almost preordained since risks cause growth in CGF and CV
- The graph enables computation of an implied rise in the mean when we raise the CV
- Note: This thought process not yet coded in tool



CV Growth Implies CGF Growth			
CV	Value	Implied CGF	CGF Increase
CV_1	20.0%	1.135	
CV_2	35.0%	1.213	0.078

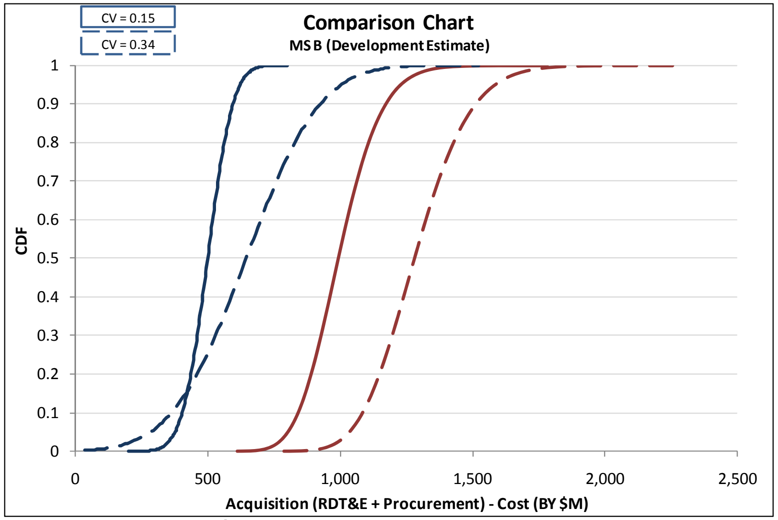
Evolution of Requirements

Tasking	Orig.	New
CV	x	
CGF		x
Single Estimate vs. Historical [Benchmarking Mode]	x	
Parametric (e.g., Normal with Mean and CV)	x	
Empirical (i.e., Risk Monte Carlo output of up to10,000 trials)		x
Point Estimate (i.e., risk analysis not yet done)		x
1st est. vs. 2nd est. (or 2 phases) and Historical [Reconciliation mode]		x

- Modest but useful additions to the tool were suggested by
  - Industry experience (Empirical and Reconciliation)
  - Consideration of NCCA's future uses (Reconciliation)
  - Reviewers (Point Estimate input)

S-Curve Tool Design Considerations

- Tradeoffs Addressed
  - Capability vs. Complexity
    - Minimizing the cost of complexity
  - Permissive vs. Restrictive Controls
    - “Pistol in the playroom” vs. “the Soup Nazi”
- Effective Design
  - Color / geography / brevity
- Organization and clarity
  - Make the flow intuitive
  - Be consistent and mnemonic
- Clear Structure (sideline the side issues)
  - Computations and data
- Error anticipation *and* detection



S-Curve Tool Development Epiphanies

Legend: Blue font signals broader application

- Probability distributions
  - Lognormal: Better understanding of the lognormal and related normal
    - CV rules of thumb for shift from median to mean or mode
    - "Pivoting" on the median vs. "pivoting" on the mean
  - Alternate specification: Normal and Lognormal
    - Any two of:
      - Mean, median, or mode; CV; and any percentile
    - Lognormal two solutions
      - Mean and percentile, or mode and percentile

Subsequent training session IN 06B "Probability Distributions for Risk Analysis" and paper "The Perils of Portability: CGFs and CVs"

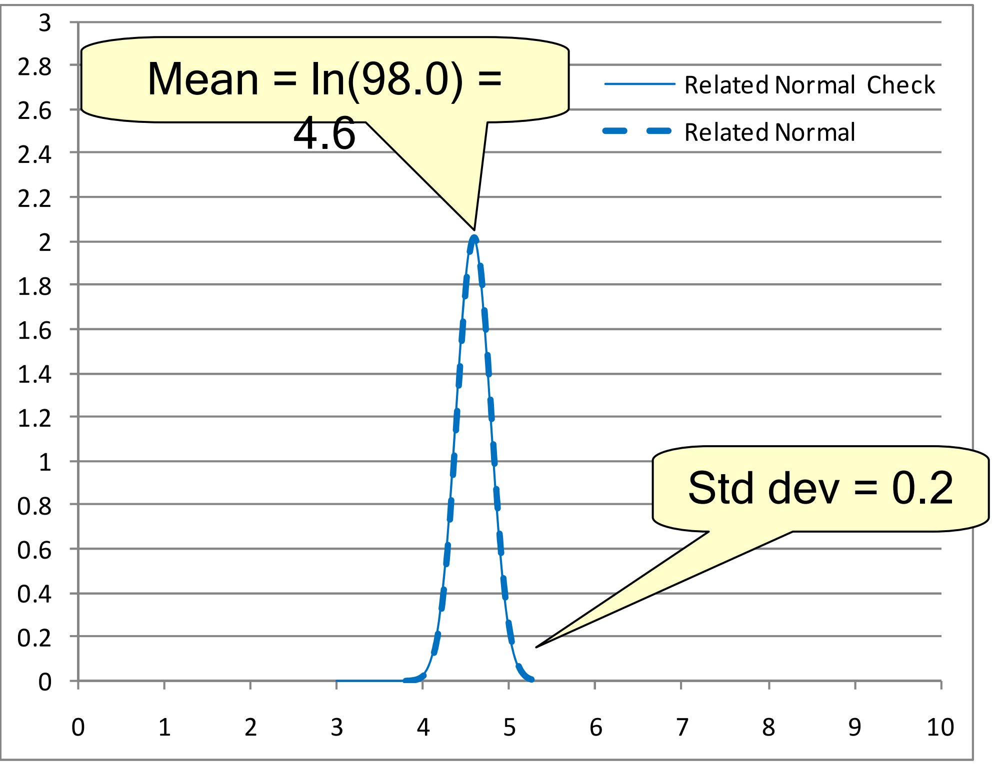
We learned that the lognormal is even harder than we'd guessed, it sometimes has two solutions

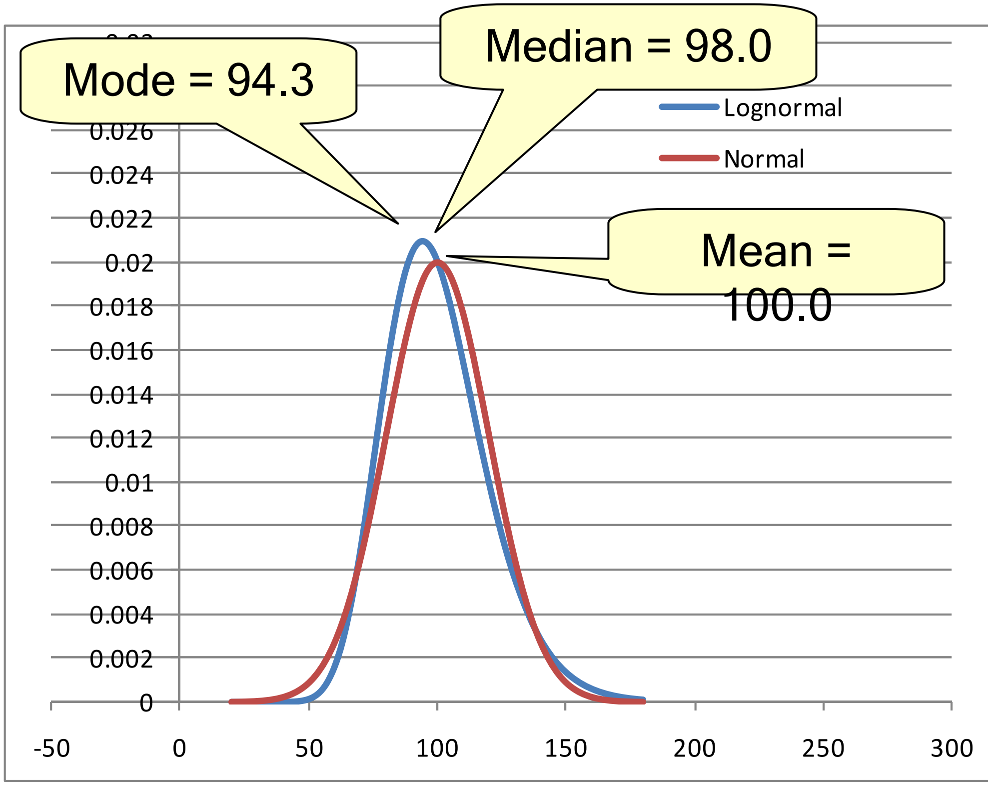
- Data analysis
  - Standard deviation vs. CV vs. pseudo-CV (s.d. divided by *median*) for historical CGFs
  - CV of CGFs vs. CV of Cost
  - Empirical percentile of 1.0 CGF vs. implied percentile given CV and CGF
  - SAR Summary data vs. SAR data vs. Contract data
- Model
  - CGF-only historical adjustment: not just a translation, because we determined to believe the CV *not* the standard deviation
  - Graphical sampling
    - Developed to solve noisy PDF problem (got worse, not better, with more trials)
    - Reduces size, increases speed of model
    - Order of operations: transformation of stats quicker than stats on transformed data

Related Normal Example

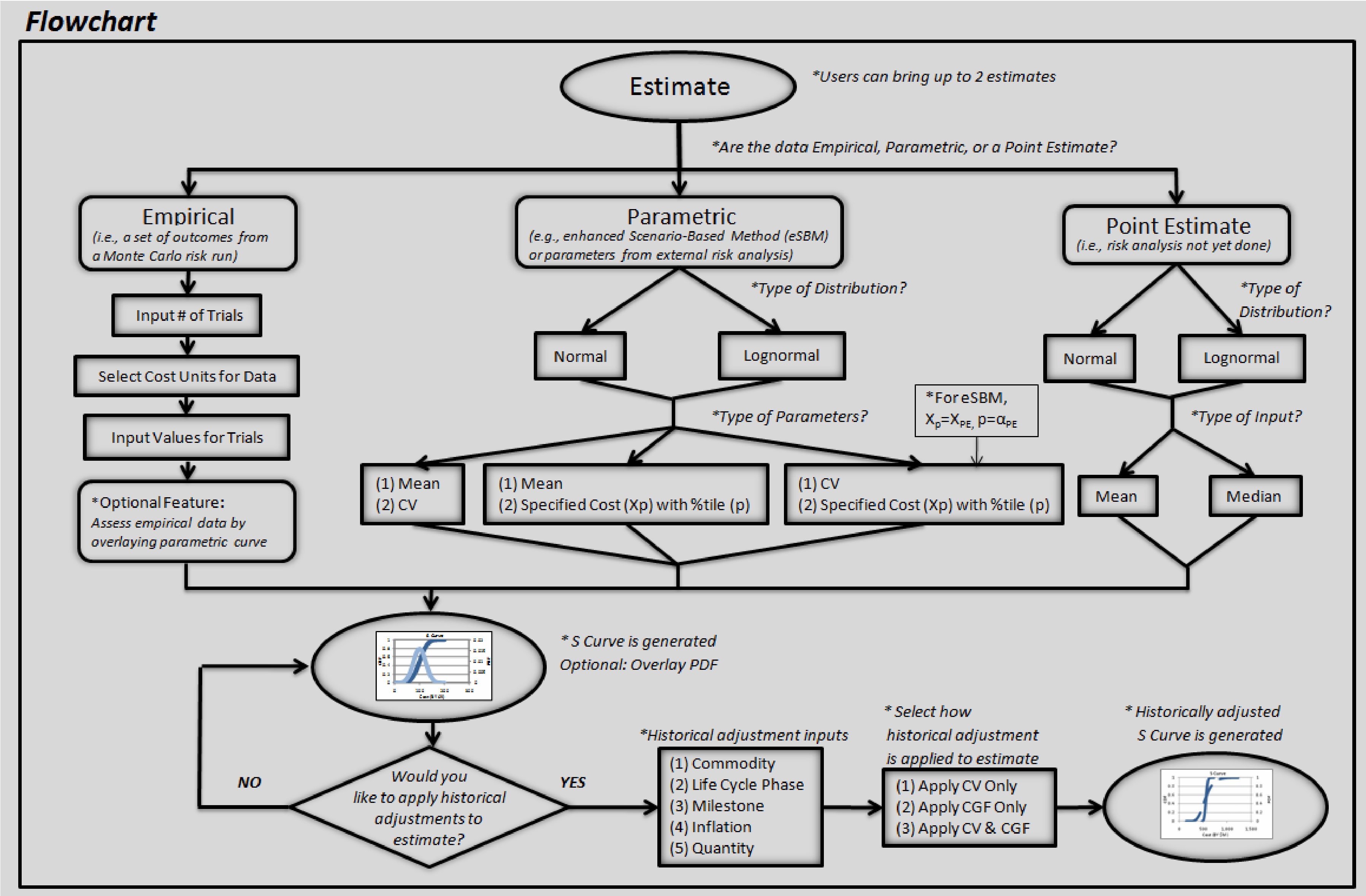
- Mean = 100, CV = 20%
  - Mode shift = -3.8% (wrt median)
  - Mean shift = +2.0% (wrt median)

