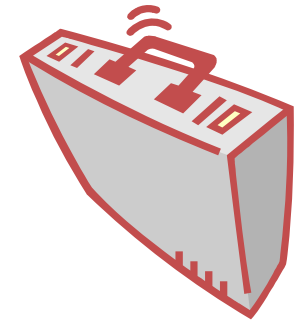


The Perils of Portability:

CGFs and CVs



“How does it feel? / To be on your own / With no direction home / Like a complete unknown / Like a rolling stone”

-Robert Zimmerman

SCEA/ISPA Conference, Thursday, June 9th, 2011

Peter J. Braxton, Richard C. Lee, Kevin M. Cincotta,
Jack Smuck, Megan Guild, Technomics, Inc.



Richard L. Coleman (retired)

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

Abstract

With the provisions of the Weapon Systems Acquisition Reform Act of 2009, or WSARA, raising the importance and visibility of risk analysis, it is an opportune time to revisit how risk benchmarks based on historical data are both derived and applied. This paper will focus on two key summary risk metrics: (1) the cost growth factor, or CGF, which captures historical cost growth for individual programs and is used collectively to capture the pervasive tendency to underestimate cost; and (2) the coefficient of variation, or CV, as a measure of the uncertainty in both historical and new estimates. In particular, it will examine the central question of how best to derive historical CVs to apply to new programs. It will build on work developing an S-Curve Tool for the Naval Center for Cost Analysis (NCCA) and recent papers by the authors on that subject: “Development and Application of CV Benchmarks” (DoDCAS 2011) and “Testing S-Curves for Reasonableness: The NCCA S-Curve Tool” (SCEA 2011).

The paper is divided into three sections. The first section briefly summarizes the basis for the historical benchmarks developed, namely Selected Acquisition Report (SAR) Summary Sheets for Department of Navy (DON) Major Defense Acquisition Programs (MDAPs). We discuss the development of a SAR database, compare SARs to Contract Data, and examine the provenance of the SAR Baseline Estimates.

The second section presents issues in SAR data analysis affecting how the benchmarks are derived. The CGF calculation is often dependent on a quantity adjustment, and we explore the Laspeyres, Paasche, and Fisher index approaches. We show the standard deviation, percentiles, and CV of the CGFs, and present an alternate method of deriving benchmark CVs using maximum likelihood estimation (MLE) regression.

The third section discusses some of the algorithms by which the benchmarks are applied in the NCCA S-Curve Tool. These include specification of baseline S-curves via various combinations of distribution parameters; the application of historical CGF and/or CV benchmarks to adjust the baseline S-curve; and solving an ancillary issue of graphing “noisy” empirical probability density functions (pdfs).

Throughout, we examine the so-called “perils of portability” and what information about the underlying historical data is maintained, distorted, or lost when we excerpt only summary metrics like CGF and CV. We illustrate these discrepancies numerically and graphically using the Milestone B and C subsets of the DON SAR data.

Too Many Ideas, Not Enough Papers

- This paper is an outgrowth of “Testing S-Curves for Reasonableness: The NCCA S-Curve Tool”
 - AKA “Getting Your S Together,” “Pimp My S-Curve,” “Where Did This S Come From?” “When the S Hits the Fan,” etc.
- Alternate title: “Cool Stuff We Learned While Building the NCCA S-Curve Tool”
 - ...and are still learning...
- Naturally builds on “Development and Application of CV Benchmarks,” Dr. Brian Flynn, et al., DoDCAS, 2011
 - Department of Navy (DON) Selected Acquisition Report (SAR) data analysis
 - North Atlantic Treaty Organization (NATO) Alliance Ground Surveillance (AGS) case study
- Some content also overflows into Training Integration Track
 - IN 06B “Probability Distributions for Risk Analysis”

Idle Thoughts on Portability

- Cost Estimators love a factor¹
 - “SE/PM is xx% of Prime Mission Equipment”
 - Eminently quotable, memorable, and portable
 - Ditto rates (e.g., hours/SLOC) and ratios (e.g., Glue Code SLOC/COTS package)
- The problem(s):
 - Why do we so readily believe a zero y-intercept?²
 - Factors convey neither quantification of uncertainty nor statistical significance
 - The two main advantages of regression-based parametrics!
 - Even CERs with standard summary statistics lack the ability to convey confidence and prediction intervals without the sum squared deviations of the Xs
 - Factors become readily divorced from their basis (or never had any to begin with!)
 - Factors are quotients, and quotients are generally statistically ill-behaved
 - OLS regression coefficients, by contrast, are distributed t
- The solution:
 - Include all the original data with “live” analysis instead of just the resultant equations
 - No longer any practical restriction of computer storage and computational capacity
 - Still need to protect sensitive data by restricting direct access and/or “disguising” programs

Like CGFs
and CVs!

1. “CERs – Rates, Factors, and Ratios” in Cost Estimating Body of Knowledge (CEBoK), Module 3 Parametric Estimating, SCEA, 2010.

2. “‘To b or Not to b’: The y-intercept in Cost Estimation,” R.L. Coleman, J.R. Summerville, P.J. Braxton, B.L. Cullis, E.R. Druker, SCEA/ISPA, 2007.

S-Curve Tool Development Epiphanies

- 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

Legend:
Blue font signals
broader
application

Previous paper “Testing the Reasonableness of S-Curves” and training session IN 06B “Probability Distributions for Risk Analysis”

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

- Data analysis

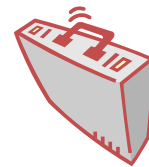
- Standard deviation vs. CV vs. pseudo-CV (s.d. divided by *median*) for historical CGFs
- CV of CGFs vs. CV of Cost
- Empirical percentile of 1.0 CGF vs. implied percentile given CV and CGF
- SAR Summary data vs. SAR data vs. Contract data

- 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

CGFs and CVs Outline

- SAR Data, or *What's our basis?*
 - SAR database, or *How do we keep all these numbers straight?!*
 - SAR Summary data vs. SAR data vs. Contract data, or *How much did the program really grow?*
 - Provenance of Baseline Estimates, or *Dispelling the myth that we never budget to the ICE*
- SAR Data Analysis, or *Deriving our benchmarks*
 - Quantity-adjustments, or *Who the hell are Laspeyres, Paasche, and Fisher?!*
 - Standard deviations and percentiles, or *Perils of Portability part 1*
 - CGF and CV derivations, or *Perils of Portability part 2*
- NCCA S-Curve Tool algorithms, or *Applying our benchmarks*
 - Specification of distributions, or *I got to use the quadratic formula at work today!*
 - Historical adjustments (CGF and CV), or *Perils of Portability part 3*
 - Graphing of probability distributions, or *Why does my PDF look like an EKG?!*
- Open Questions and Future Work, or *Where do we go from here?*



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

SAR Data

- SAR database
 - Database contents summary (n = 100, DON MDAPs)
 - Initial hypothesis testing results
 - Transition from Excel to Access
 - Data integrity gains
 - Data anomalies
 - Coding a database median function!
- SAR Summary data vs. SAR data vs. Contract data
 - F/A-18E/F Super Hornet example
 - RAND data and Development / Procurement split
- Provenance of Baseline Estimates
- Future SAR data collection efforts

SAR Database – DON Programs

Source

- SAR Summary Sheets
- Total program acquisition cost
 - R&D, procurement, MILCON
- Tied to acquisition milestones
 - Planning Estimate (PE) for MS A
 - Development Estimate (DE) for MS B
 - Production Estimate (PdE) for MS C
 - Historically, equivalent to milestones I, II, and III
- Base-Year \$ and Then-Year \$
- From 1985 to 2009

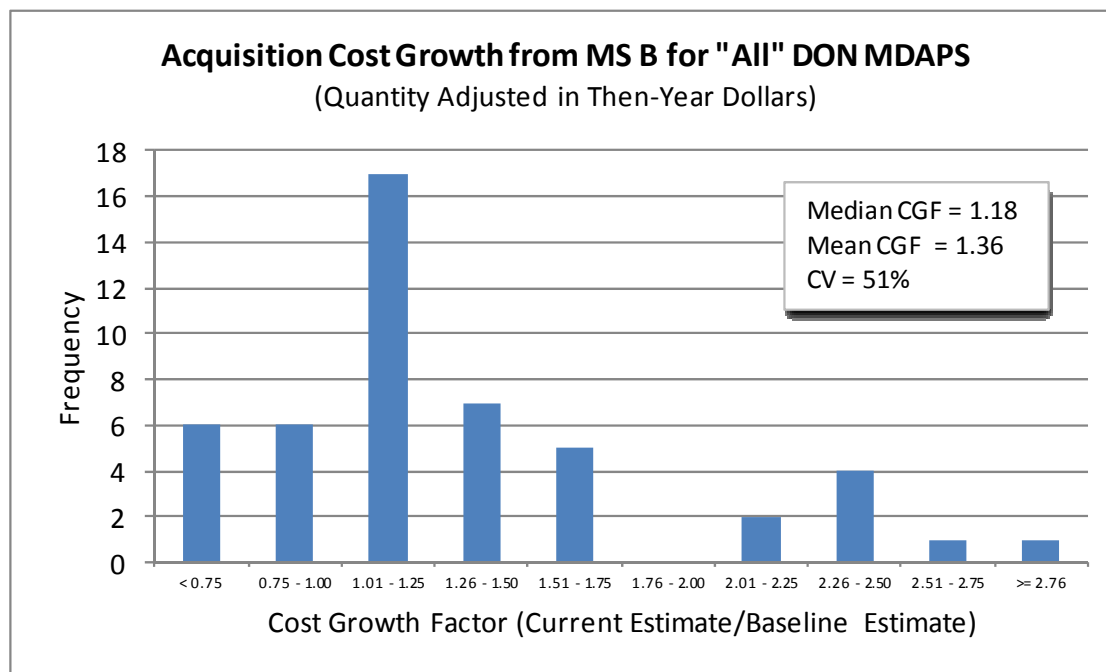
Focus

- DON MDAPS only
- 83 programs, 100 observations
- Baseline Estimates date from 1969 to 2003
 - Mostly completed programs, but a few on-going such as LPD 17 and LCS
 - Ships, submarines, missiles, and aircraft predominate
 - Excludes notables such as A-12 and Presidential Helicopter

MS B: All DON MDAPs

- Distribution skewed to right
- Adjustments for changes in quantity and inflation decrease values of CGFs and CVs
- CVs sensitive to outliers
 - E.g., removing Harpoon decreases quantity-adjusted TY\$ CV to 0.45
 - 2nd oldest datum (1970 baseline)

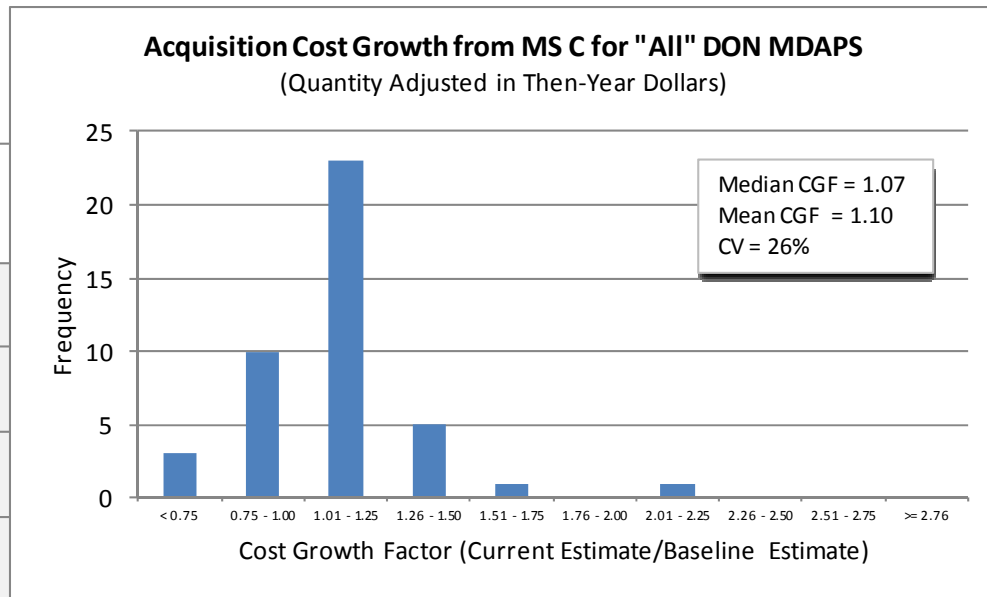
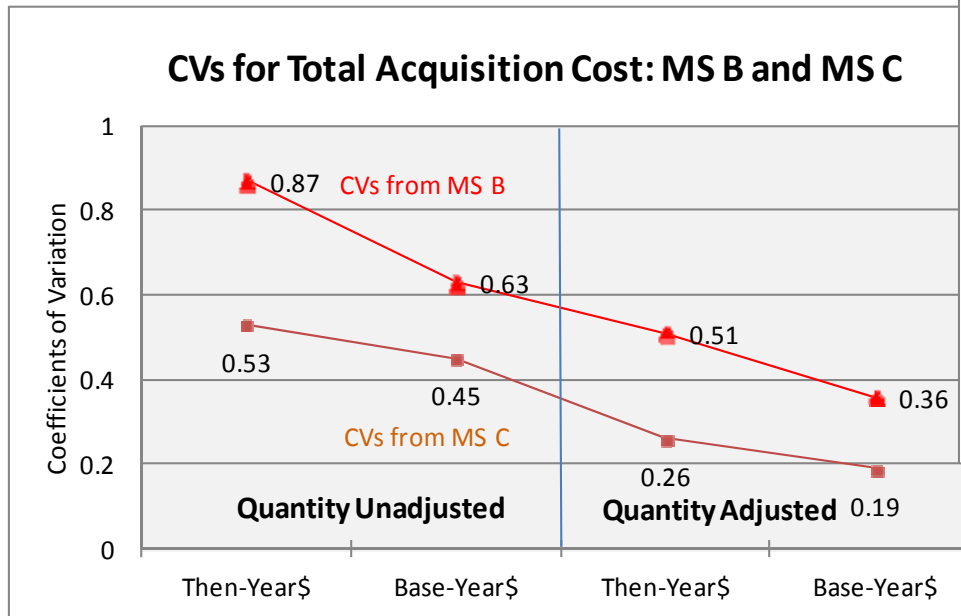
Cost Growth Factors & CVs for All DON MDAPs at MS B for 1969 & Later; n = 50				
Statistics	(Without Qty Adjustment)		(Quantity Adjusted)	
	Base-Year\$	Then-Year\$	Base-Year\$	Then-Year\$
Mean	1.48	1.84	1.23	1.36
Standard Deviation	0.94	1.60	0.44	0.69
CV	0.63	0.87	0.36	0.51



MS C: All DON MDAPs

- CVs uniformly lower
- Cost growth factors less compared to **DE** values
 - Mean (1.10 versus **1.36**)
 - Median (1.07 versus **1.18**)
 - Similar trend for the 9 programs with both DEs and PdEs
- Distribution less skewed

Cost Growth Factors & CVs for All DON MDAPs at MS C for 1969 & Later; n = 43					
Statistics		(Without Qty Adjustment)		(Quantity Adjusted)	
		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



3. "Development and Application of CV Benchmarks," Brian Flynn, Paul Garvey, Peter Braxton, Richard Lee, DoDCAS, 2011.