CV (lognorm)	mode shift factor	mean shift factor	percentile of mean
10%	0.990	1.005	52.0%
20%	0.962	1.020	53.9%
30%	0.917	1.044	55.8%
40%	0.862	1.077	57.6%
50%	0.800	1.118	59.3%

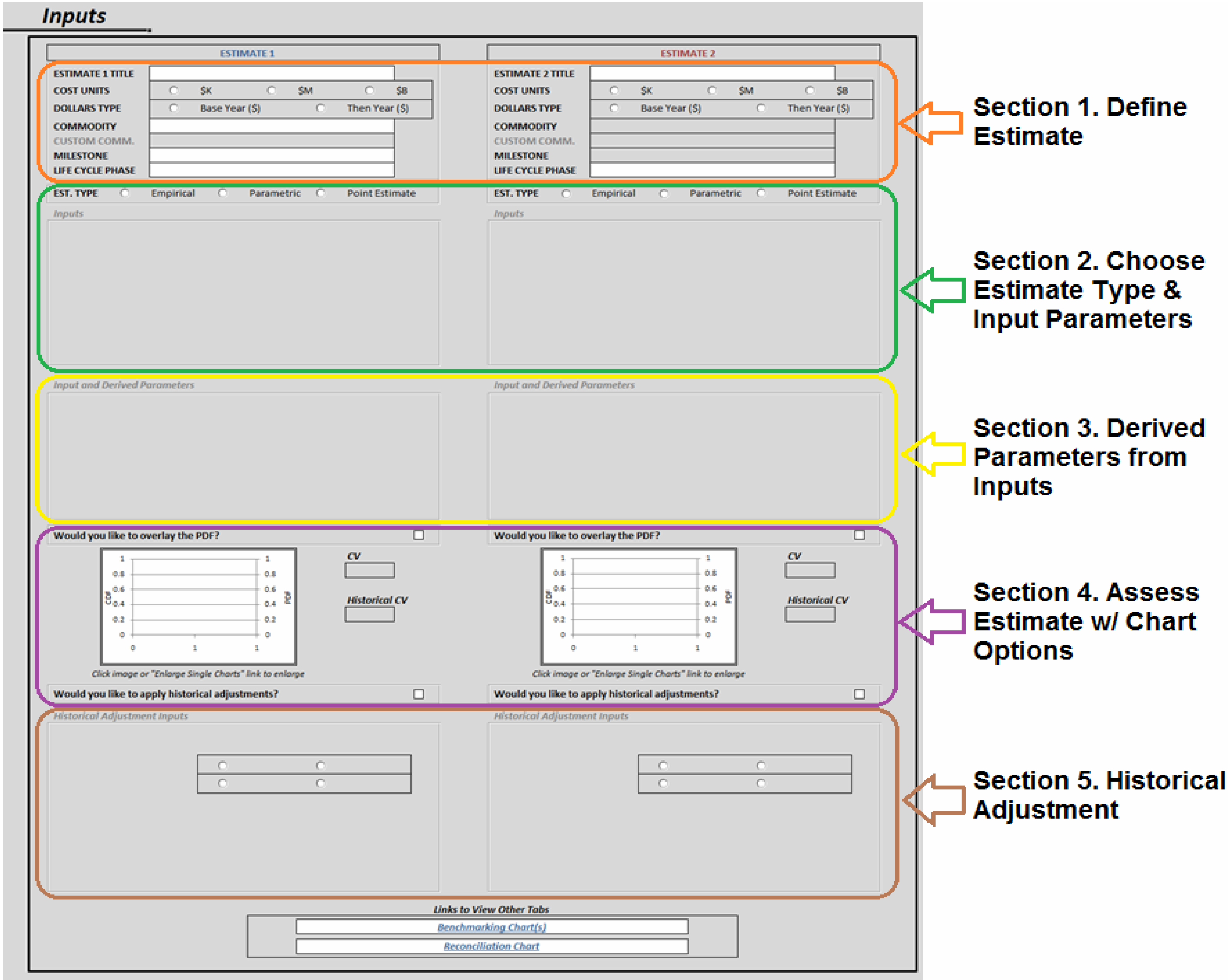




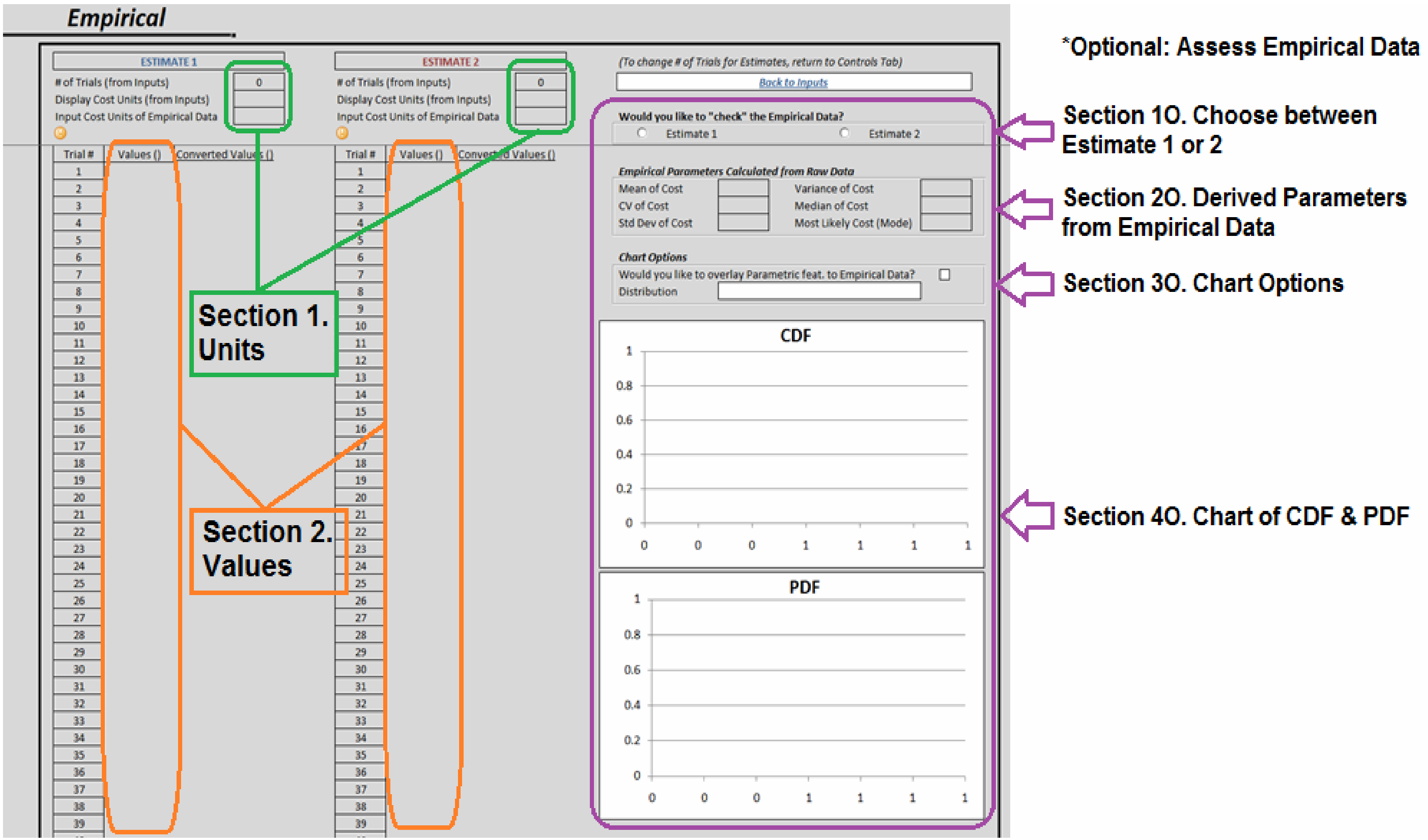
S-Curve Tool Flowchart



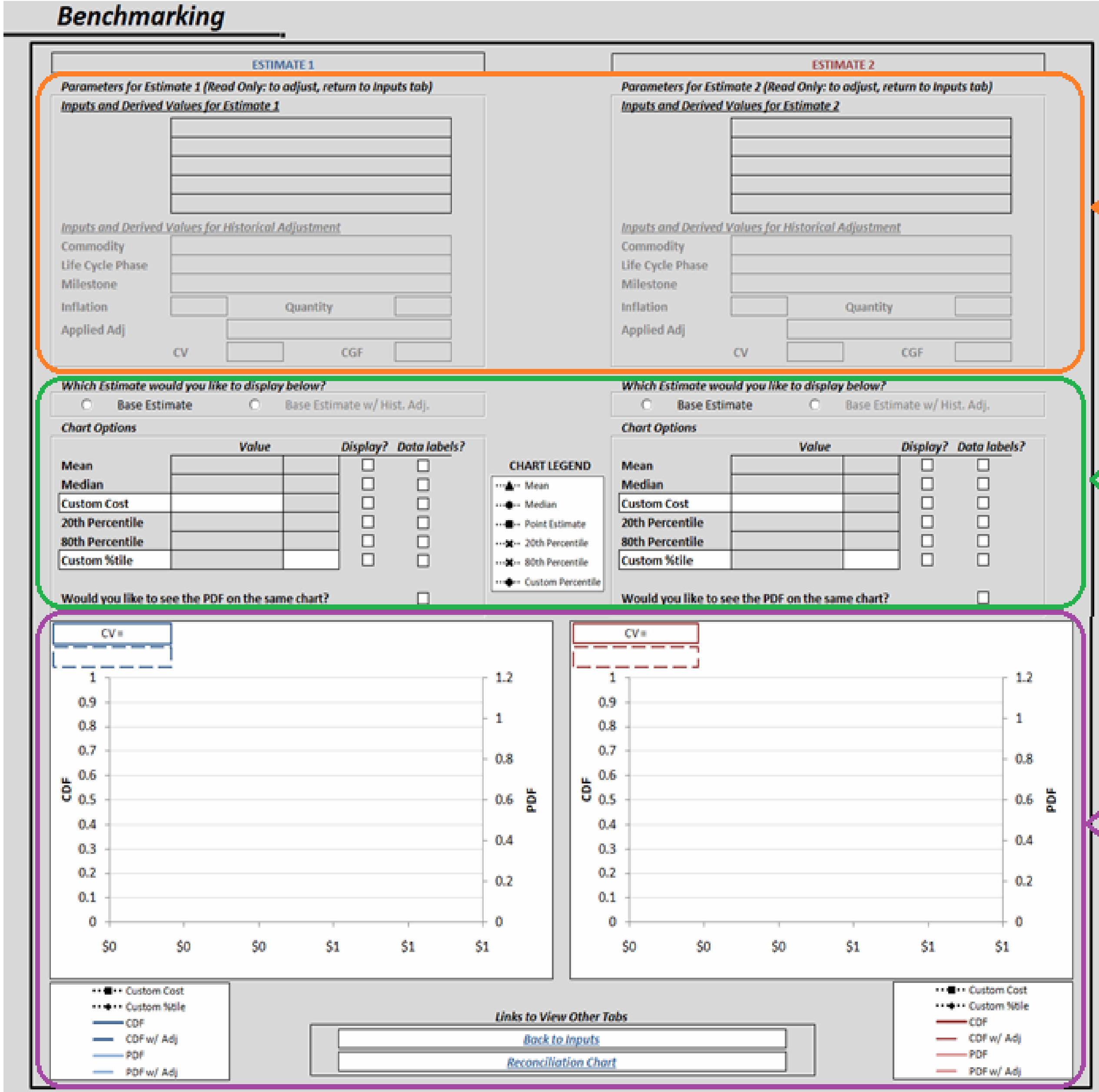
S-Curve Tool Inputs tab



S-Curve Tool Empirical tab



S-Curve Tool Benchmarking tab

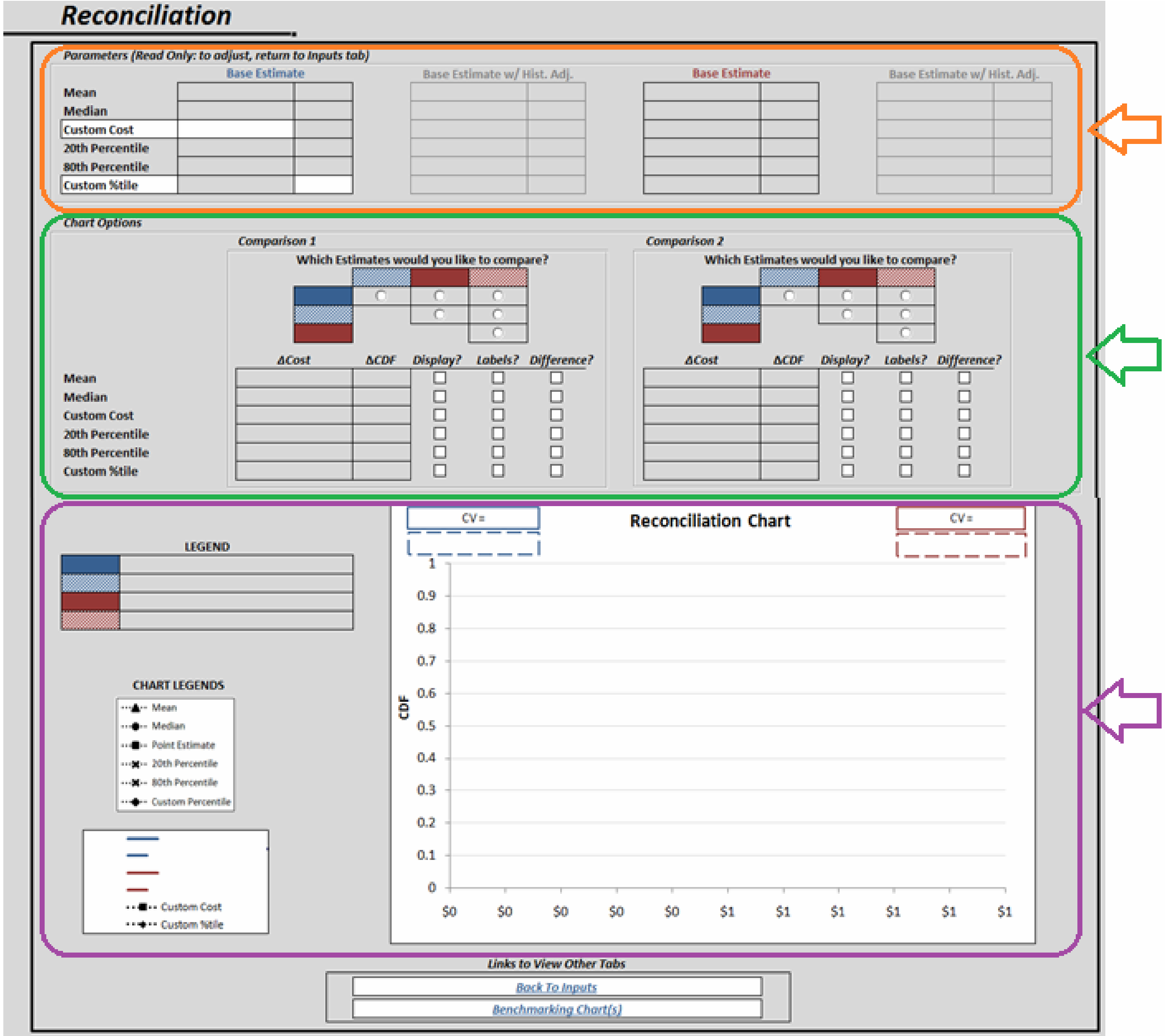


Section 1. Parameters from "Inputs" Tab

Section 2. Select estimate & apply chart options

Section 3. View chart based on selections from chart options

S-Curve Tool Reconciliation tab



Section 1. Derived Parameters from "Inputs" tab

Section 2. Select which estimates to compare & apply chart options

Section 3. View chart based on selections from chart options

S-Curve Tool Demo

“Let’s go the Excel…!”

Backup

Evolution of Risk at an OEM– Second Stage (2004)

- First ICE and Risk WG established
- Recommended significant scope increases in risk and cost
- Dictated cost team involvement in risk evaluation and quantification
- Identified pockets of excellence in cost and risk
- Suggested improvements in change order estimation, especially in cumulative effects (never fully exploited)

Evolution of Risk at an OEM – Third Stage (2006)

- Second ICE/Risk WG established
  - Dictated use of S-Curves
  - Several in-use methods demonstrated
    - Heavy on boosterism
    - Resistance to standardization due to
      - Differing commodity & legacy company of sectors
      - Angling for advantage
      - “Confederation philosophy” of corporate
- Chartered a sub-group on risk modeling

Evolution of Risk at an OEM – Fourth Stage (2006)

- Special Risk Modeling Sub-WG established
  - Competing methodologies demonstrated but no "under-the-hood" checks
  - Monte-Carlo strongly recommended but not required
- Considerable resistance to findings, Not Invented Here (NIH) syndrome was strong

Evolution of Risk at an OEM – Fifth Stage (2007)

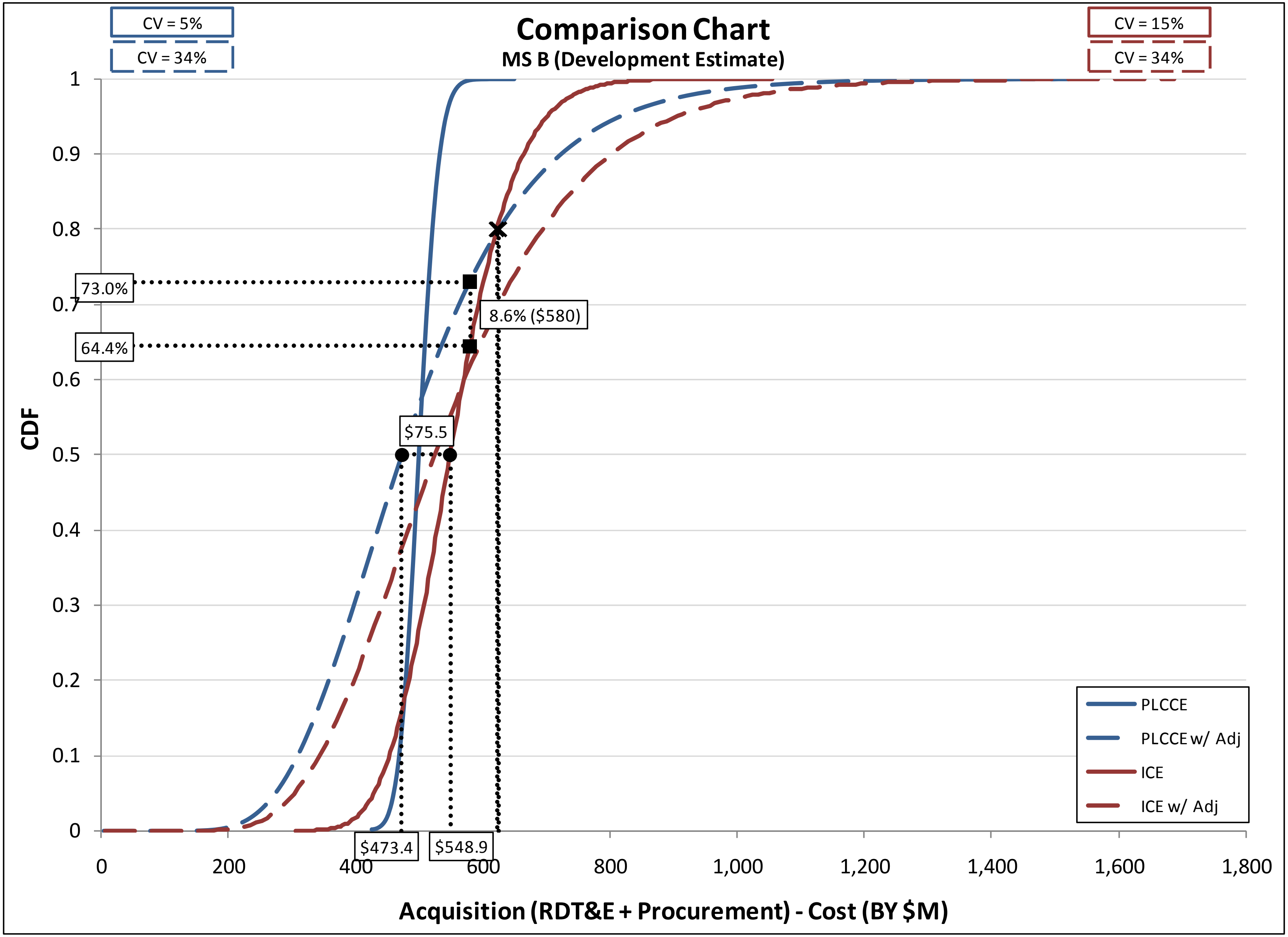
- Second Risk WG Established
  - Competing methodologies processed a sample problem and were compared
- Method of Moments was found to produce symmetric results (as implemented) and was banned
  - MoM model also allowed specification of CL for all BOEs which was found in practice to produce *systematic overstatement*
- Formats standardized including
  - CV call out on all s-curves
  - Departures from standard methodology to be expressly reported
- "People" and "Process" teams addressed training and retention and standard processes

Cost Risk Analysis in the DoN[1]

- The Cost Review Board (CRB) produces the Service Cost Position [2, 3, 4]
  - " … all resources … regardless of funding source[3]
  - Insight into Cost Drivers, Cost risk and uncertainty, Total Ownership Cost[4]
- Prior to WSARA[5], comparison of point estimates was usual
  - Comparison of PLCCE to ICE
- After WSARA, comparison of the range of potential cost outcomes became the standard
  - Using the S-curve, the cumulative distribution function (CDF) of cost

1. Service Cost position (SCP) – Process and Discussion, CAPT J. Baratta, DoNCAS September 2009
2. OSD Memorandum, 12 March 2009, Required Signed and Documented Component Level Cost Position for Milestone Reviews
3. ASN (RD&A) & ASN (FM&C) Memorandum, 7 Jan 2010, Department of the Navy Service Cost Position
4. SECNAVINST 5223.2, 16 Dec 2008, Department of the Navy Cost Analysis
5. Weapon Systems Acquisition Reform Act of 2009, (Public Law 111-23)

Reconciliation Mode of S-Curve Tool



S-Curves in Industry

Excerpts from

Risk Quantification and the S-Curve

Richard L. Coleman, Peter J. Braxton, Eric R. Druker, Patti J. Tisone

Presented at the Northrop Grumman Cost, Pricing and Supply Chain Conference October 2008

Risk

What is Risk?

- Risk is a word that's tossed about with little common agreement as to its meaning
  - In lay terms, risk refers to bad things that may happen
- In both finance and cost estimating, values both grow (or shrink) and fluctuate
  - In finance, risk refers more to *variability*, less to *bias*
    - Risk analysis tends to focus on how much the value fluctuates
  - In cost estimating, risk refers more to *bias,* less to *variability*
    - Risk analysis tends to focus on how much the value grows
- In cost, various forms of risk analysis are in use; all attempt to cope with the questions "how much will this cost estimate grow?" and "plus or minus what?"
  - Risk Cube
    - Probability of Failure (Pf), Consequence of Failure (Cf), Expected value of failure (Ef = Pf * Cf)
  - Inputs risk
    - "If weight grows x, cost grows y"
  - Outputs risk
    - "Programs/WBS items like this tend to grow P percent"
  - Expert-Based
    - "This program/WBS item could be as low as L, probably will be M, but could grow to H"
- However done, risk isn't "something that may happen"
  - Risk is "something that will happen"
  - The only question is "how much"

Opportunities - *The Softer Side of Risk*

- Mathematically, risks are positive (increase cost) or negative (decrease cost)
  - In practice, events are divided into "Risks" or "Opportunities"
  - Each has a probability of occurrence, usually called *p*
- Opportunities are not numerous in ICEs
  - About $^1/_5$ of "Risks," or less
  - This is because Opportunities are almost all "taken" by Proposal Teams (PTs)
    - "Opportunities are like donuts … there's never any left on the table"
  - Further, Proposal Teams are prone to "bake in" Opportunities without regard to *p*, the probability of occurrence
    - The PT bakes them in at full value, ignoring that they may not happen
    - This turns them into Risks at the complementary probability of 1 - *p*

How Much? *Measuring Bias in Cost Estimating*

- Bias* in the risk adjusted cost estimate is specified using one of the below measures of central tendency
  - Mode
    - Not common, not particularly useful
  - Median = 50th %-ile
    - Common, but not as good a metric as the mean
  - Mean = Expected Value
    - Becoming more common, it is the long-term outcome

* Customarily we are prone to think of “negative bias” as meaning “the estimate is probably low” but we express the risk as positive if we mean “the estimate will probably grow.”