Summary of Findings

Conjectures

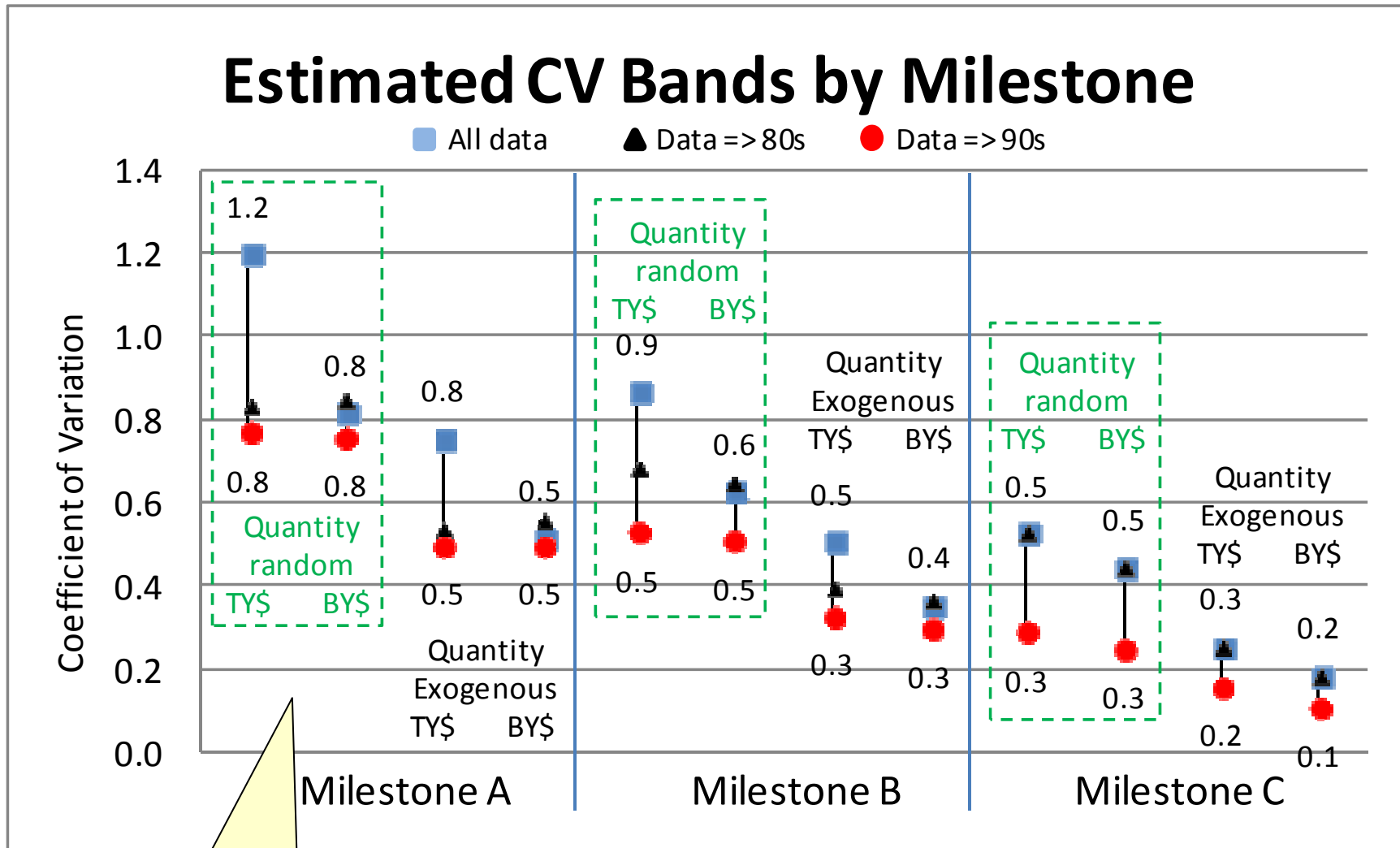
- Estimation Consistency
 - CVs from ICEs jibe with acquisition experience
 - Ad hoc observation suggests underestimation of CVs, at times, in the international defense community
- Decline During Acquisition
 - CVs decrease throughout acquisition lifecycle
 - Strongly supported (MS B to MS C)
- Platform Homogeneity
 - CVs equivalent for aircraft, ships, and other platform types
 - Strongly supported, especially for MS B

Conjectures

- Adjustment Decline
 - CVs decrease when adjusted for changes in quantity and inflation
 - Strongly supported
- Secular Invariance
 - CVs steady long-term
 - Rejected
 - Evidence of secular decline
 - However, small sample sizes lessen confidence

3. "Development and Application of CV Benchmarks," Brian Flynn, Paul Garvey, Peter Braxton, Richard Lee, DoDCAS, 2011.

SAR CVs by Milestone and Era



Note: Milestone A numbers are "by analogy"

SAR Data – Initial Excel Format

- Mix of programs (e.g., DDG 51) and estimates (i.e., PE, DE, PdE)
- Mix of source data and calculated values
- Quantity-adjusted indices calculated in additional rows instead of additional columns
 - Some ability to detect anomalies
 - For Phoenix AIM-54C, the quantity adjustment was paradoxically greater than the current estimate, leading to a negative Laspeyres index!
- Blank and near-blank rows cause problems for sorting and filtering

Name of Program	Base Year	Baseline Type	Type System	Baseline Estimate			Cost Growth Multiplier (Without Qty Adjustment)		Cost Growth Multiplier (Quantity Adjusted)	
				Base-Year\$	Then-Year\$	Quantity	Base-Year\$	Then-Year\$	Base-Year\$	Then-Year\$
DDG-51 Destroyers (Arleigh Burke Class)	1987	PdE	Ship	\$16,953.7	\$20,117.5	23	3.34	4.00	1.22	1.21
		PdE							1.60	1.70
		PdE							1.40	1.44
DDG-51 Destroyers (Arleigh Burke Class)	1984	DE	Ship	\$12,454.4	\$18,479.6	18	4.04	4.35	1.13	1.17
		DE							1.47	1.63
		DE							1.29	1.38

Before

SAR Database (Access)

Filters:

Commodity:

- Aircraft
- Electronics
- Gun
- Missile
- Satellite
- Ship
- Torpedo
- Vehicle
- Submarine

Clear

Baseline Type (Milestone):

DE (MS B)

Clear

Year (>=):

Clear

Program Name (includes):

Clear

Data & Calculations:

View Data

Include Calculations

View Stats:

Fisher

Use TY Dollars

Allows grouping of commodities (e.g., Ships and Subs)

Allows filtering by era (e.g., ≥ 1980)

Displays key statistics and exports query results to Excel

Required VBA custom function!

μ	σ	CV	Median	n
1.36	0.69	0.51	1.18	50

- Relational database, stores information once and only once
- Stores source data (e.g., BE, CE, QΔ, BY\$/TY\$), calculates derived data (e.g., CGF, CV)
- Flexible ad hoc queries
- Data integrity: Detected several minor errors in Excel analysis (e.g., inadvertent transposition of Laspeyres, Paasche, and Fisher indices)

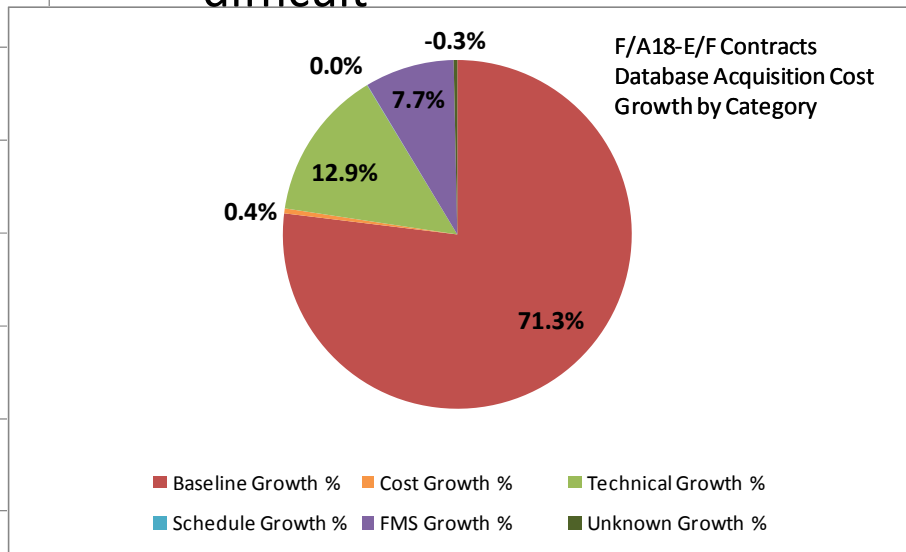
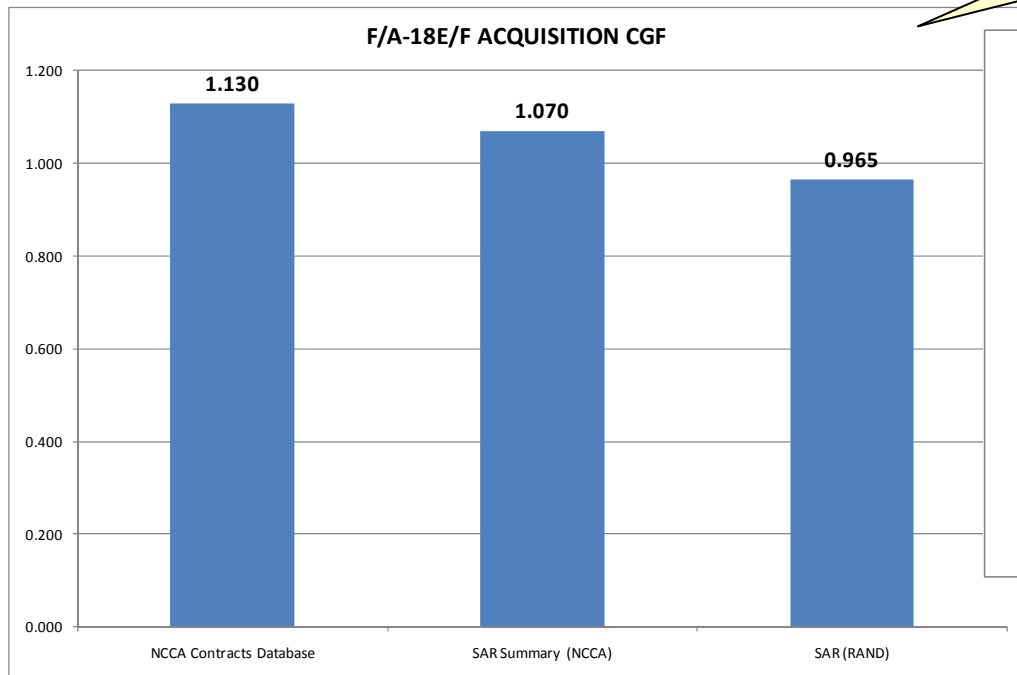
ExcelRow	Program	BaseYear	BaselineType	Commodity	BaseYear\$Initial	ThenYear\$Initial
1	DDG-51 Destroyers (Arleigh Burke Class)	1987	PdE	Ship	\$16,953.70	\$20,117.50
2	DDG-51 Destroyers (Arleigh Burke Class)	1984	DE	Ship	\$12,454.40	\$18,479.60
3	DDG-51 Destroyers (Arleigh Burke Class)	1981	PE	Ship	\$6,443.50	\$10,953.50
4	DDG-1000 Destroyers (Zumwalt Class)	2005	DE	Ship	\$31,547.90	\$36,296.30
5	CVN-78 Aircraft Carriers (Gerald R. Ford Class)	2000	DE	Ship	\$28,701.20	\$36,082.10
6	CVN-77 (1 ship) from CVN-68 Aircraft Carriers (Nimitz Class)	1995	PdE	Ship	\$4,557.10	\$5,540.80
7	LPD-17 Amphibious Transport Dock (San Antonio Class)	1996	DE	Ship	\$9,018.10	\$10,761.80