How Much? *Measuring* Variability *in Cost Estimating*

- Variability is the measure of how much different the result may be from what we expect. It is also known as dispersion
- Variability is measured in one of six ways
  - Standard Deviation
  - Coefficient of Variation (Standard Deviation over Mean)
    - The standard deviation expressed as a percent of the mean ("± y%")
  - 20th, 50th, 80th %-iles expressed as three dollar values
  - Three-Point Estimate (Low, Most Likely and High) expressed as three dollar values
    - Similar to 20/50/80 method
    - Believed by most risk analysts to be understated and to really represent the 10/90 or the 20/80
    - The most likely often claimed (with little basis) to be the Mode
  - PDFs
  - S-curves (CDFs)

Variability Measures

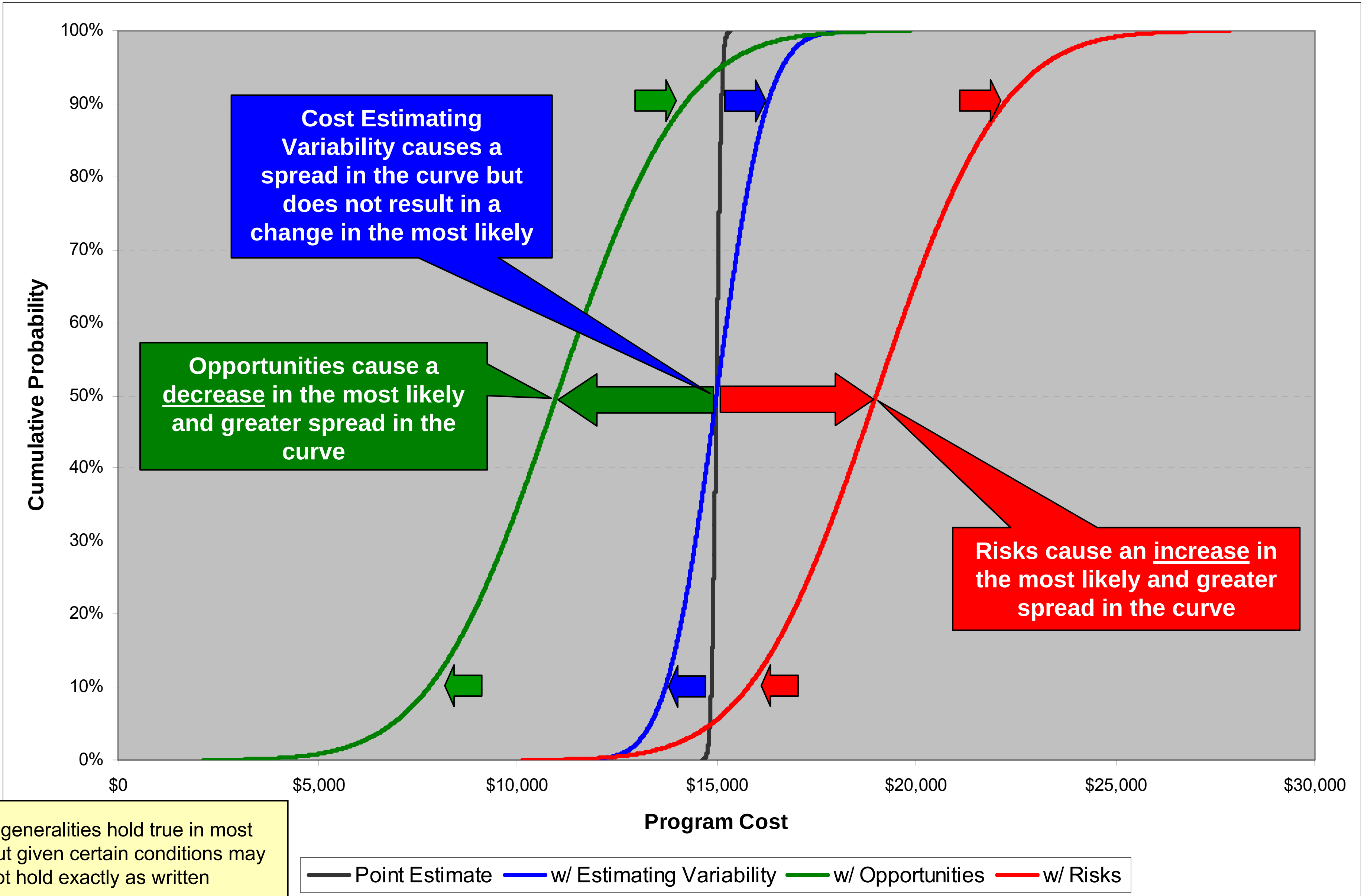
- Standard Deviation
  - Asymmetry not portrayed
  - Hard to interpret in degree
- Coefficient of Variation (CV = Standard Deviation over Mean)
  - Shows degree of variability
  - Easily compared across programs
  - Asymmetry not portrayed
- 20^{th}, 50^{th}, 80^{th} %-iles expressed as three dollar values
  - Shows asymmetry, but requires "mental arithmetic"
  - Hard to compare across programs
  - Similar to Standard Deviation and CV method because the 20 and the 80 are just a bit narrower (60% between them) than ± one Standard Deviation (68.3%)
- Three-point estimate (Low, Most Likely and High) expressed as three dollar values
  - Shows asymmetry, but requires "mental arithmetic"
  - Hard to compare across programs
  - Similar to 20/50/80 method
  - Believed by most risk analysts to be understated and to really represent the 10/90 or the 20/80
  - The most likely often claimed (with little basis) to be the Mode
- PDFs
  - Difficult to generate
  - Good display, shows asymmetry *extremely* well
  - Cannot be read directly, it must be 'integrated by eye"
- S-curves (CDFs)
  - Difficult to generate
  - Good display, shows asymmetry *poorly*
  - Can be read directly

S-Curves

S-Curves

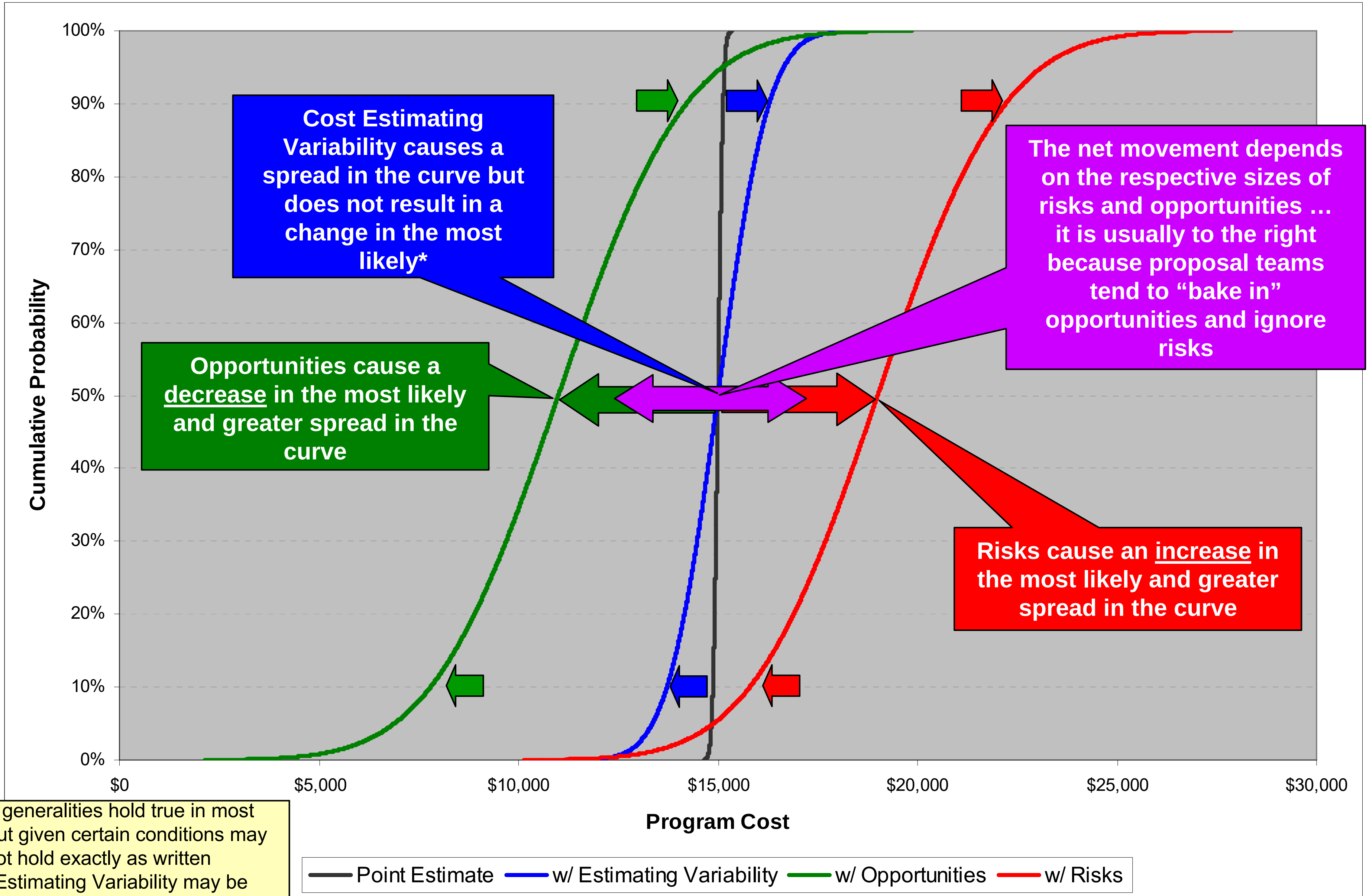
- The statistical name for an S-curve is the Cumulative Distribution Function (CDF) for the total cost of a program
  – Each point on the curve indicates the *cumulative* probability (y-value) that the cost will be ≤ that amount (x-value)
- S-curves for cost estimates are derived from three sources:
  – The base cost estimate
    - In industry, the Proposal Team Cost Estimate
    - In DoD, the PLCCE
  – Cost Estimating Variance
    - The uncertainty of the Proposal Team Cost Estimate (or PLCCE)
  – Risks and Opportunities
    - Events or changes in assumptions that can cause the costs incurred on a program to rise above or below the estimate
    - Can also be characterized as "Lowest, Most Likely and Highest" ("three-point estimate")
- S-curves are generated from the above list of inputs by the use of "Monte Carlo"
  – Monte Carlo just means simulation of a number of individual outcomes and plotting the histogram of their values
- S-curves are typically used by management to understand the range of potential costs for a program
  – Enables appropriate business decisions
  – Aids in negotiations, especially contract type and "geometry"

S-Curves – The Shaping Forces



These generalities hold true in most cases but given certain conditions may not hold exactly as written

S-Curves – The Shaping Forces



These generalities hold true in most cases but given certain conditions may not hold exactly as written
Cost Estimating Variability may be accompanied by a (small) bias)

(1) Cost Estimating Variance

- The first step in producing an S-curve is determining cost estimating variance
  - Cost estimating variance is an error band or S-curve caused by variance in the cost estimate rather than risks and opportunities
  - It may be accompanied by a bias
- There are a number of ways this variance can be determined (in order of "preference"*):
  A. Statistical analysis of the data used to build the Proposal Team Cost Estimate
  B. Statistical analysis of the data used to build an Independent Cost Estimate
  C. Statistical variability of similar analyses – a specific close analogy
  D. Historical data at a top level – a general analogy
  E. Subject Matter Expert (SME) interviews - SMEs are interviewed and give ranges
  F. A mix of the above

*** Order of preference is arguable. Many respectable analysts believe that (E) SME interviews are preferable to (D) historical data at a top level. It amounts to which of the two competing views we believe:**

- **Experts have sufficient exposure and are effective at assimilating that exposure and producing cumulative results**
- **Judiciously used history is a good guide to the present and the future**

(2) Risks and Opportunities

- The second step is to generate a list of risks and opportunities (what could go worse / better)
  - The "risk register" (if 2+ sources, a "conflated risk register")
- Risks cause costs to rise
  - Opportunities cause costs to drop
- In practice, risks outweigh opportunities (~ 4-to-1) causing a net rise
  - Unless the Proposal Team Cost Estimate is very conservative, which is rare
  - Historical data[1] shows that ~12.5% of DoD programs come in "at or under"
- Risks and opportunities can be:
  - Discrete: Specific events with a probability of occurrence and a cost impact (called *Bernoulli* random variables, or "Pf/Cf" or "categorical")
    - Example: Schedule/Technical risks
  - Continuous: Events that always occur, but with varying cost impact
    - Example: Learning curve, precious material prices, SLOC growth,
    - Continuous risks have mass on the "opportunity" and the "risk' end
- Risks & Opportunities are probabilistic and always produce variability
  - Spreads the 20th and 80th %-iles from the 50th %-ile
  - In ICEs their contribution is less that the Cost Estimating Variance
    - But in historical data they are larger[2,3], which should worry us