After

SAR Data vs. Contract Data

- F/A-18E/F Super Hornet example
- NCCA Contracts database quantity 472, SAR quantity 515
- NCCA Contracts database reflects growth on contract
 - Removed Baseline and FMS growth to reflect only unanticipated changes
 - Includes *all* program contracts and subcontracts, including those not capture by SAR
 - Reflects actuals, not EAC, though Production values adjusted *pro rata*
- SAR Summary Total Acquisition only
- RAND SAR data BY only

Achieving an “apples-to-apples” comparison is difficult



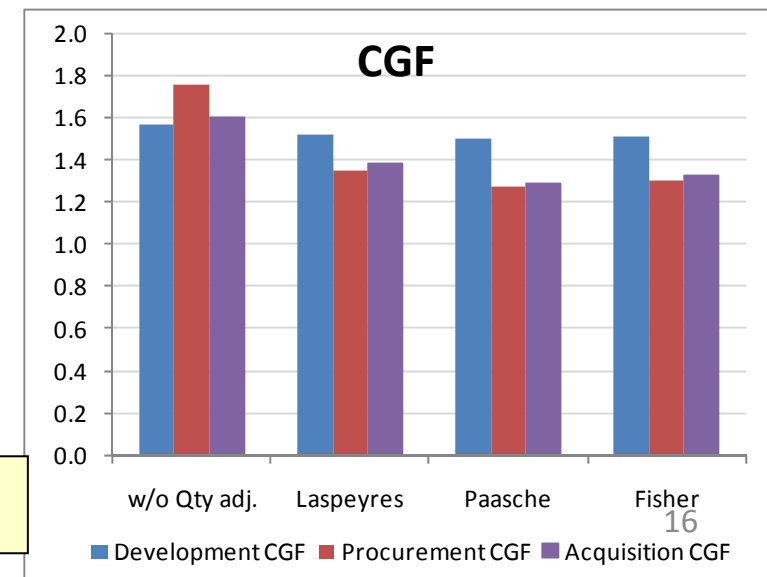
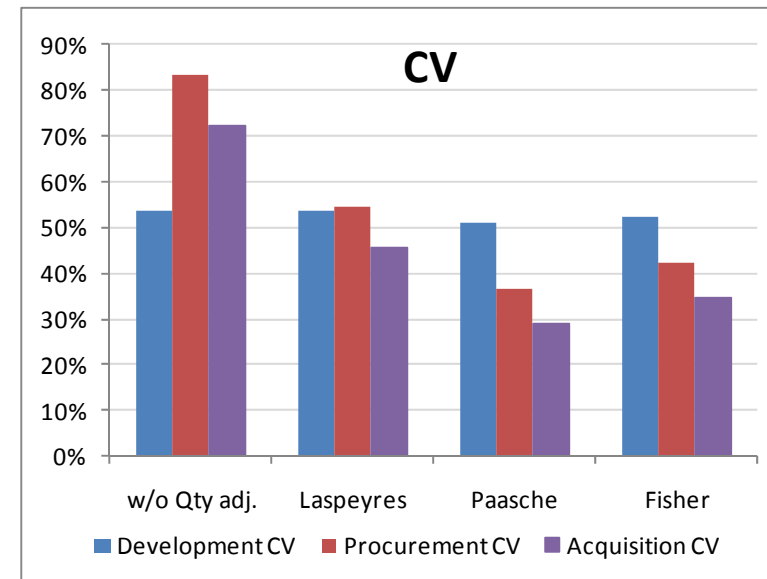
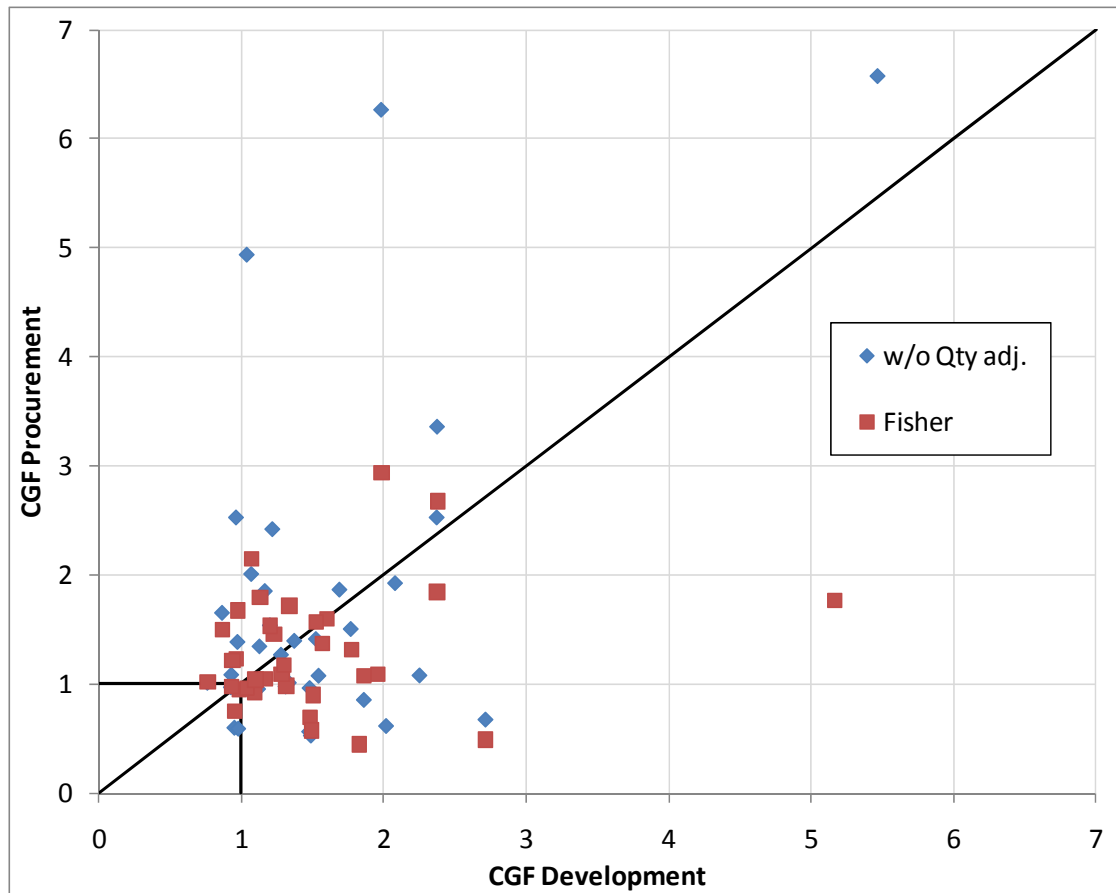
SARs vs. SAR Summary Sheets

- NCCA analysis of SARs for DON MDAPs based on SAR Summary Sheets
 - Total Acquisition only
- S-Curve Tool allows option of historical benchmarks broken out by Development and Procurement
 - By Phase (R&D / Production) or Appropriation (RDT&E / Procurement)?
 - Many users demand this, but Occam's Razor might say (initial) comparisons are best done at Total Acquisition level
 - One solution is to add SAR detail data to the database
 - Corresponding DON guidance TBD
- RAND study⁴ offers Development / Procurement split
 - Mostly non-overlapping set of MDAPs (primarily AF)

4. *Sources of Weapon System Cost Growth (MG-670)*, Joseph G. Bolten, Robert S. Leonard, Mark V. Arena, Obaid Younossi, Jerry M. Sollinger, RAND Project Air Force, 2008.

RAND Study Quick-Look

- Little correlation between Development and Procurement CGFs
- Seems reasonable to use Acquisition CVs (lower bound) and CGFs (middle) as a proxy



4. Sources of Weapon System Cost Growth (MG-670), Joseph G. Bolten, Robert S. Leonard, Mark V. Arena, Obaid Younossi, Jerry M. Sollinger, RAND Project Air Force, 2008.

Provenance of Baseline Estimates

- Comparison drawn for programs for which data were available
 - 6 ACAT ID programs (OSD CAIG ICE)
 - 4 ACAT IC programs (NCCA ICE)
- With one exception, SAR BE, POE, and ICE were quite close
 - Attributed in part to Reconciliation process

Analysis of Deltas									
SAR BE		Program Office's Acquisition Cost Estimate		ICE (CAIG for ID; NCCA for IC)		Ratio of POACE to SAR BE	Ratio of POACE to SAR BE	Ratio of ICE to SAR BE	Ratio of ICE to SAR BE
in BY\$	in TY\$	in BY\$	in TY\$	in BY\$	in TY\$	in BY\$	in TY\$	in BY\$	in TY\$
\$2,877	\$3,093	\$2,817	\$3,032	\$3,130		0.98	0.98	1.09	
\$4,123	\$4,310	\$4,123		\$4,104		1.00		1.00	
\$45,633	\$71,081	\$45,500		\$47,400		1.00		1.04	
	\$8,636		\$8,400		\$8,580		0.97		0.99
\$26,494	\$31,429	\$24,490		\$26,810		0.92		1.01	
\$31,548	\$36,296	\$32,800		\$39,100		1.04		1.24	
\$10,627	\$11,425	\$10,727				1.01			
\$43,490	\$46,826	\$43,000				0.99			
\$4,263	\$4,890	\$4,245		\$4,349		1.00		1.02	
\$2,977	\$3,290	\$3,019	\$3,284		\$3,505	1.01	1.00		1.07
						Means =	0.99	0.98	1.07
								1.03	1.03 without outlier

SAR Data Analysis

- Quantity-adjustments:
 - Laspeyres, Paasche, and Fisher \$ and indices
- Standard deviations and percentiles
 - Standard deviation vs. CV vs. (median-based) pseudo-CV
 - Empirical percentile of 1.0 CGF vs. implied percentile given CV and CGF
- CGF and CV derivations
 - CV of CGFs and pseudo-iid mental models
 - Confidence intervals for CV (normal or lognormal assumption)
 - CV of Cost and CE vs. BE graphs
 - White test for heteroskedastic error terms
 - MLE regressions, error functional forms, and the size effect

Quantity Adjustments

- Quantity viewed as either:
 - Random (no adjustment); or
 - Exogenous (adjustment)
- Three possible quantity adjustments:

Method	Description	Baseline \$	Current \$	CGF
Laspeyres	Adjust current estimate to reflect baseline quantities	BE	$CE - Q\Delta$	$\frac{CE - Q\Delta}{BE}$
Paasche	Adjust baseline estimate to reflect current quantities	$BE + Q\Delta$	CE	$\frac{CE}{BE + Q\Delta}$
Fisher	"Split the difference" between baseline and current quantities	$\sqrt{BE \cdot (BE + Q\Delta)}$	$\sqrt{(CE - Q\Delta) \cdot CE}$	$\sqrt{\frac{(CE - Q\Delta) \cdot CE}{BE \cdot (BE + Q\Delta)}}$

3. "Development and Application of CV Benchmarks," Brian Flynn, Paul Garvey, Peter Braxton, Richard Lee, DoDCAS, 2011.

Geometric mean is used for multiplicative comparisons

Cost Growth Calculations

Example: CG-47 Class (MS B)



- Baseline Estimate (BE) of 1978
 - 16 ships at \$9.01B (BY\$) and \$14.08B (TY\$)
- Current Estimate (CE) of 1992
 - 27 ships at \$14.11B (BY\$) and \$23.28B (TY\$)
 - Deltas in BY\$
 - \$5.10B total & \$5.49B quantity
 - Deltas in TY\$
 - \$9.20B total & \$11.74B quantity