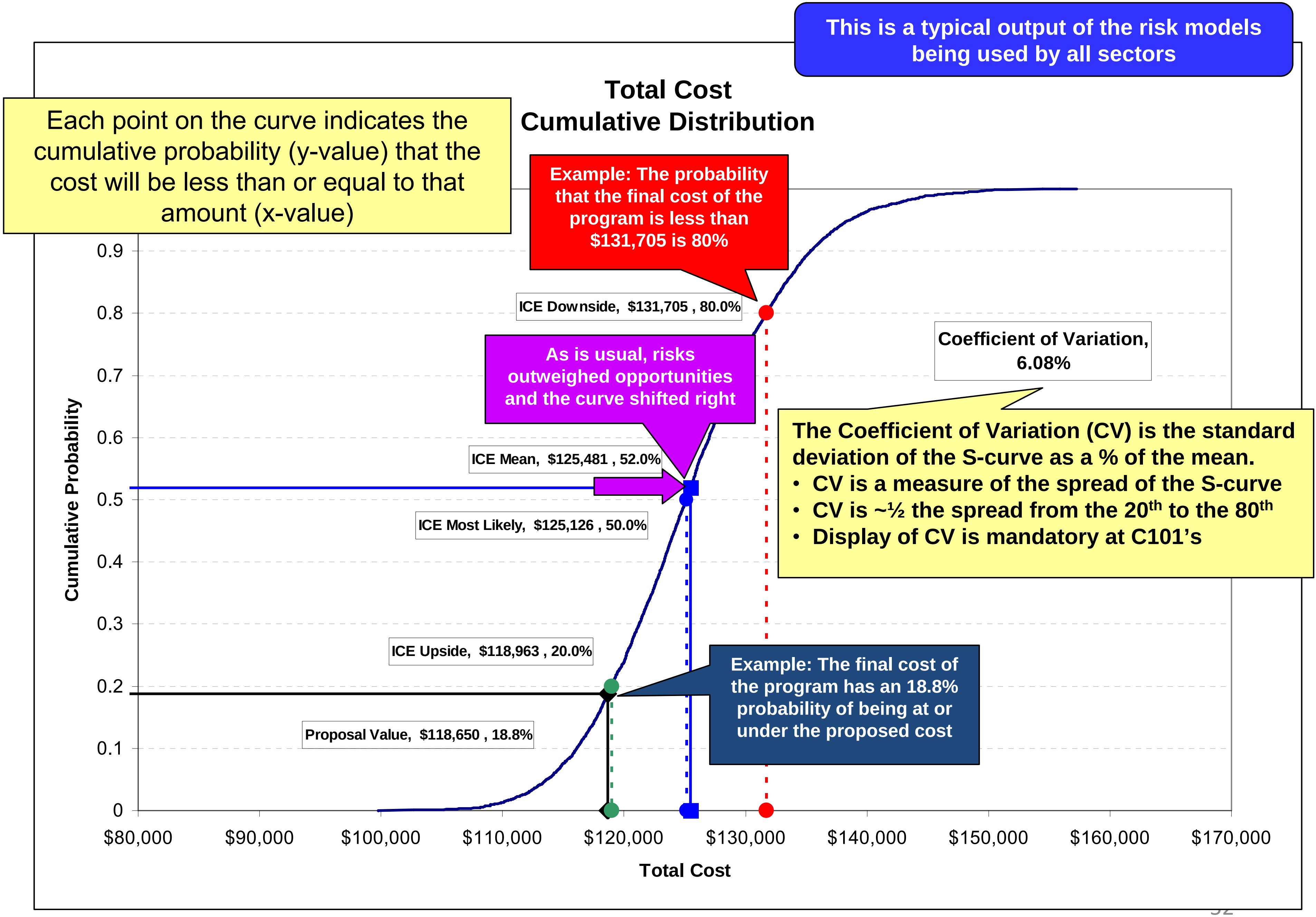
1.*Risk in Cost Estimating General Introduction & The BMDO Approach*, R. L. Coleman, J. R. Summerville, M. DuBois, B. Myers, 33rd DoDCAS 2000
2. ibid
3. *NAVAIR Cost Growth Study*, R. L. Coleman, M.E. Dameron, C.L. Pullen, J.R. Summerville, D.M. Snead, 34th DoDCAS and ISPA/SCEA 2001

(3) The Monte Carlo

- The third step is running a Monte Carlo simulation to produce the S-curve
  - We don't have to do a Monte Carlo to get the expected value, we need it to get various percentiles[4]
- The Risk Register is usually combined with the cost estimating variance to produce the final S-curve in one step
- Once the S-curve is produced, we generally draw attention to the 20th - 50th - 80th percentiles
  - The cost being put forth by the proposal team is also placed on the S-curve
  - See next slide for detailed illustration

4. ***Making Risk Management Tools More Credible - Calibrating the Risk Cube*, J. R. Summerville, R. L. Coleman, M. E. Dameron, SCEA 2006, DoDCAS 2007, Awarded ISPA/SCEA Best paper on Risk**

Reading an S-Curve



A Quick Calculator

- This formula, while not intended for general use, is a powerful teaching device that once explained, gives a good mental image of the interplay between risk and confidence
- The formula shows that you can interpolate between the 20th and 50th (or 50th and 80th) percentiles by noting the ratio of the mean of the risk register to the CV
- As a formula:

Given:

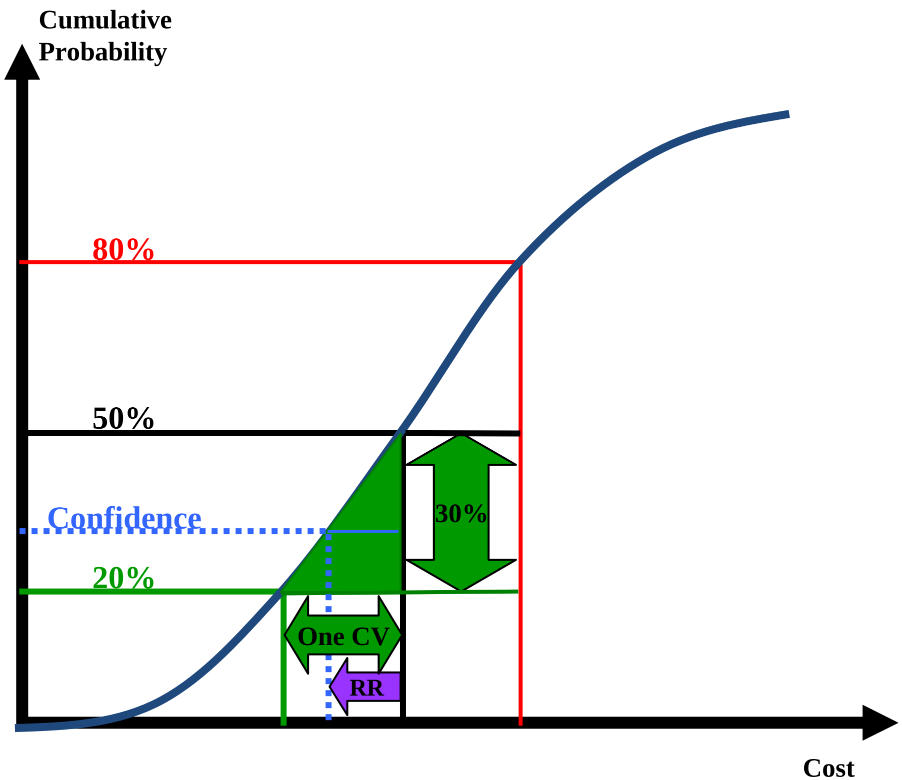
Mean Risk = the expected value of the risks as a percentage of the proposal value

CV = the standard deviation of the proposal value as a % of the proposal value

Then:

Confidence = 50 – 30 * [(Mean Risk)/CV]

- In other words, if you know the CV and the mean of the risk register, the confidence level of the estimate is easily approximated*

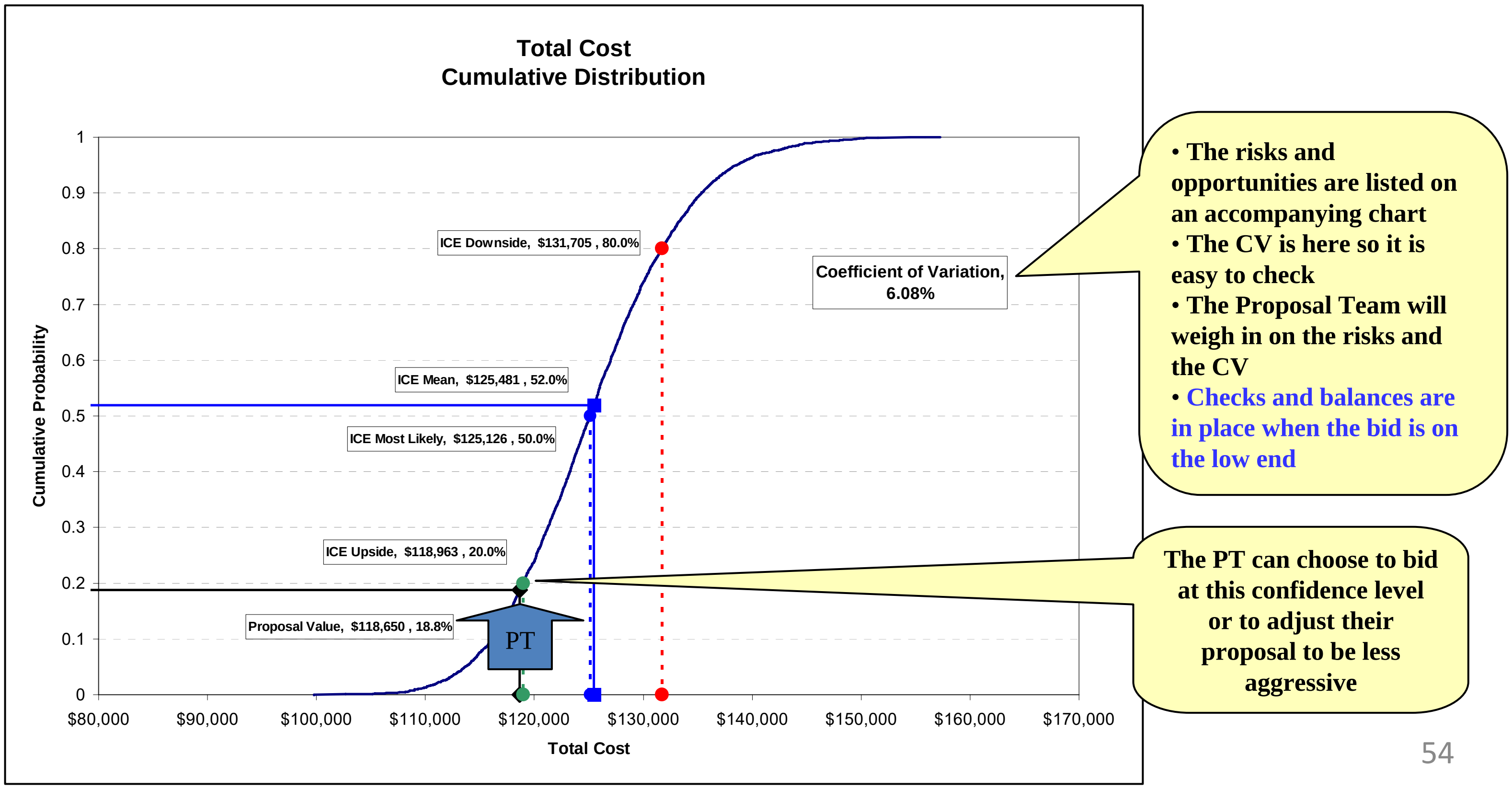


* To be exact, it's the 50th percentile of the risks, not the mean, and the interpolation is not exactly linear, but this gets you very close … the Monte Carlo will give the "exact" answer

What Should the Reviewer Look at?

An Aggressive Proposal and Some Considerations

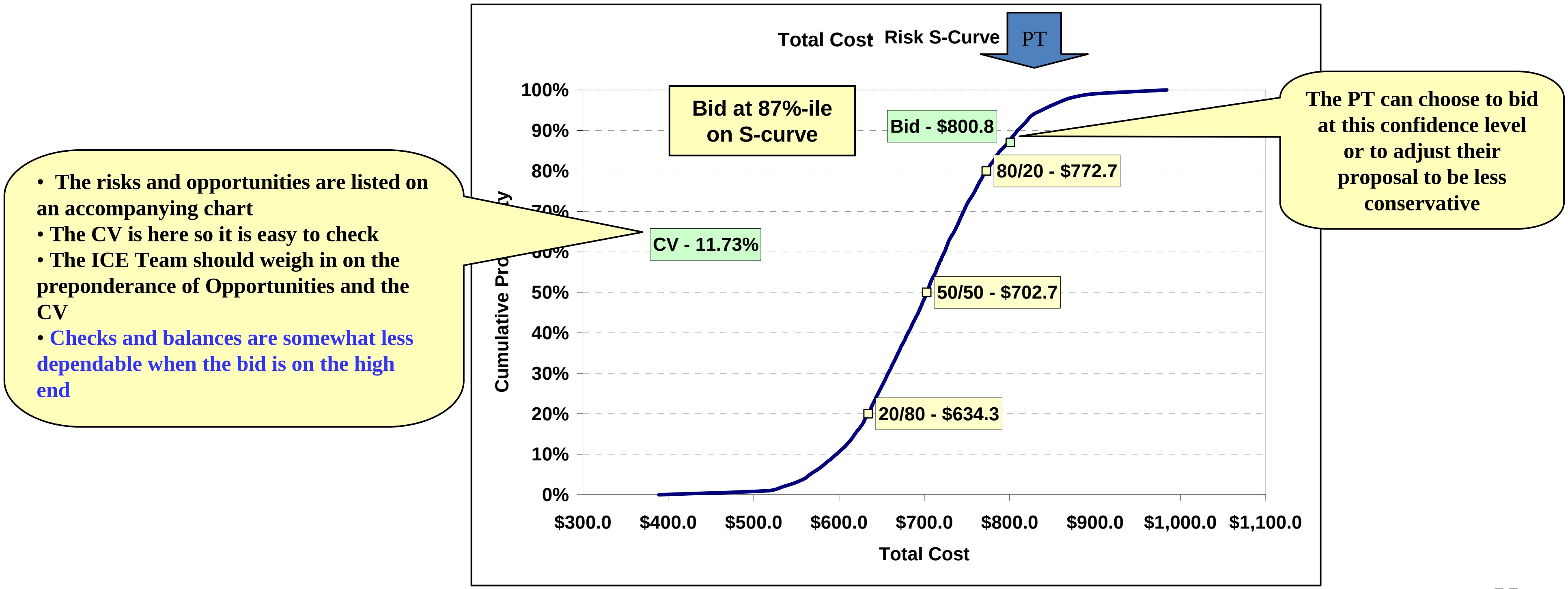
- Below is a recent S-curve for an aggressively-bid program
- The Proposal Team value is on the <u>low</u> end of the S-curve
  - Risks outweighed opportunities by 4-to-1, which is about average
  - The burden of review in this case is to be sure that:
    - Opportunities were not missed and risks are realistic (not overstated)
    - The CV is not <u>understated</u> causing the bid to seem unduly low confidence
  - The business proposition to be considered is whether this bid can be executed if won … was it unduly influenced by PTW or optimism?