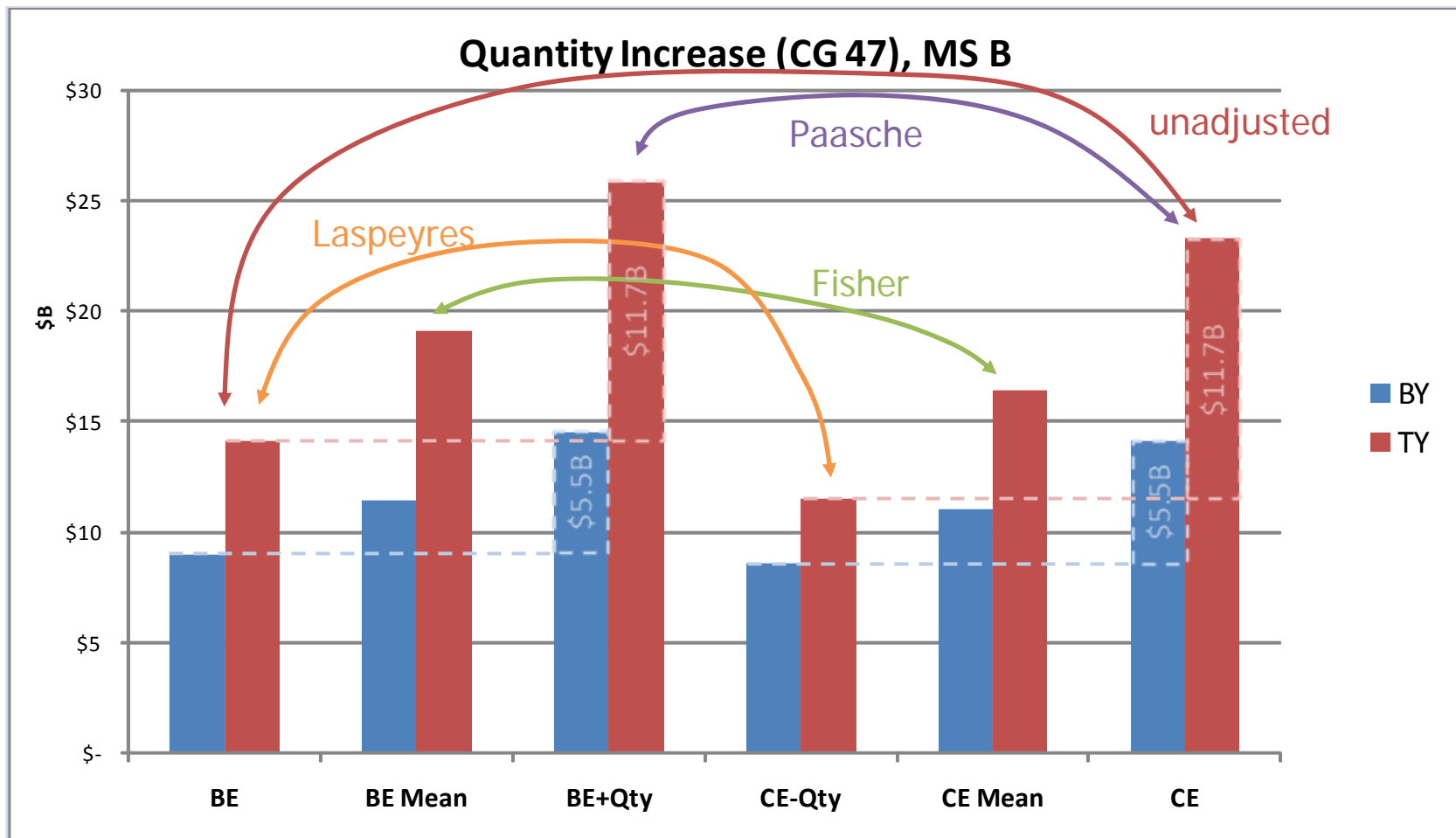
Paasch

Cost Growth Factors

- Unadjusted for quantity Δ
 - Then-year dollars
 - $\$23.28B / \$14.08B = 1.65$
 - Base-year dollars
 - $\$14.11B / \$9.01B = 1.57$
- Adjusted for quantity Δ , using OSD methodology
 - Then-year dollars
 - $\$23.28B / (\$14.08B + \underline{\$11.74B}) = 0.90$
 - Base-year dollars
 - $\$14,11B / (\$9.01B + \underline{\$5.49B}) = 0.97$

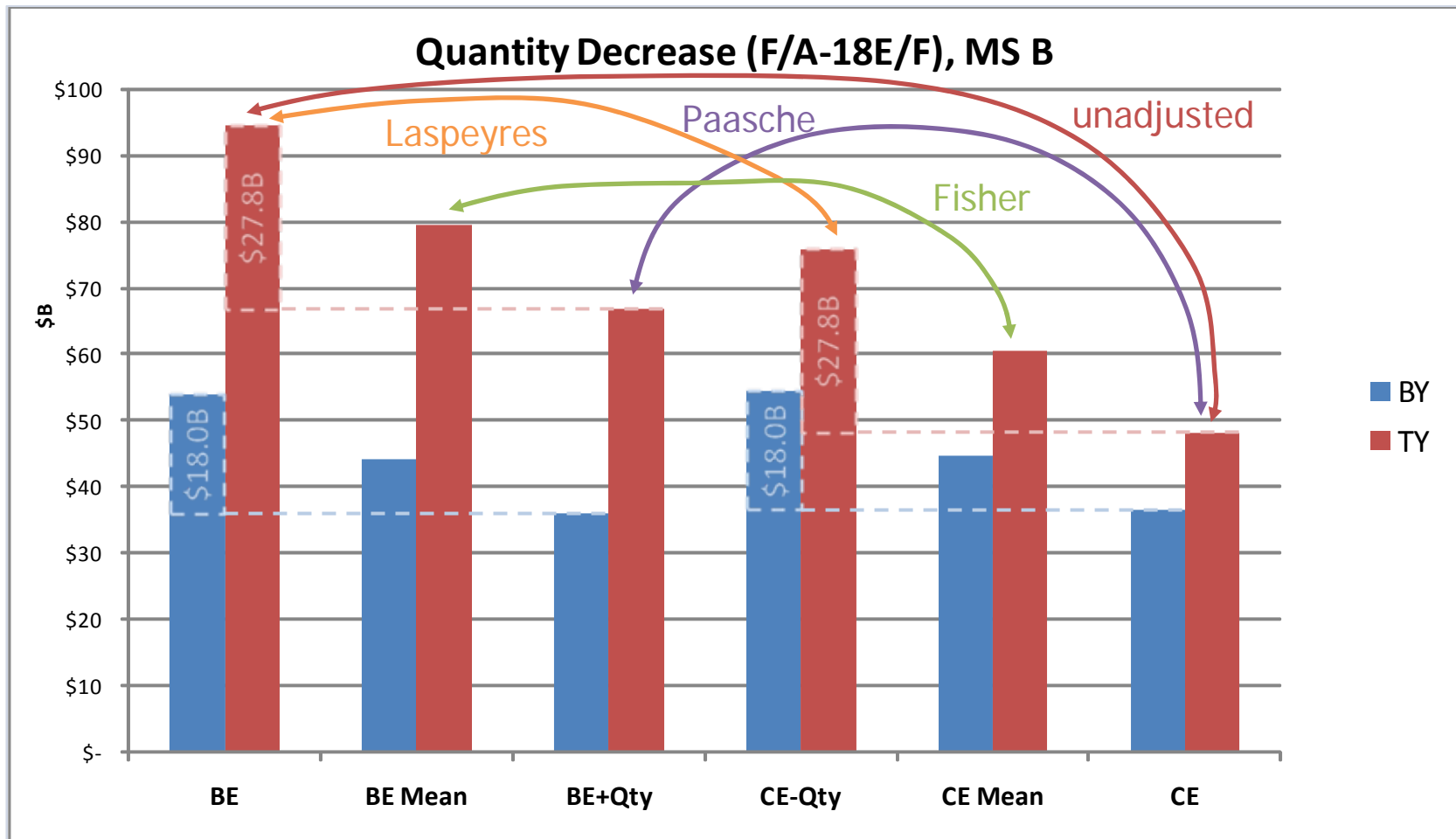
3. "Development and Application of CV Benchmarks," Brian Flynn, Paul Garvey, Peter Braxton, Richard Lee, DoDCAS, 2011.

CGF Calculations Illustrated



CG 47	BE	BE Mean	BE+Qty	CE-Qty	CE Mean	CE	Qty
BY	\$ 9,013.7	\$ 11,434.4	\$ 14,505.1	\$ 8,620.2	\$ 11,029.3	\$ 14,111.6	\$ 5,491.4
TY	\$ 14,083.5	\$ 19,070.2	\$ 25,822.5	\$ 11,537.9	\$ 16,388.0	\$ 23,276.9	\$ 11,739.0
BY	1.57			0.96	0.96	0.97	
TY	1.65			0.82	0.86	0.90	
	unadjusted			Laspeyres	Fisher	Paasche	
Qty	16		27	16		27	11

CGF Calculations Illustrated



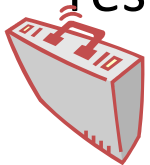
F/A-18E/F	BE	BE Mean	BE+Qty	CE-Qty	CE Mean	CE	Qty
BY	\$ 53,959.6	\$ 44,068.9	\$ 35,991.1	\$ 54,538.2	\$ 44,659.2	\$ 36,569.7	\$ (17,968.5)
TY	\$ 94,583.0	\$ 79,503.6	\$ 66,828.2	\$ 75,846.2	\$ 60,394.9	\$ 48,091.4	\$ (27,754.8)
BY	0.68			1.01	1.01	1.02	
TY	0.51			0.80	0.76	0.72	
	unadjusted			Laspeyres	Fisher	Paasche	
Qty	1000		515	1000		515	-485

CV – A Means to an End

- We want to answer questions such as:
 - What standard deviation should I put around my cost estimate?
 - At what percentile is my point estimate?
- CV is a convenient, **portable** quantity that helps us answer these questions
 - *Under certain distributional assumptions*
- First we divide by the baseline estimate to obtain CGF
 - Attempt to normalize programs of different sizes and outcomes
 - We'll return to this shortly as **Peril of Portability #2**
- The standard deviation of CGFs is our key proxy for the spread of our estimates
 - We can take it as is, or express it as a multiple of the mean (CV) or median (“pseudo-CV”)

Determination of Standard Deviation

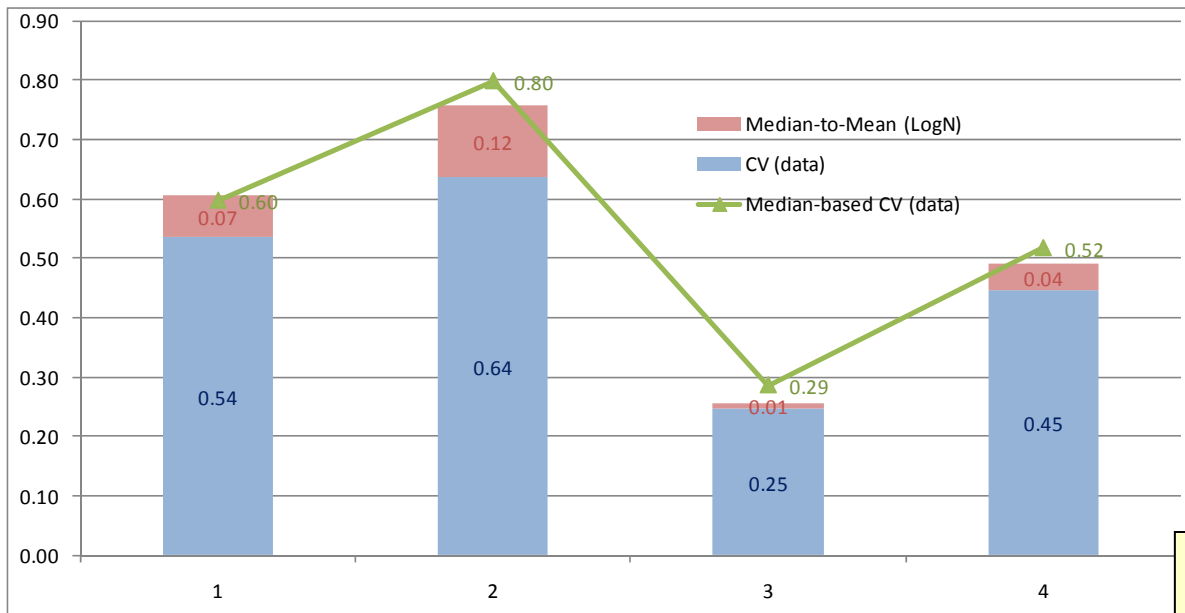
- Standard Deviation is CV times Mean
 - CV derived from standard deviation of CGF divided by mean of CGF
- A CV-like measure (“pseudo-CV”) could be derived from SAR data:
 - Standard deviation of CGF divided by median of CGF
- Applying lognormal distribution yields similar but not identical results



Put another way, the empirical mean-to-median ratio is different from that implied by the CV [Peril #1A]

$$\sqrt{1 + CV^2}$$

- Ships and Submarines example below



Blue = CV of CGFs
 Red = Addl delta to account for median-to-mean factor for lognormal (see next slide)