Cost Estimate S-Curve

A Conservative Proposal and Some Considerations

- Below is a recent S-curve for a conservatively bid program
- The Proposal Team value is on the high end of the S-curve
  - Opportunities outweighed risks by $90M to $20M, which is uncommon
    - The PT deliberately created opportunities and left them in, to allow a cushion, and put in MR as well
  - The burden of review in this case is to be sure that
    - Risks were not missed and opportunities are realistic (not overstated)
    - The CV is not understated which would overstate the confidence
  - The business proposition to be considered is whether this bid will win … was it unduly influenced by pessimism or the false assurance of incumbency



S-Curves Bottom Line

- “It’s the ROS, stupid!”
  - We care about the S-curve of cost because it’s directly related to Return On Sales (ROS)
  - S-curve, together with contract geometry, defines possible ROS outcomes
  - Corporate hurdle rates often specify ROS at 80th percentile
- “I’m too low on the S-curve, what do I do?”
  - Add/increase Management Reserve (MR), if allowed
  - Bid more conservatively/less aggressively (e.g., flatter learning curve)
  - Do nothing! (strategic bid or mitigated by contract geometry)
- “I’m too high on the S-curve, what do I do?”
  - Reduce/eliminate MR, if applicable
  - Bid more aggressively/less conservatively (e.g., steeper learning curve)
  - Do nothing! (sole source or punitive contract geometry)

“Say it ain’t so, Joe”

- Do *not*:
  - Arbitrarily adjust estimates, thereby invalidating sound Basis Of Estimate (BOE) documentation and setting the PM up for failure
  - Arbitrarily adjust well-documented risks and opportunities, thereby painting a rosy picture and setting the PM up for failure
  - Change the scale of the graph to make the S-curve look steeper or flatter

The Last Thought

- Risk analysis and S-curves cannot be reduced to a “Process” that will never go wrong
- Risk analysis and S-curves are no better than the validity of the risks and opportunities that go into them
- The most dangerous thing about risk analysis and S-curves is that the final product is indistinguishable as to quality … it is only by the scrutiny of the inputs that they can be trusted

Opportunities and Cost Estimating Variability

toy problem to illustrate three issues:

1) "baking in" opportunities,

2) the effect of Cost Estimating Variance,

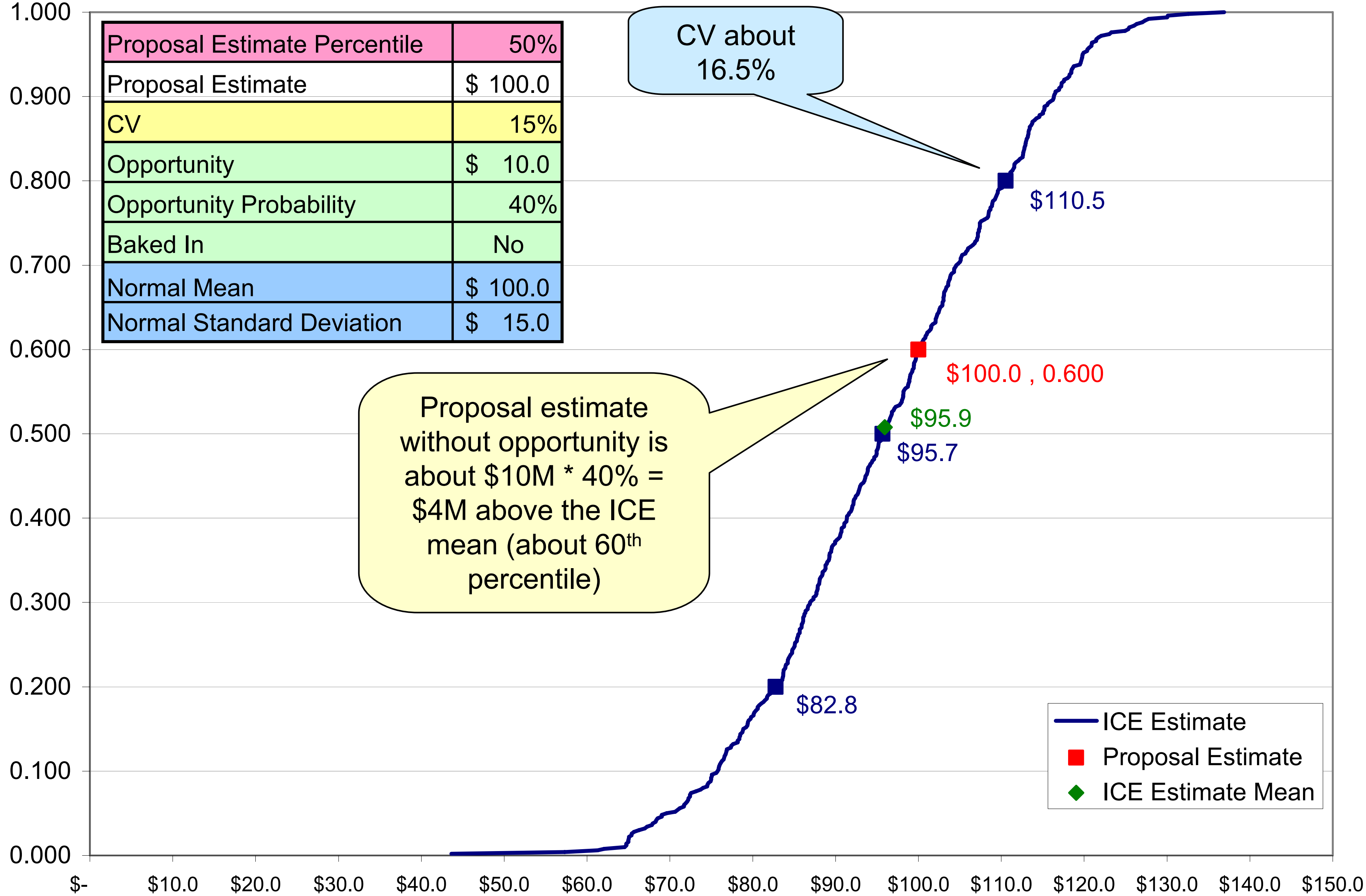
and

3) the mischief of arbitrarily assigning Confidence Levels to estimates

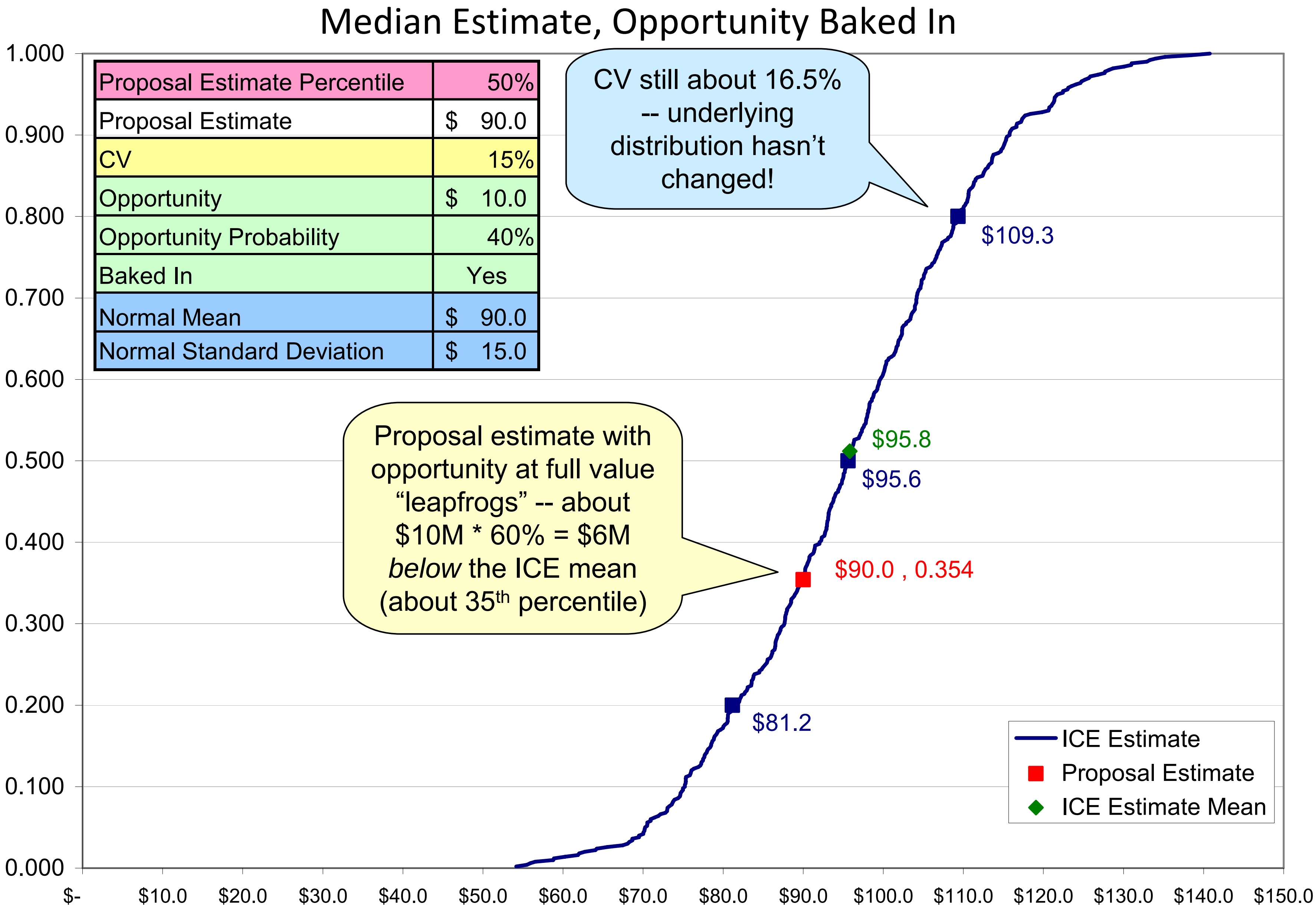
Base Case

Median Estimate, Opportunity Not Baked In

Proposal Estimate Percentile	50%
Proposal Estimate	$ 100.0
CV	15%
Opportunity	$ 10.0
Opportunity Probability	40%
Baked In	No
Normal Mean	$ 100.0
Normal Standard Deviation	$ 15.0

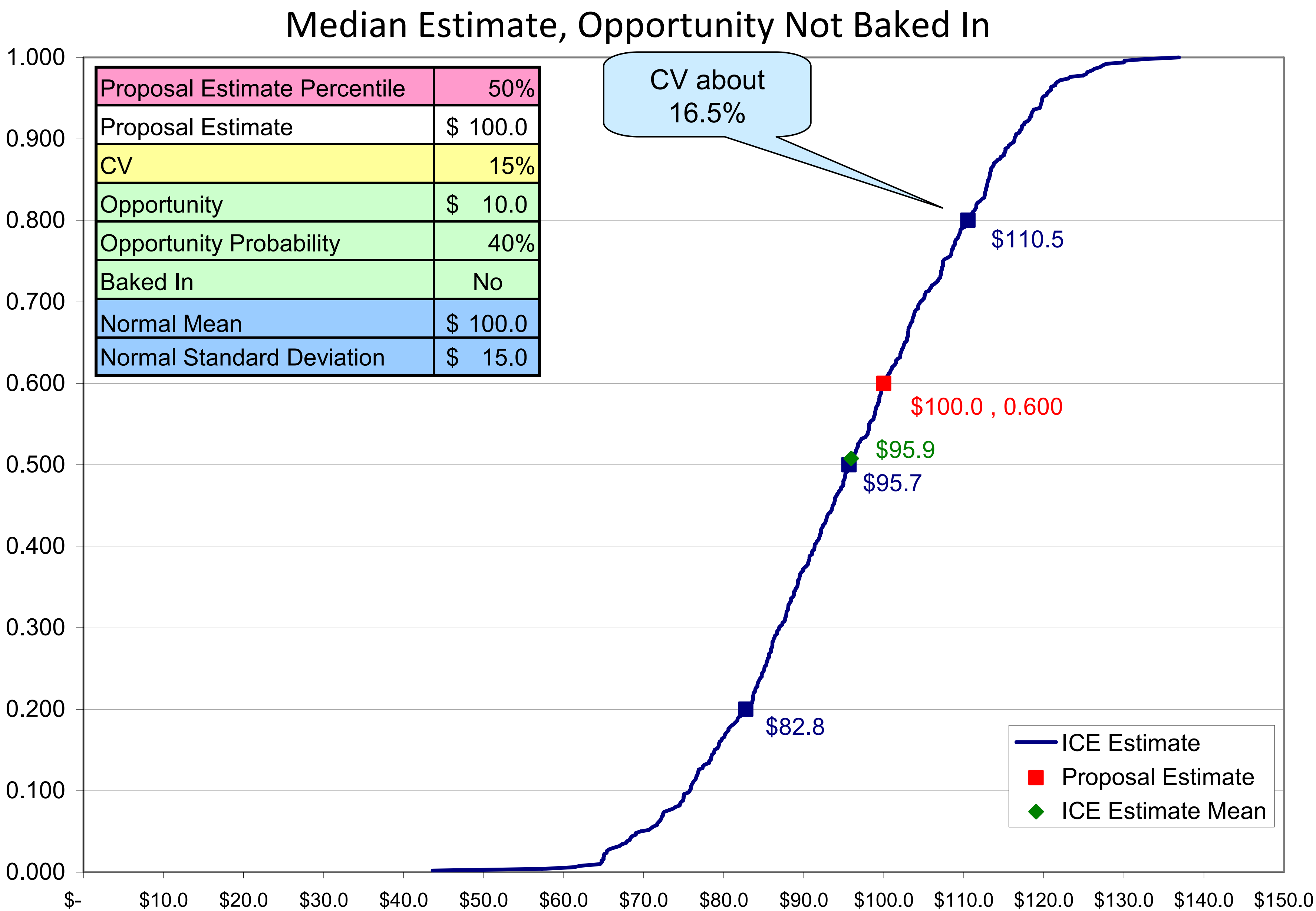


“Opportunity Taken” Case



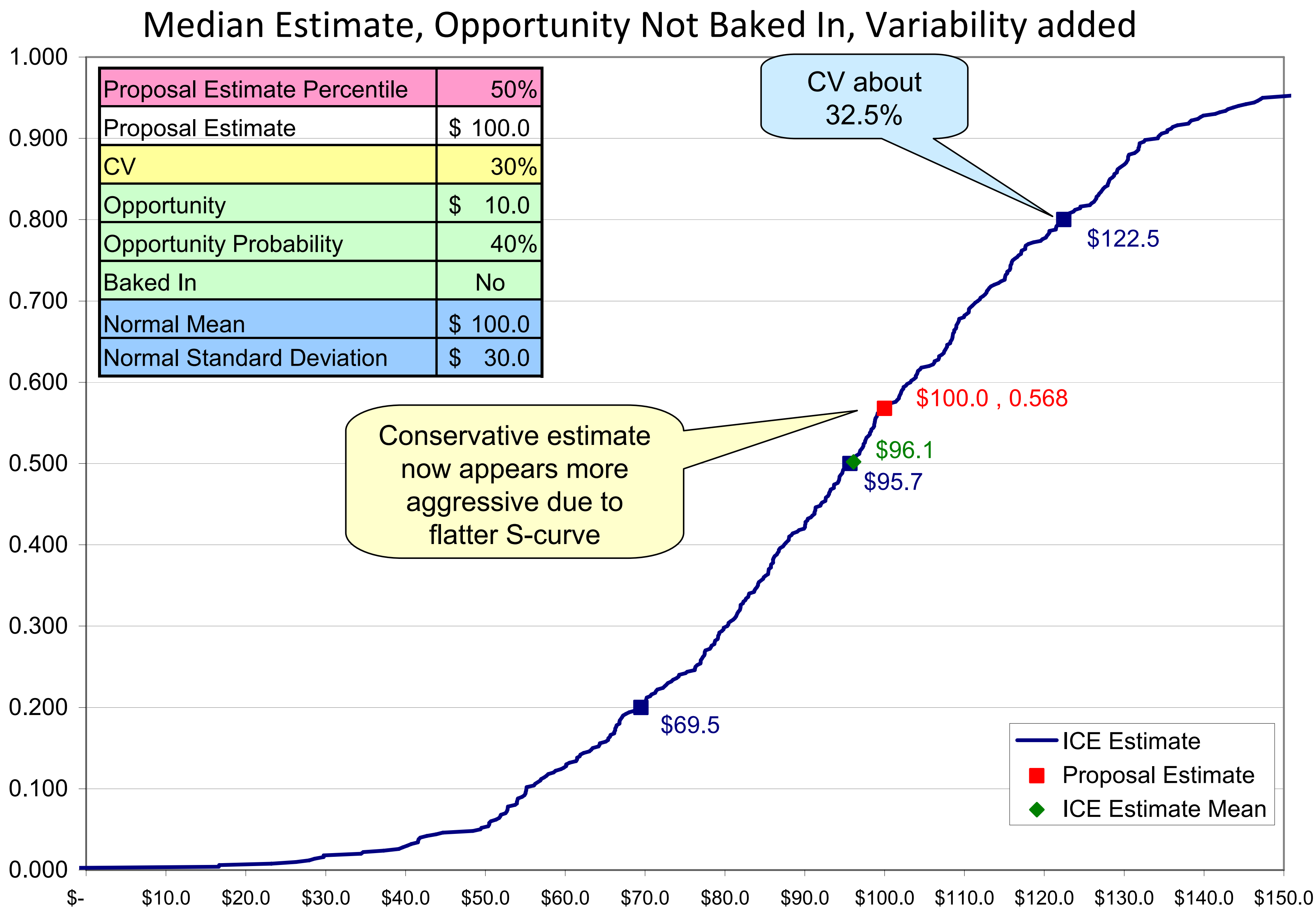
Proposal Estimate Percentile	50%
Proposal Estimate	$ 90.0
CV	15%
Opportunity	$ 10.0
Opportunity Probability	40%
Baked In	Yes
Normal Mean	$ 90.0
Normal Standard Deviation	$ 15.0

Base Case



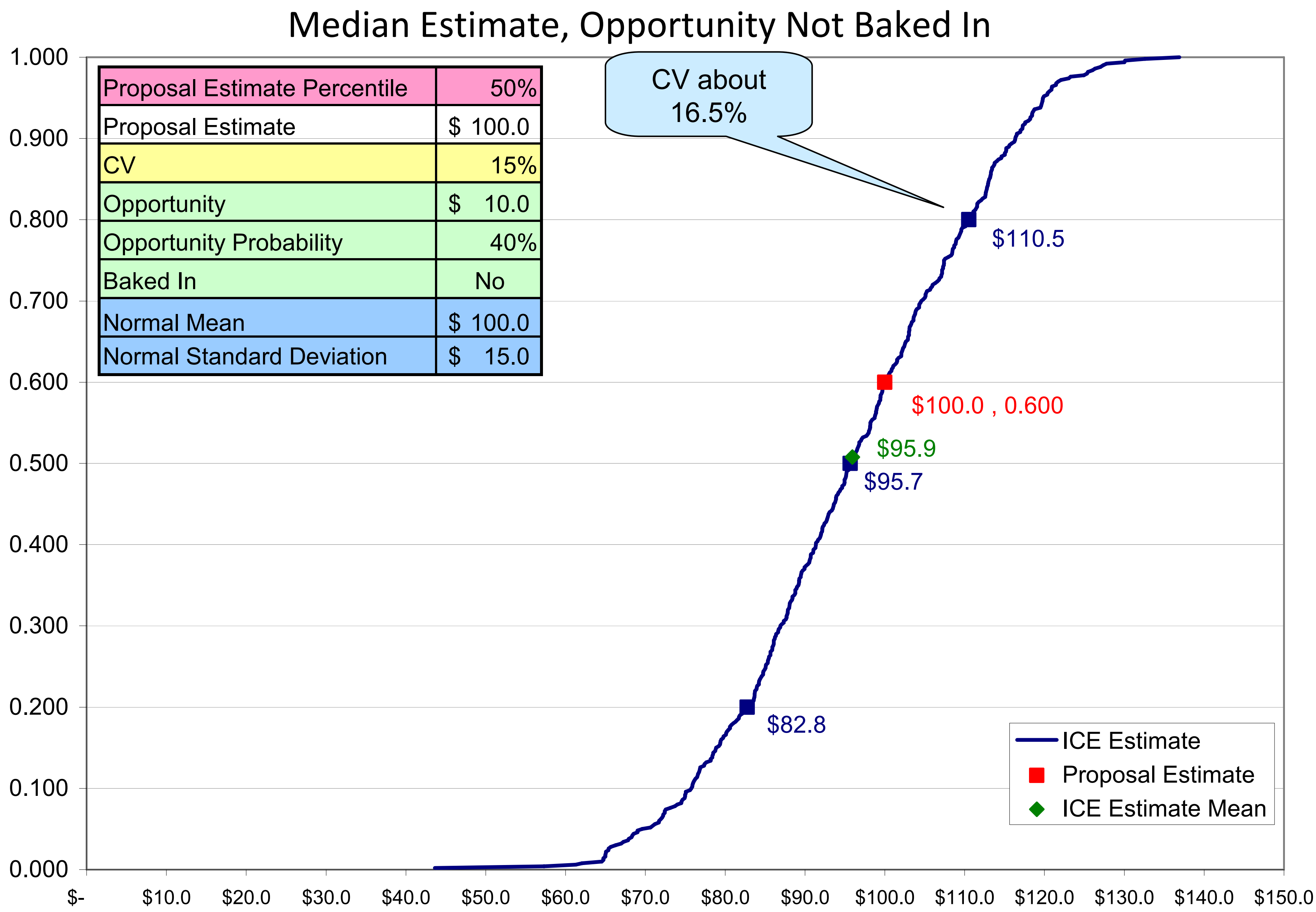
Proposal Estimate Percentile	50%
Proposal Estimate	$ 100.0
CV	15%
Opportunity	$ 10.0
Opportunity Probability	40%
Baked In	No
Normal Mean	$ 100.0
Normal Standard Deviation	$ 15.0

Wider Variance



Proposal Estimate Percentile	50%
Proposal Estimate	$ 100.0
CV	30%
Opportunity	$ 10.0
Opportunity Probability	40%
Baked In	No
Normal Mean	$ 100.0
Normal Standard Deviation	$ 30.0

Base Case

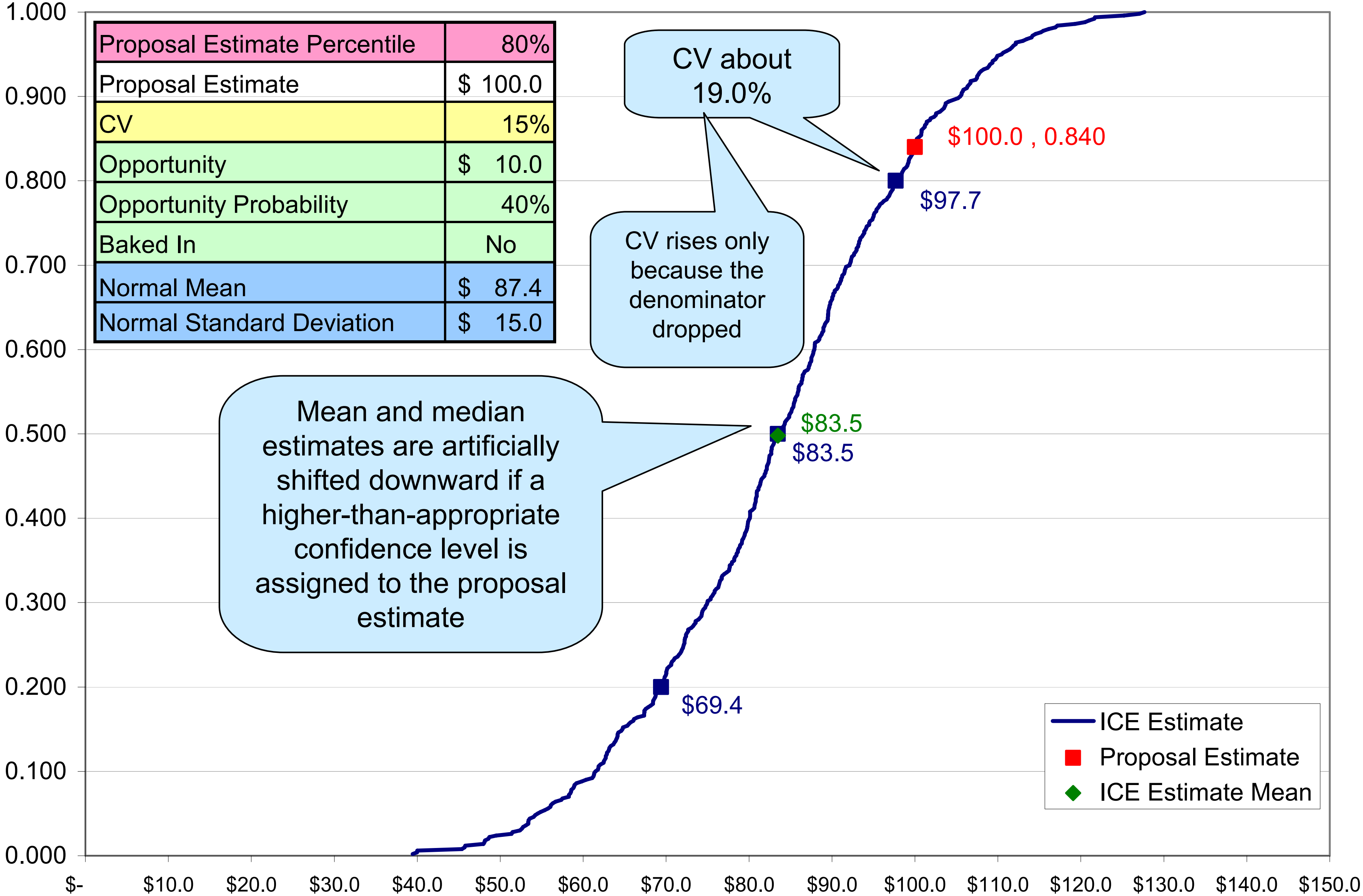


Proposal Estimate Percentile	50%
Proposal Estimate	$ 100.0
CV	15%
Opportunity	$ 10.0
Opportunity Probability	40%
Baked In	No
Normal Mean	$ 100.0
Normal Standard Deviation	$ 15.0