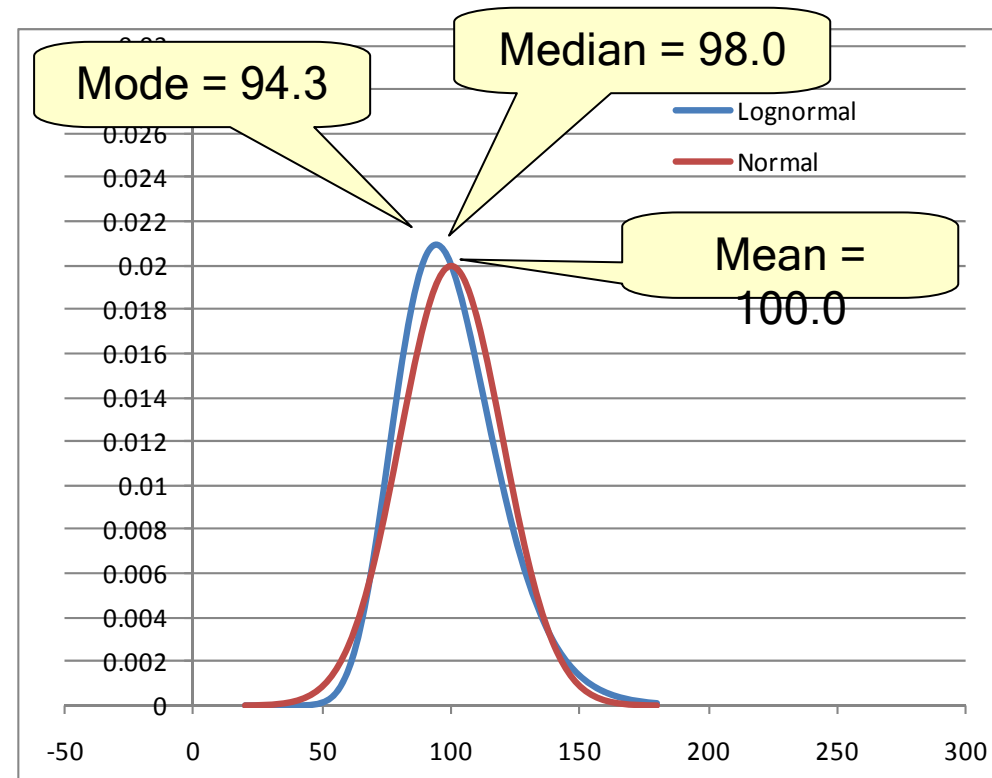
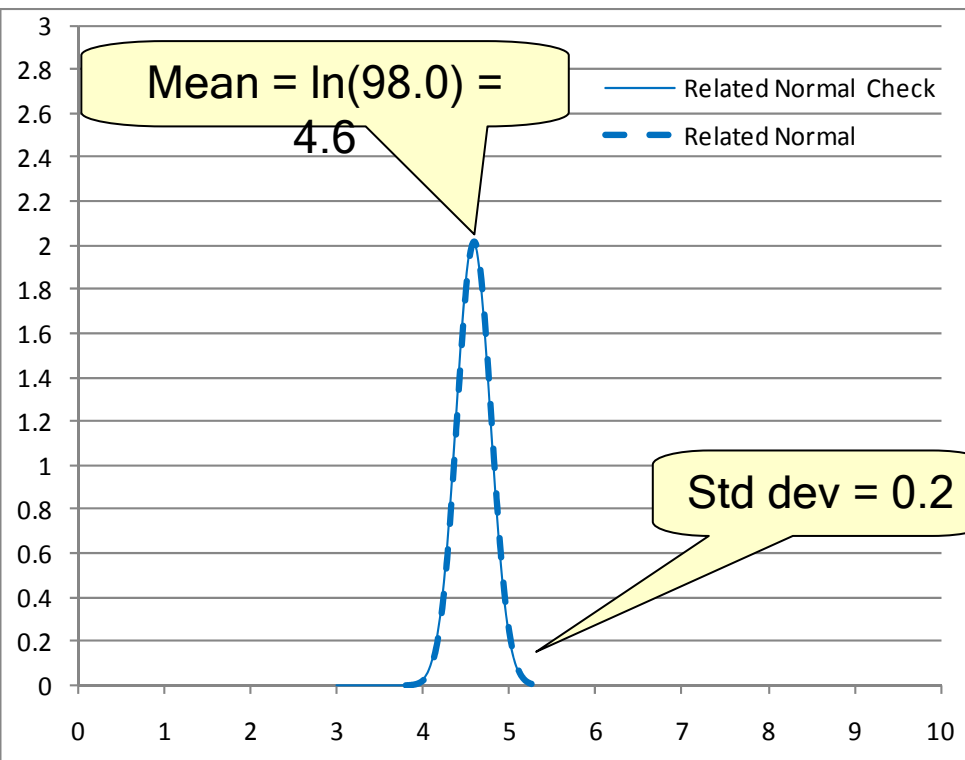
Green = Pseudo-CV of CGFs

5. IN 06B “Probability Distributions for Risk Analysis,” Peter Braxton, SCEA/ISPA, 2011.

Related Normal Example

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

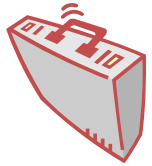
CV (lognorm)	mode shift factor	mean shift factor	percentile of mean
10%	0.990	1.005	52.0%
20%	0.962	1.020	53.9%
30%	0.917	1.044	55.8%
40%	0.862	1.077	57.6%
50%	0.800	1.118	59.3%



5. IN 06B "Probability Distributions for Risk Analysis," Peter Braxton, SCEA/ISPA, 2011.

CGFs and Percentile of 1.0

- The Percentile of 1.0 for historical CGF data is of interest
 - Roughly speaking, the proportion of programs that come in at or under budget



Empirical percentile of 1.0 is different from percentile of 1.0 implied by CGF and CV [Peril #1B]

- At what percentile would the “point estimate” have to be so that multiplying by the CGF would bring it up exactly to the mean
 - Parametric, assume normal or lognormal (different answers)
 - Tables for related conversion of percentile to CGF in Backup
- Interesting side note: There is no universal definition for Percentile⁶!
- Excel’s PERCENTILE() function does not even use the NIST-recommended method!
 - We favor linear interpolation between closest ranks
 - Thankfully, all methods give about the same answer for large datasets

α_{PE} of
eSBM

6. <http://en.wikipedia.org/wiki/Percentile>

Confidence Interval for CVs

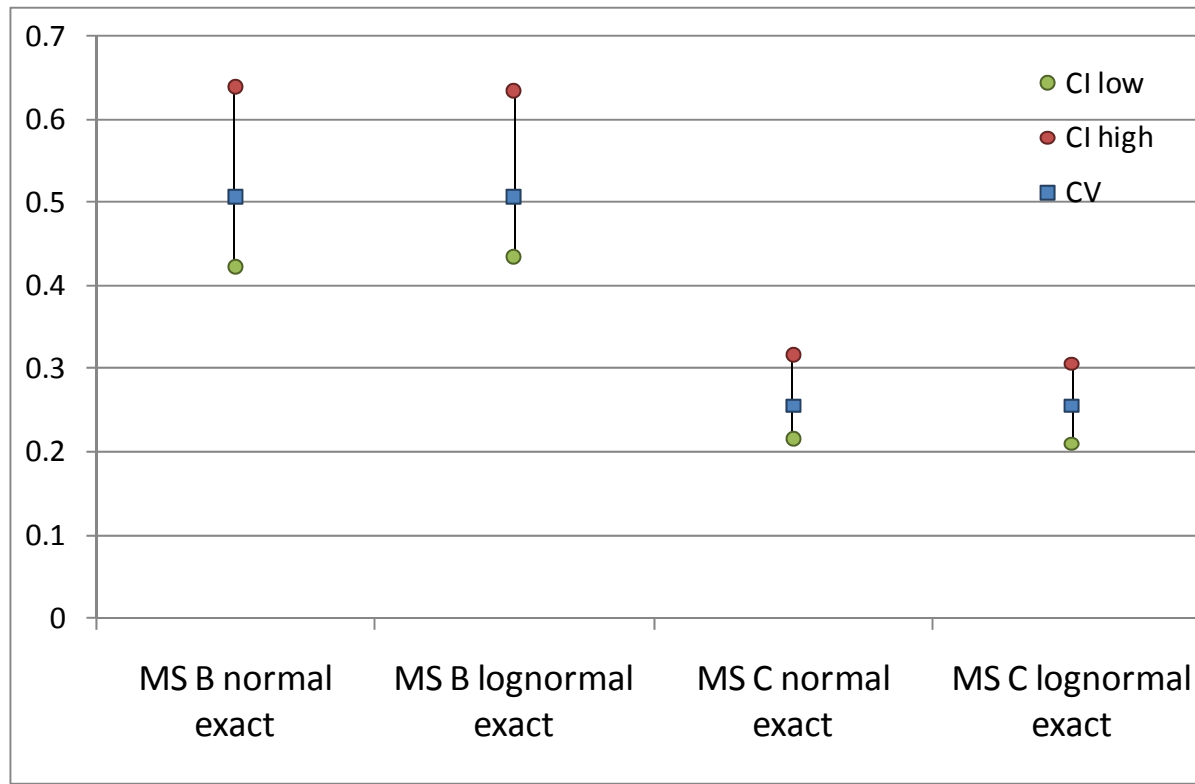
- Since we are using CVs as a benchmark, we'd like to know how close the true CV is to our sample CV
 - If the CGFs are in fact normally or lognormally distributed (as in Mental Model #2), the known confidence interval (CI) for CV can be applied
 - It is not immediately clear based on the previous histograms that this is true for the DON SAR data – see subsequent Kolmogorov-Smirnov test results
- Cited paper gives exact and approximate formulae
 - Normal or lognormal distributions
 - Exact formula involves non-central t distribution!
 - Not readily calculated in Excel
 - Approximate formula uses chi square distribution instead
 - Web calculator is also available
 - Requires only α , n , x -bar, and s
- Issue: For low confidence levels, the approximate confidence interval does not contain the mean!
 - CI formulae appeal to asymptotic properties, but we often have small samples

<http://www1.fpl.fs.fed.us/covnorm.html>

<http://www1.fpl.fs.fed.us/covln.html>

7. "Confidence Bounds for Normal and Lognormal Distribution Coefficients of Variation," Steve Verrill, Forest Products Laboratory, Research Paper FPL-RP-609, 2003.

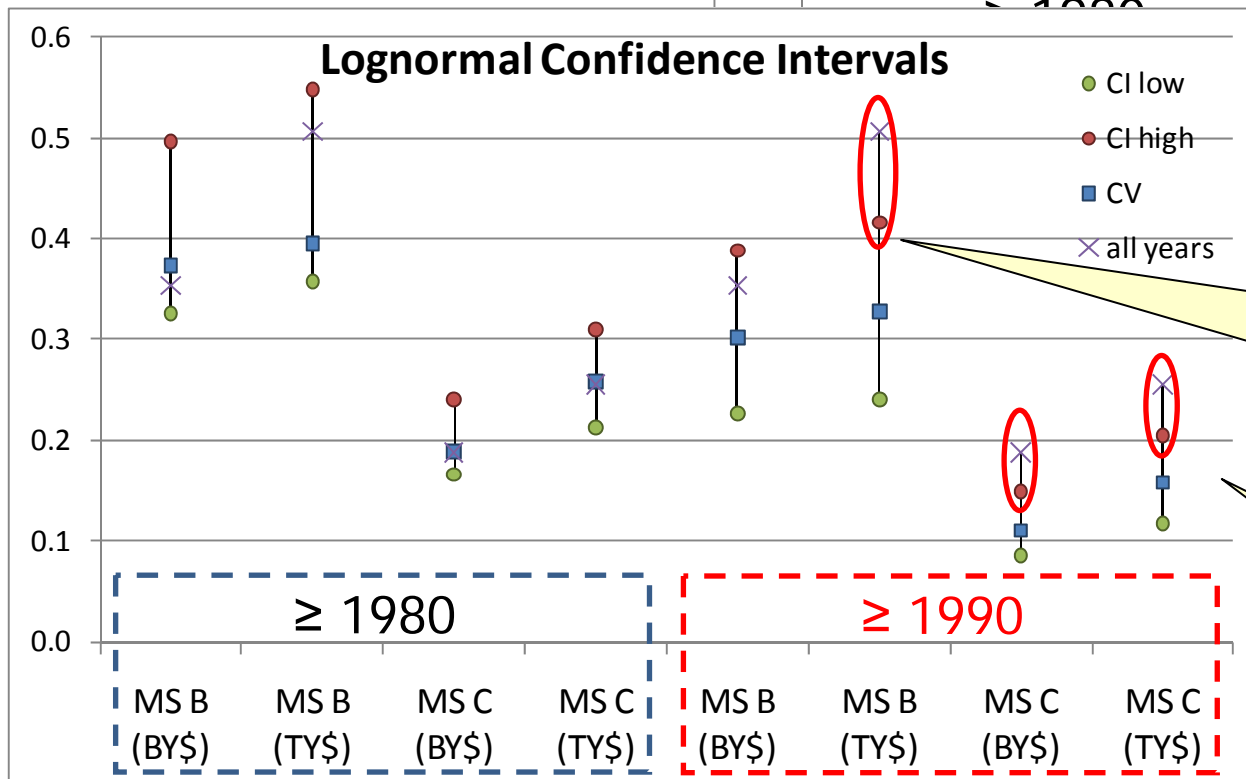
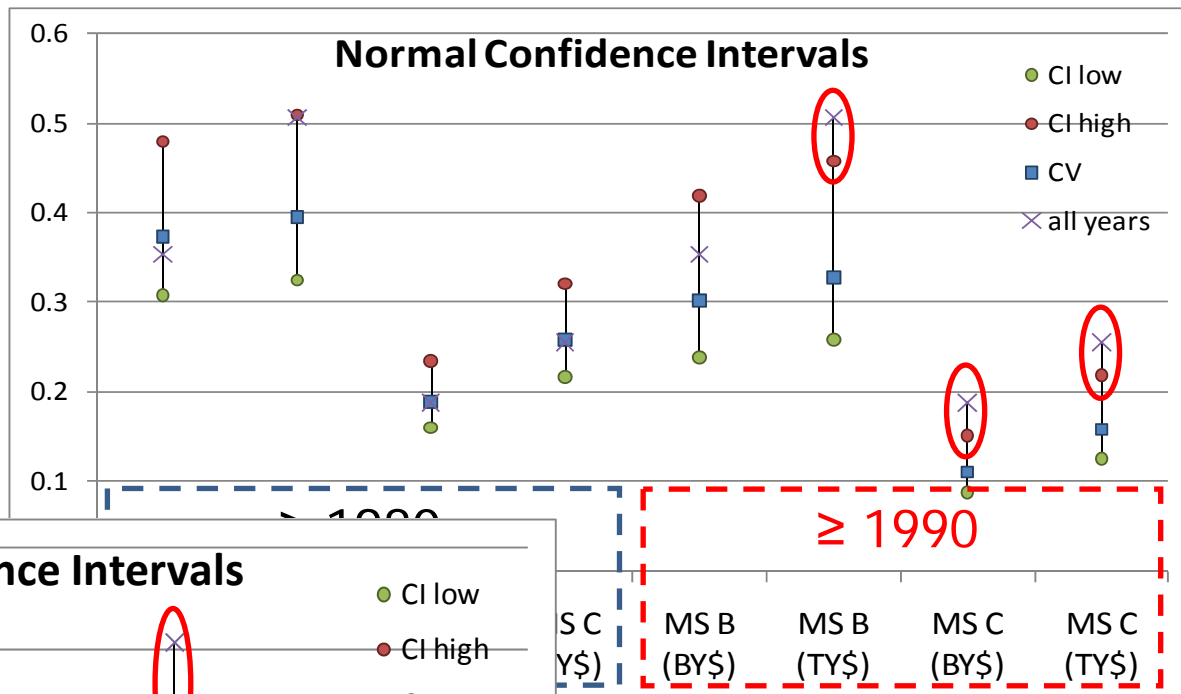
Example CV CIs (TY\$ Fisher)



		approximate		exact				
90%		CI low	CI high	CI low	CI high	mean	std dev	CV
MS B n = 50	normal	0.42	0.61	0.42	0.64	1.36	0.69	0.51
	lognormal	0.43	0.63	0.43	0.63			
MS C n = 43	normal	0.22	0.31	0.22	0.32	1.10	0.28	0.26
	lognormal	0.21	0.31	0.21	0.31			

Application of CIs – Secular Decline

We can conclude that the secular decline is statistically significant at $\alpha = 0.10$ for ≥ 1990 MS B (TY\$ only) and MS C



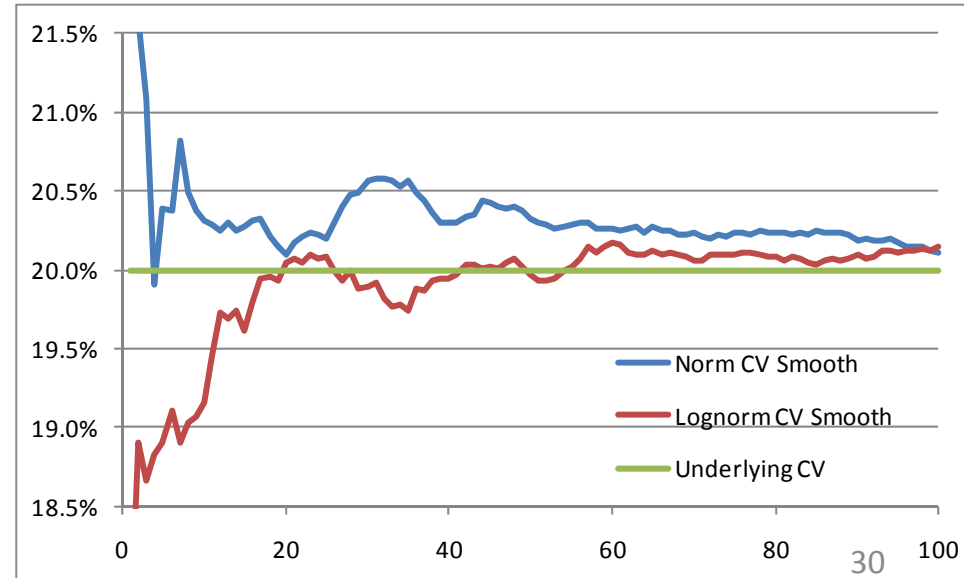
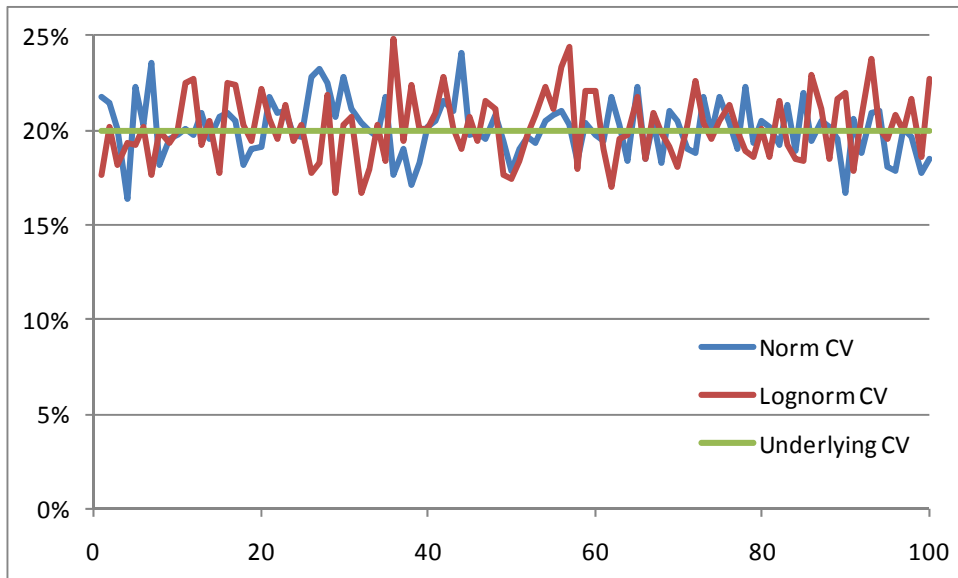
SC (Y\$) MS B (BY\$) MS B (TY\$) MS C (BY\$) MS C (TY\$)

Effect more pronounced in TY\$ due to exclusion of stagflation in the late '70s and early '80s

Lognormal a slightly better fit for SAR data

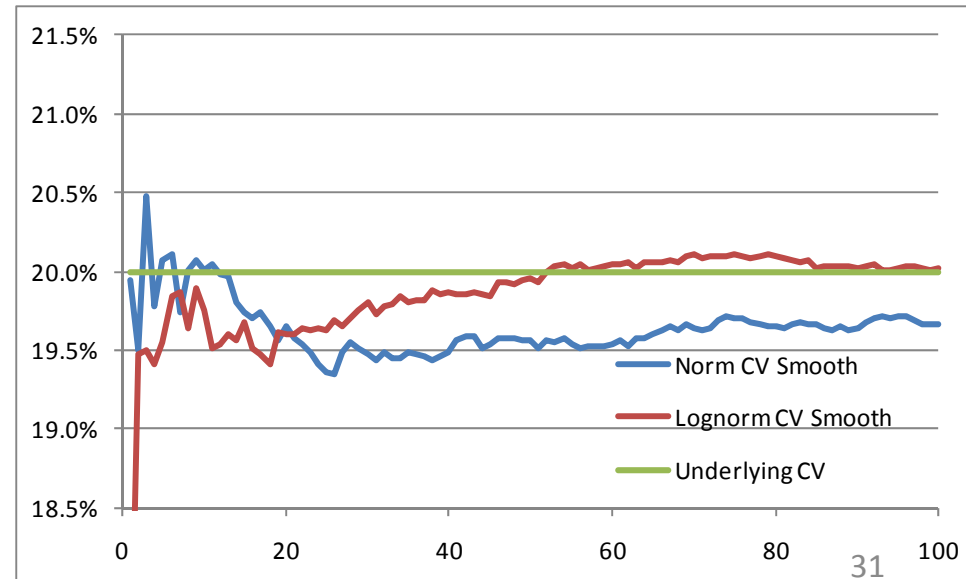
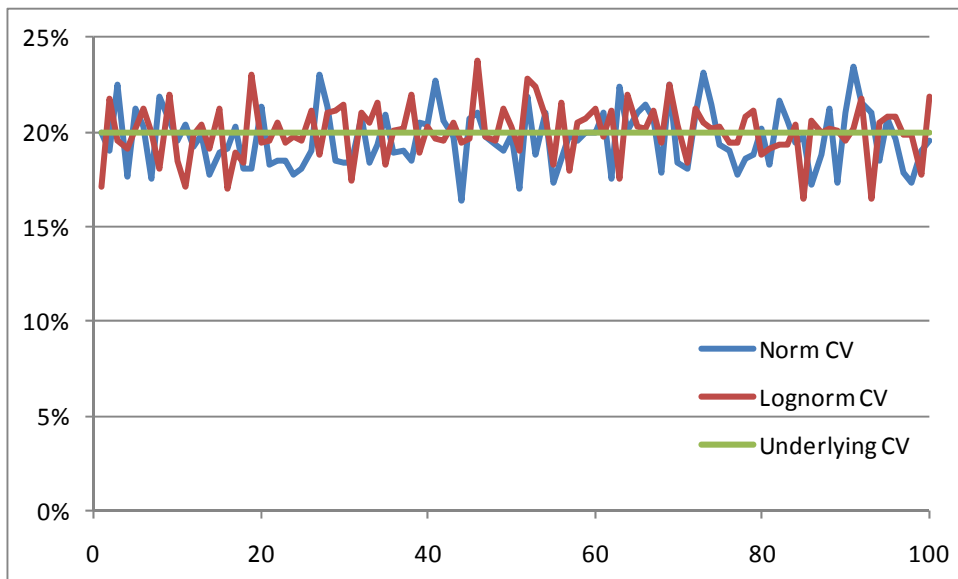
CV of CGFs – iid Programs

- Mental Model #1
 - All historical programs are **independent and identically distributed (iid)**
 - Each program has the same point estimate (EST), \$100M, say; the same mean cost, \$120M, say; and the same standard deviation, \$24M, say
 - We test both normal and lognormal distributions
- In this case, the CV of the observed CGFs is a good predictor of the CV of original cost ($\$24M/\$120M = 20\%$)
 - In fact, the CV of observed CGFs is *identical* to the CV of observed costs for each set of trials (tautological, since simply scaled by a constant factor)
 - Note that we have not yet examined a distribution of CVs, nor constructed a confidence interval around the sample CV



CV of CGFs – Pseudo-iid Programs

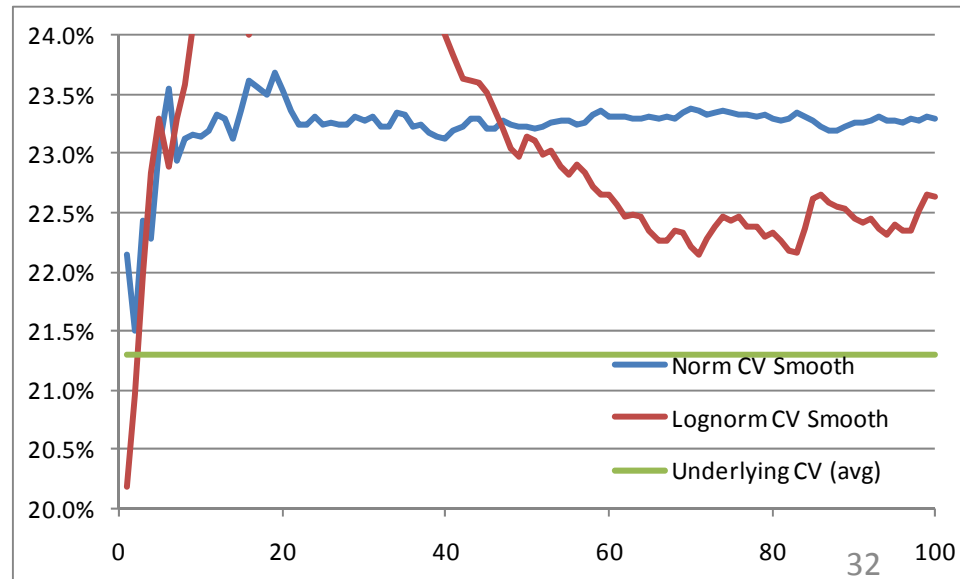
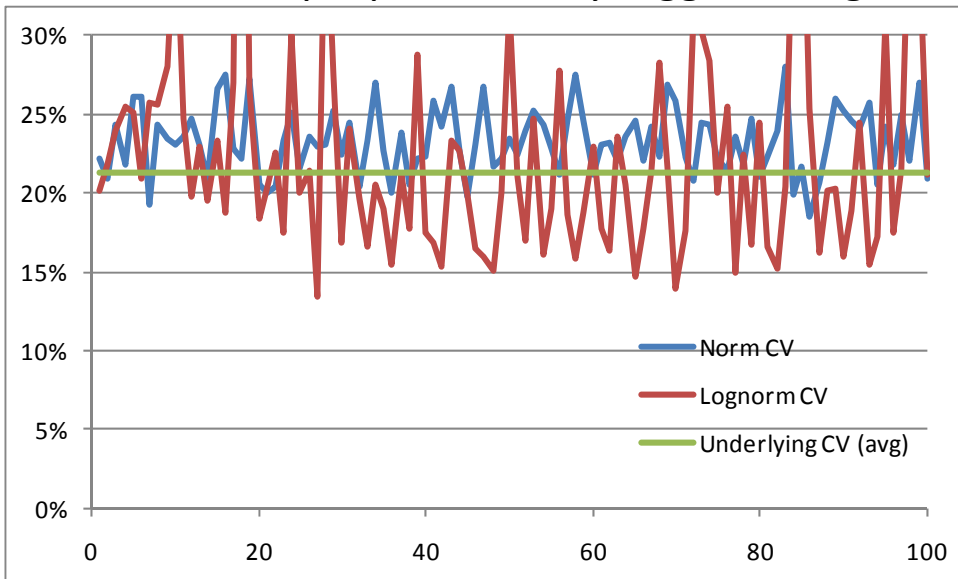
- Mental Model #2
 - Historical programs are of **varying sizes**, say EST is uniform between \$50M and \$150M, but they experience **identical (percentage) cost growth**, say with a mean of 1.2 and a standard deviation of 0.24 (same as previous)
 - We test both normal and lognormal distributions
- In this case, CV of CGFs is (tautologically) identical to CV of cost (for each program) and is a good predictor of true CV
 - CV of cost for the sample is no longer comparable to CV of cost for a program due to the variation in size
 - Note that we have not yet examined a distribution of CVs, nor constructed a confidence interval around the sample CV