False Sense of Security

Higher Confidence for Base Estimate, Opportunity Not Baked In

Proposal Estimate Percentile	80%
Proposal Estimate	$ 100.0
CV	15%
Opportunity	$ 10.0
Opportunity Probability	40%
Baked In	No
Normal Mean	$ 87.4
Normal Standard Deviation	$ 15.0



Independent Cost Evaluation and the Living Risk Register

SCEA NG Panel Presentation

3 June 2009

Richard L. Coleman
Director, Independent Cost Estimation
Northrop Grumman Information Systems

Evolution of Cost Risk Assessment at NGC

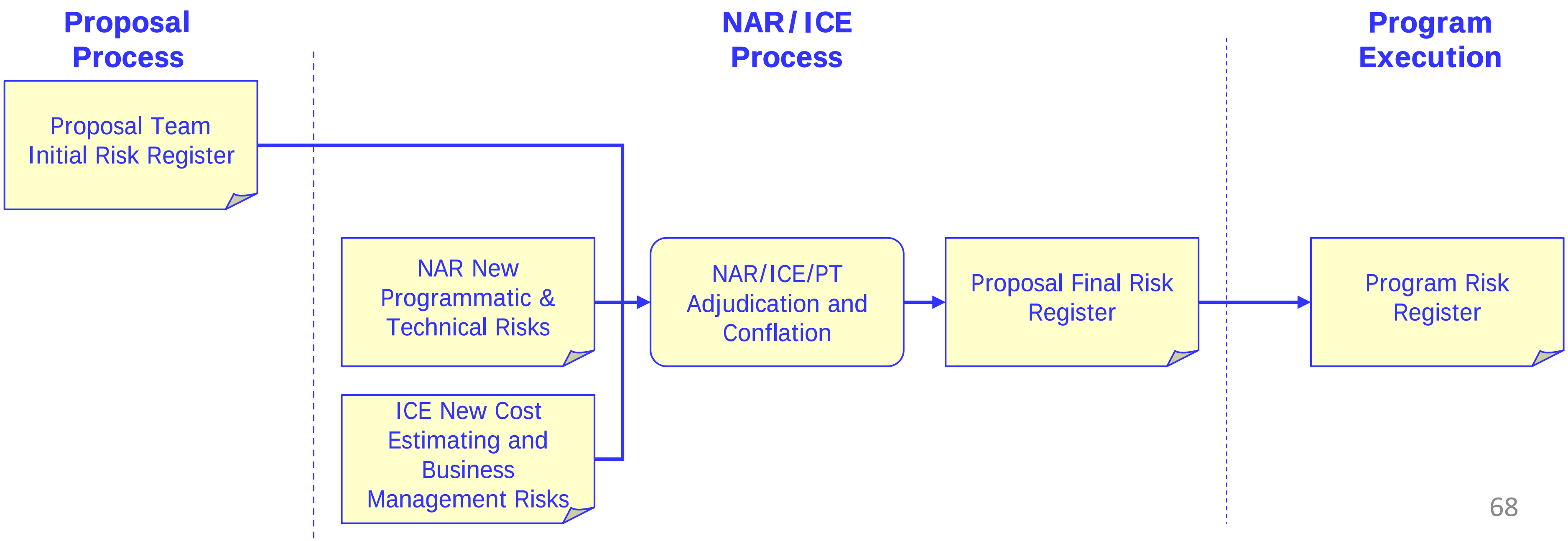
- Prior to institution of ICE Process - Comparison of Proposal Estimate to Pricing Analysis
  - Point estimates
- Independent Cost Evaluation
  - Early days – Sanity check - Primarily qualitative evaluation of risk
  - Has evolved to become much more quantitative
- Range of potential cost outcomes
- S-curve … cumulative cost probability distribution (e.g., Monte Carlo)

Independent Cost Evaluation (ICE)

- An ICE is an independent evaluation of the proposal cost estimate
  - It is not intended to be a complete Independent Cost Estimate, but may entail one, if time and resources allow
- Independence is achieved through the appointment of an ICE Team lead who is organizationally independent of the capture/proposal team
- The purpose of the ICE is to identify and quantify:
  - Significant cost risks and opportunities
  - The range of potential cost and financial outcomes
- The ICE is conducted in parallel with a Non-Advocate Review (NAR), which is an independent programmatic and technical review
- The NAR and ICE jointly review the Proposal Risk Register and the Proposal to adjust existing risks and opportunities and add new risks and opportunities
  - The Proposal Team brings their own experience and intimate knowledge of the proposed solution
  - The NAR and ICE bring a balance to the Proposal Team's natural optimism as well as corporate-wide experience on NARs and ICEs

Risk Output from the NAR and ICE

- The Proposal Team incorporates the risks identified by the NAR and ICE review into the Proposal Risk Management process
- Program startup processes then take the risk register forward into the Program Risk Register as a living document
- The result is a robust, independently reviewed Program Risk Register that combines the expertise of the Proposal Team with the broad view and independence of the review teams



Examples of Risks & Opportunities Identified

- BOE–based risks & opportunities[1]
  - Unjustified estimate
  - Unjustified adjustments to analogies
  - Computation errors (common in Learning Curves)[2]
- Business Management/Financial risks and opportunities
  - Expiring quotes
  - Quantity discount errors
  - Escalation/De-escalation
  - Warranty
  - Terms and Conditions[3]
- Technical risks
  - Inadequate/improper technical solution
  - Missing SOW
  - Schedule aggressiveness
  - Service Level Agreements
- Program Management risks
  - Missing/inadequate processes
  - Staffing difficulties

Top 5 Risks		
Risk	**Category**	**Expected Value ($)**
Expiring Supplier Hdwe Quotes(Discount Elim.-Inflation)	ICE	1,341,557
Change to Alt. Supplier	ICE	398,578
Quality: Unjustified Estimate Data Adjustments	ICE/BOE	101,016
Customer Rework to Pre-Award Effort	NAR	85,241
Pension Changes	ICE	55,929
Remaining Risks (primarily BOE-related)		494,127
Total		2,476,447

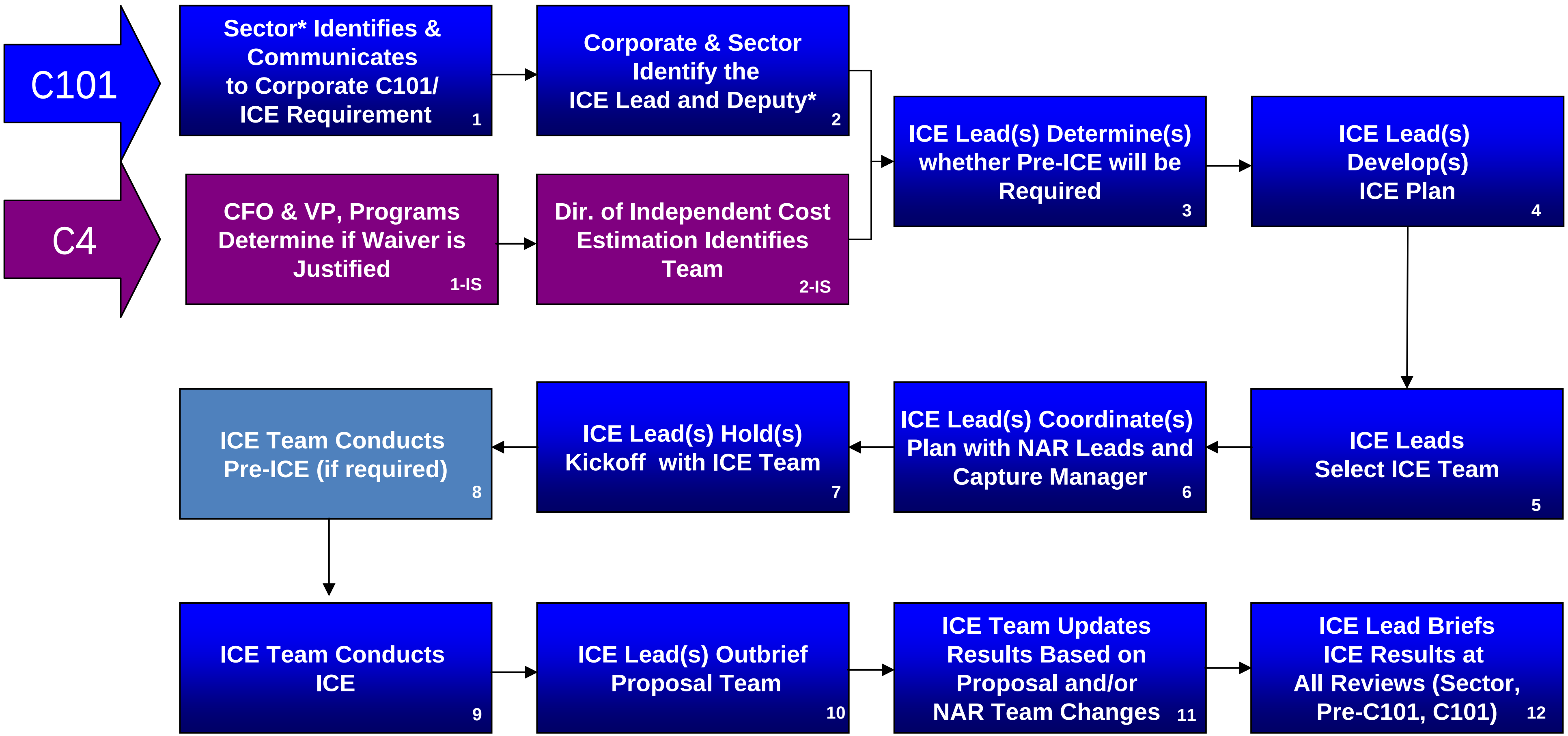
Top 5 Opportunities		
Opportunity	**Category**	**Expected Value ($)**
Expiring Supplier H/W Quotes (Add'l Discounts - Deflation)	ICE	-1,411,685
Expiring Supplier S/W Licenses (Add'l Discounts - Deflation)	ICE	-596,247
Production Learning Curve Error	ICE/BOE	-118,442
Conservative ST&E Estimate	ICE/BOE	-9,000
Conservative ILS Mgt Estimate	ICE/BOE	-309
Remaining Opportunities		0
Total		-2,135,683

Summary	
Risks	2,476,447
Opportunities	-2,135,683
MR	428,571
Total	769,335

For a more complete discussion or for examples, see:

1. RS004 – "What Percentile Are We At Now (And Where Are We Going?)" - R. L. Coleman, E. R. Druker (BAH), P. J. Braxton, B. L. Cullis, C. M. Kanick
2. RS015 – "Don't Let the Financial Crisis Happen to You: Why estimates using power CERs are likely to experience cost growth" - E. R. Druker (BAH), R. L. Coleman, P. J. Braxton
3. RS002 – "Risk-Based Return on Sales (ROS) for Proposals with Mitigating Terms and Conditions" - P. J. Braxton, R. L. Coleman, E. R. Druker (BAH), B. L. Cullis, C. M. Kanick, A. V. Bapat

C4/C101 ICE Process Flow



Corporate = VP CCPSC or Designee (Corporate Director of Pricing)
Sector = Sector CFO and VP Programs
ICE Lead(s) = ICE Team Lead (and Deputy for C101)