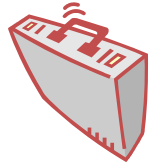
CV of CGFs – Pseudo-iid Programs

- Mental Model #3
 - Historical programs are of **varying sizes**, say EST is uniform between \$50M and \$150M, but they experience **identical (dollar-value) cost growth**, say with a mean of \$20M and a standard deviation of \$24M (same as previous)
 - True underlying CV is no longer constant:
 - $\$24M/\$70M = 34.3\%$ for smallest programs
 - $\$24M/\$170M = 14.1\%$ for largest programs
 - $\$24M/\$120M = 20.0\%$ for average programs (same as previous)
 - We test both normal and lognormal
- In this case, CV of the observed CGFs is *not* a good predictor of program CV
 - Skews high, because same symmetric error for smaller program causes disproportionately bigger swing in CGF

Note asymmetry due to quotient; average = 21.3%



CV of CGFs – The Peril



- As Mental Model #3 illustrates, CV of CGFs can be significantly different from CV of cost and therefore may not be a good proxy [Peril #2A]
- In any case, we should never compare to CV of *observed* costs, which is overstated, since it includes both:
 - Variation in program size (BE\$), and
 - Variation in program performance (overrun/underrun, captured by CGF)
- This is related to a central issue of risk analysis:
 - We are trying to characterize **within-program** risk
 - But “Cost is an unrepeatabe experiment,” and we only ever get one observation for each historical program
 - Thus, we are stuck using data from **cross-program** risk
 - We must cleverly devise a model that explains the former, while using historical data from the latter

CV of Cost – Theoretical Framework

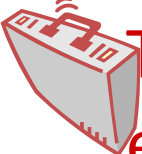
- Neither Mental Model #2 nor #3 is very satisfying, as they fly in the face of historical data
 - Variation in cost cannot be a fixed percent (#2), because the well-established “size effect” says that **larger programs** have a **smaller *percent* variation**
 - Variation in cost cannot be a fixed dollar value (#3), because clearly **larger programs** have a **larger *dollar-value* variation**
- Thus we need a model that will accommodate both these observations, which bring us to... [drum roll]
- Mental Model #4
 - Current Estimate is a linear function of Baseline Estimate with a heteroskedastic error term
 - Variance increases linearly with program size (BE\$)
- We explore this model using the DON SAR data themselves

CV of Cost – SAR Data

- White test conducted
 - Reject null hypothesis of homoskedastic error terms at $\alpha = 0.10$
- Maximum Likelihood Estimation (MLE) regressions
 - Error functional forms tried:

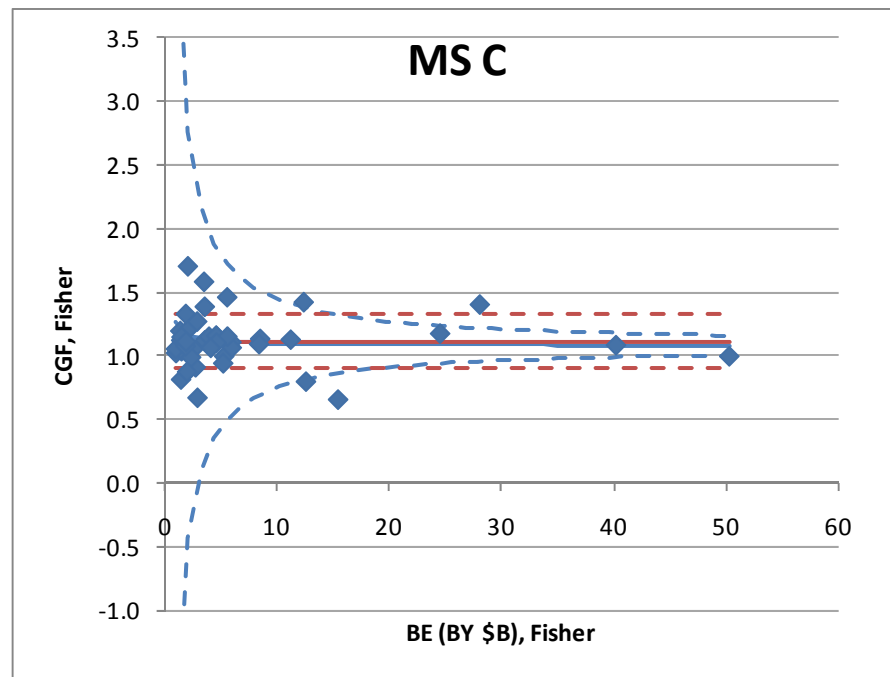
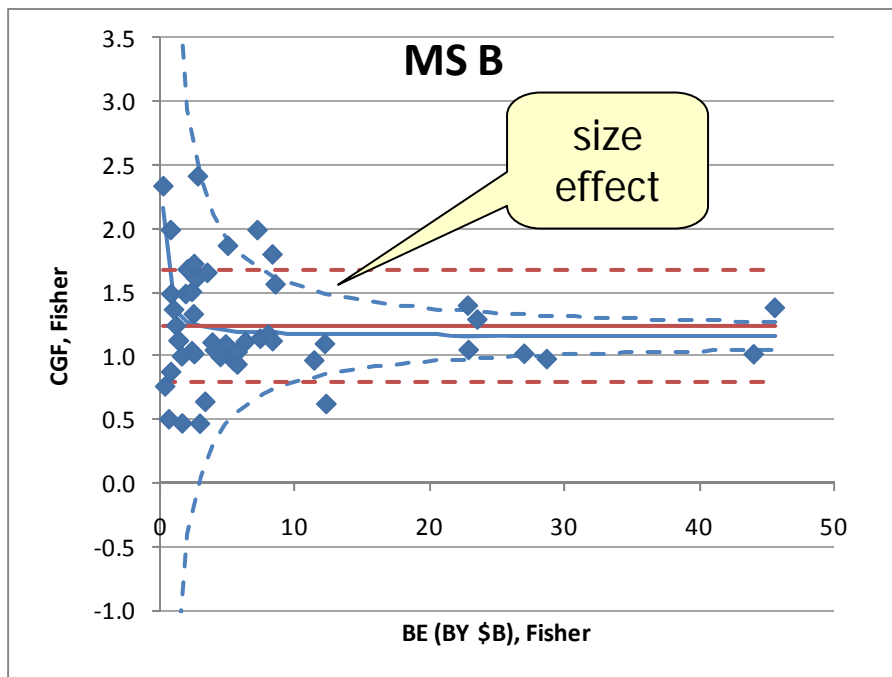
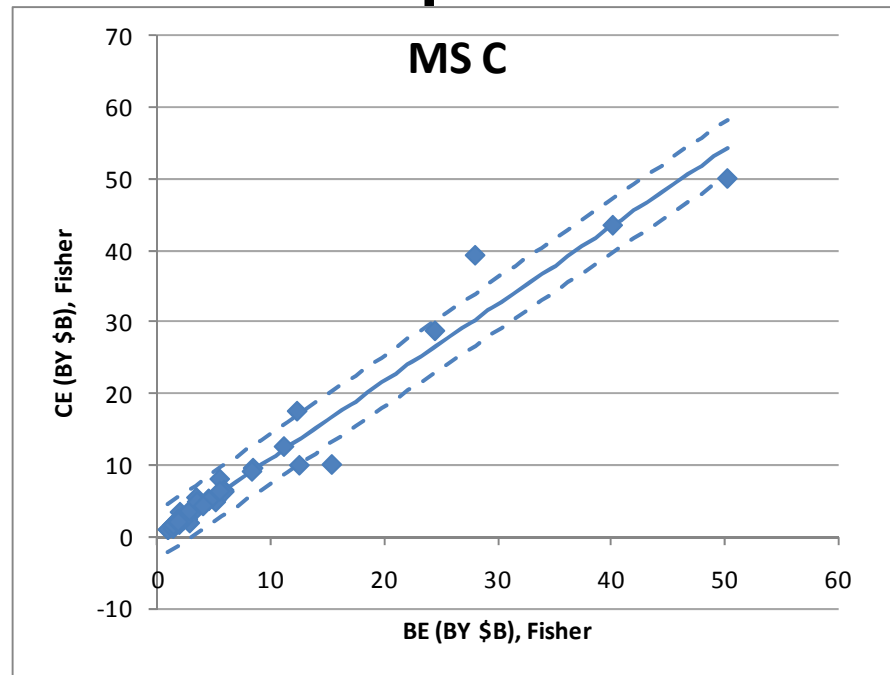
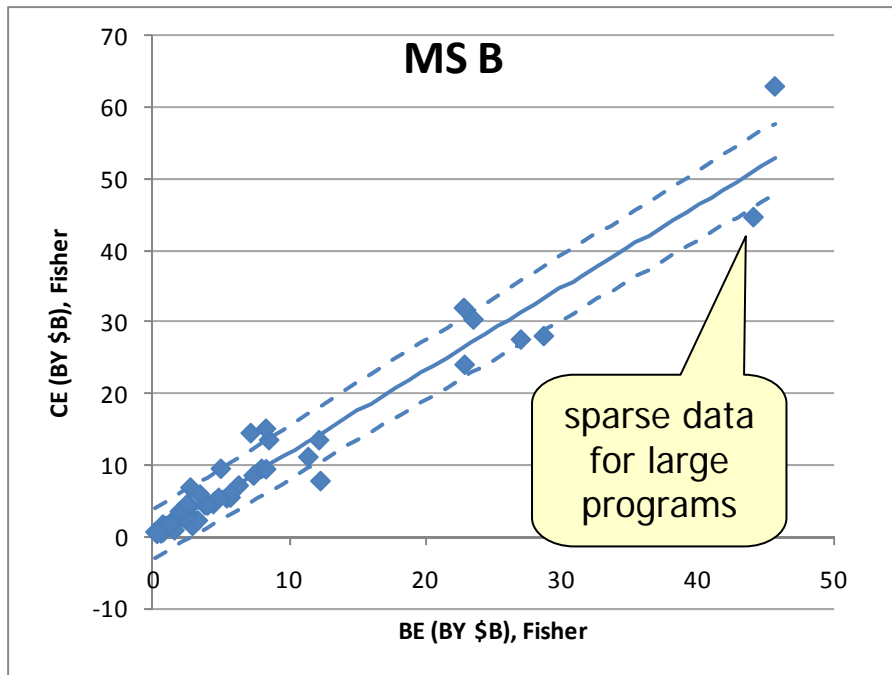


1. $\sigma^2 = kx$, error bands too tight (shrink to zero) for small programs
2. $\sigma^2 = kx^2$, error bands are linear, too wide for large programs, constant CV regardless of program size
3. $\sigma^2 = k_0 + k_1x$, error bands “just right,” models prevalent size effect reasonably, greater \$ errors but smaller % errors for larger programs

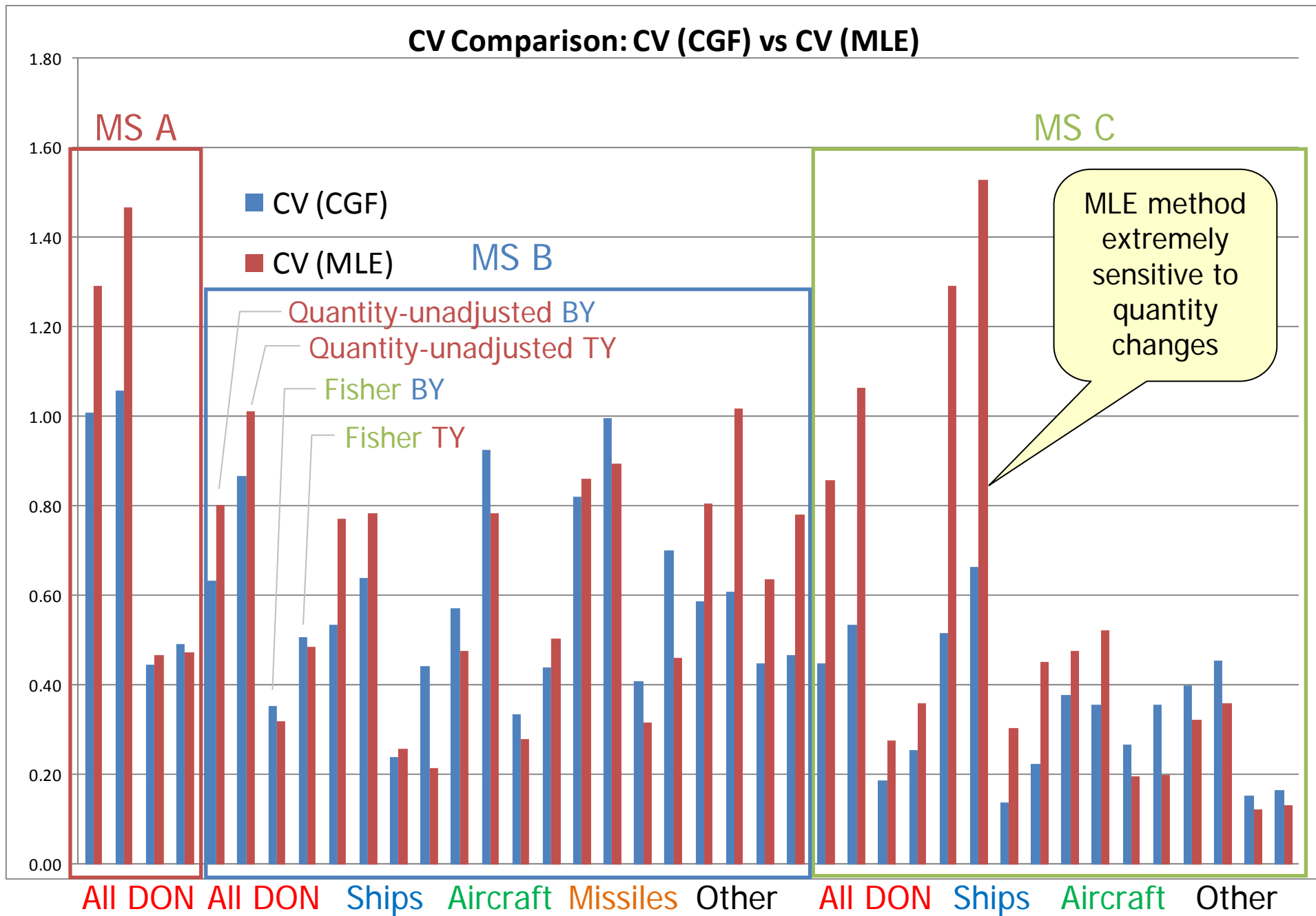
 The previous CV of CGFs method does not allow for a size effect [Peril #2B]

- Because this method uses dollars and not (unitless) quotients, it is somewhat problematic with TY\$
 - Even BY\$ need to be normalized to a common BY!

CV of Cost – Scatterplots



CV Comparison – CGF vs. MLE



CV Comparison – CGF vs. MLE

- A paired t -test shows a statistically-significant difference between the two sets of factors
 - This is a bit artistic, since we don't believe either comes from a consistent population
- *So which is better?*
- As of this writing, NCCA S-Curve Tool still uses CV of CGFs
 - Thought to be a reasonable proxy
- MLE Regression may prove to be an improvement
 - Most importantly, models *within-program* risk using *cross-program* historical data
 - Also accounts for widely-acknowledged size effect
 - Presently lacks measures of statistical significance and uncertainty
 - May explore bootstrapping
- Secular decline notwithstanding, both methods effectively convey the key message that *our CVs are too low!*
 - On the previous chart, CVs are seldom below 30%, and hardly ever below 20%, but when's the last time you saw CVs that high?!

NCCA S-Curve Tool Algorithms

- Specification of distributions
 - Normal and Lognormal, any two of:
 - Mean, median, or mode; CV; and any percentile
 - CV and a percentile case independently derived for eSBM
 - Lognormal two solutions
 - Mean and percentile, or mode and percentile
- Historical adjustments (CGF and CV)
 - CV historical adjustment: “Pivoting” at the mean vs. median
 - CGF-only historical adjustment: not just a translation
 - Historical adjustment for empirical data
 - Normal or lognormal distributions (parametric) vs. empirical distribution of CGFs (non-parametric)
- Graphing of probability distributions
 - Graphical sampling
 - Developed to solve noisy PDF problem
 - Reduces size, increases speed of model
 - Order of operations: historical adjustments and summary statistics

Alternate Specification of Normal

Given	Mean	Std Dev	CV
Mean, CV	μ	$CV \cdot \mu$	CV
Mean, Percentile (X_p, p)	μ	$\frac{X_p - \mu}{Z_p}$	$\frac{\left(\frac{X_p}{\mu} - 1 \right)}{Z_p}$
CV, Percentile (X_p, p)	$\frac{X_p}{1 + Z_p \cdot CV}$	$\frac{X_p \cdot CV}{1 + Z_p \cdot CV}$	CV

$$Z_p = \Phi^{-1}(p)$$

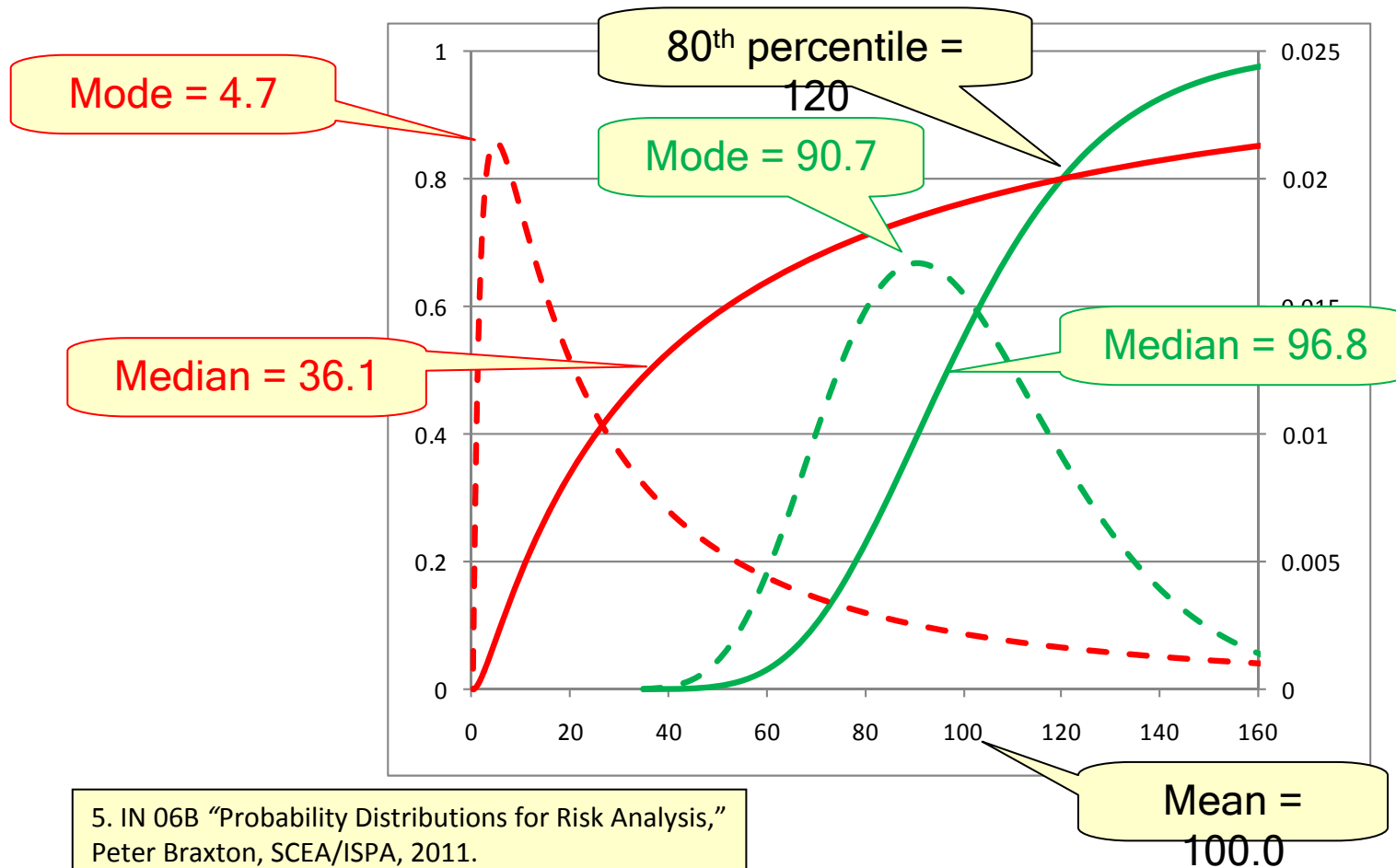
Alternate Specification of Lognormal

Given	Mean	Std Dev	CV	μ	σ
Mean, CV	$E(X)$	$CV \cdot E(X)$	CV	$\ln\left(\frac{E(X)}{\sqrt{1+CV^2}}\right)$	$\sqrt{\ln(1+CV^2)}$
Mean, Percentile (X_p, p)	$E(X)$	$CV \cdot E(X)$	$\sqrt{e^{\sigma^2} - 1}$	$\ln\left(\frac{E(X)}{\sqrt{1+CV^2}}\right)$	$Z_p \pm \sqrt{Z_p^2 - 2 \ln\left(\frac{X_p}{E(X)}\right)}$
CV, Percentile (X_p, p)	$e^{\mu + \frac{\sigma^2}{2}}$	$CV \cdot E(X)$	CV	$Z_p \sqrt{\ln(1+CV^2)}$	$\sqrt{\ln(1+CV^2)}$

$Z_p = \Phi^{-1}(p)$

Two Lognormals Example

- Mean = 100, 80th percentile = 120
 - “Regular” solution with 26% CV
 - “Extreme” solution with 258% CV!



5. IN 06B “Probability Distributions for Risk Analysis,” Peter Braxton, SCEA/ISPA, 2011.

Historical Adjustments

- Three ways to apply historical adjustments:
 - CV Only: We believe your Risk, but your Uncertainty is off
 - CGF Only: We believe your Uncertainty, but your Risk is off
 - CGF and CV: Your Risk and Uncertainty are both off
- Two modes in which historical adjustments are applied

– Empirical

$$x_i^* = CGF \left[\bar{x} + \frac{CV_{CGF}}{CV_x} (x_i - \bar{x}) \right]$$

$$x_i^* = \bar{x} + \frac{CV_{CGF}}{CV_x} (x_i - \bar{x})$$

$$x_i^* = CGF \cdot x_i$$

– Parametric

$$\mu^* = CGF \cdot \mu \quad \sigma^* = CV^* \cdot \mu^*$$

$$\mu^* = \mu \quad \sigma^* = CV^* \cdot \mu$$

$$\mu^* = CGF \cdot \mu \quad \sigma^* = CV \cdot \mu^*$$

- In the CV Only and CGF Only modes, historical adjustments inevitably cause shifts in other parameters [Peril #3A]

Beliefs Drive Results

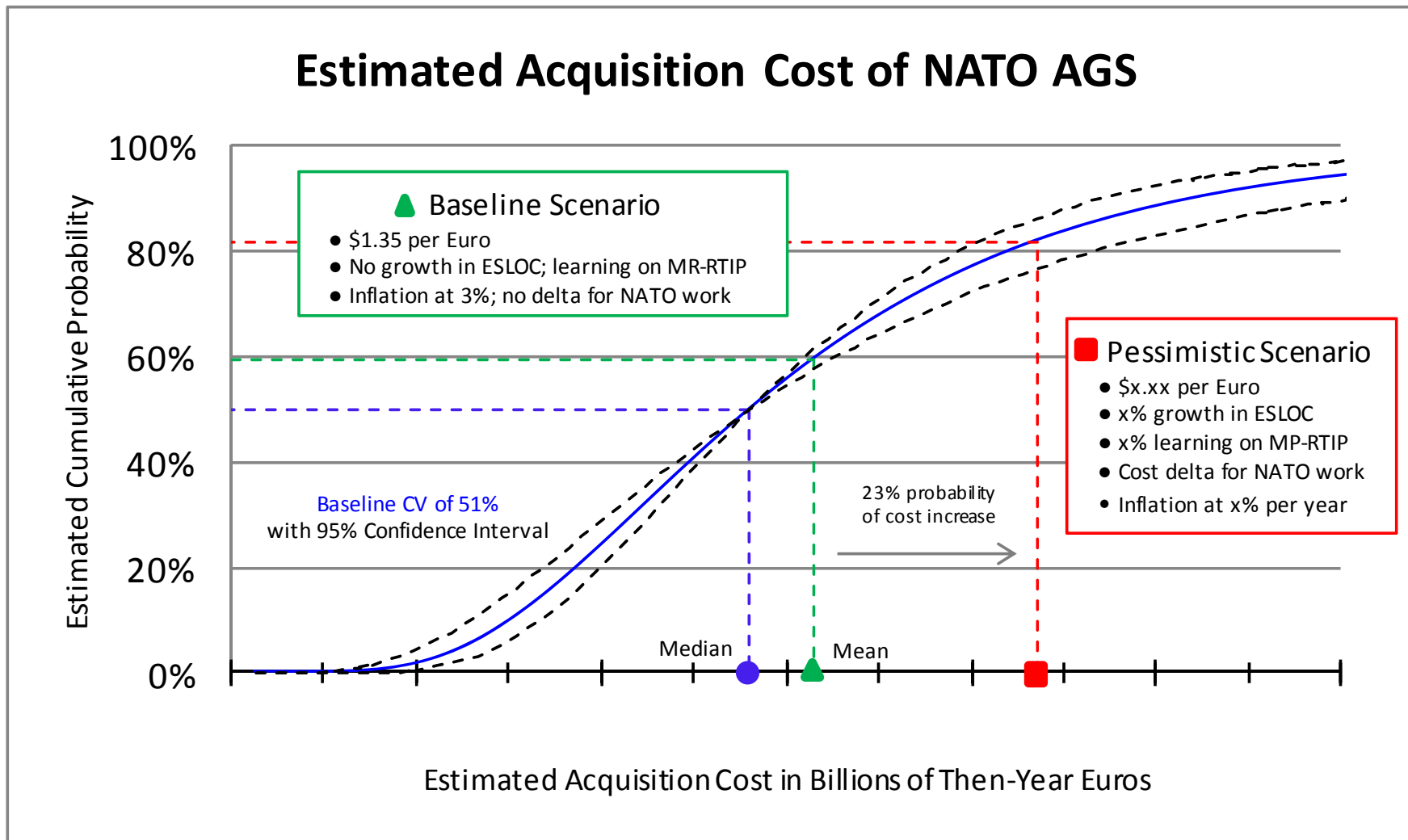
- Choices can trigger the law of unintended consequences
- The below tables show what our choices imply
- Results assume CV increases (top) and $CGF > 1.0$ (bottom)
 - If CV drops, or $CGF < 1.0$, reverse the cited effects
- For the Lognormal, the median is the natural “pivot point”
 - We still favor Case 1 over Case 2, since the Mean is what we care about

CV Only mode Pivot Point Choices	Maintain \$ of Mean (Case 1)	Maintain \$ of Median (Case 2)
Normal	No Effect	No Effect
Lognormal	Median \$ Drops Mean % Rises	Median Constant Mean \$ Rises, Mean % Rises

CGF Only mode	Believe CV	Believe s.d.
Normal/Lognormal	SD Rises	CV Shrinks

NATO AGS Example with CV

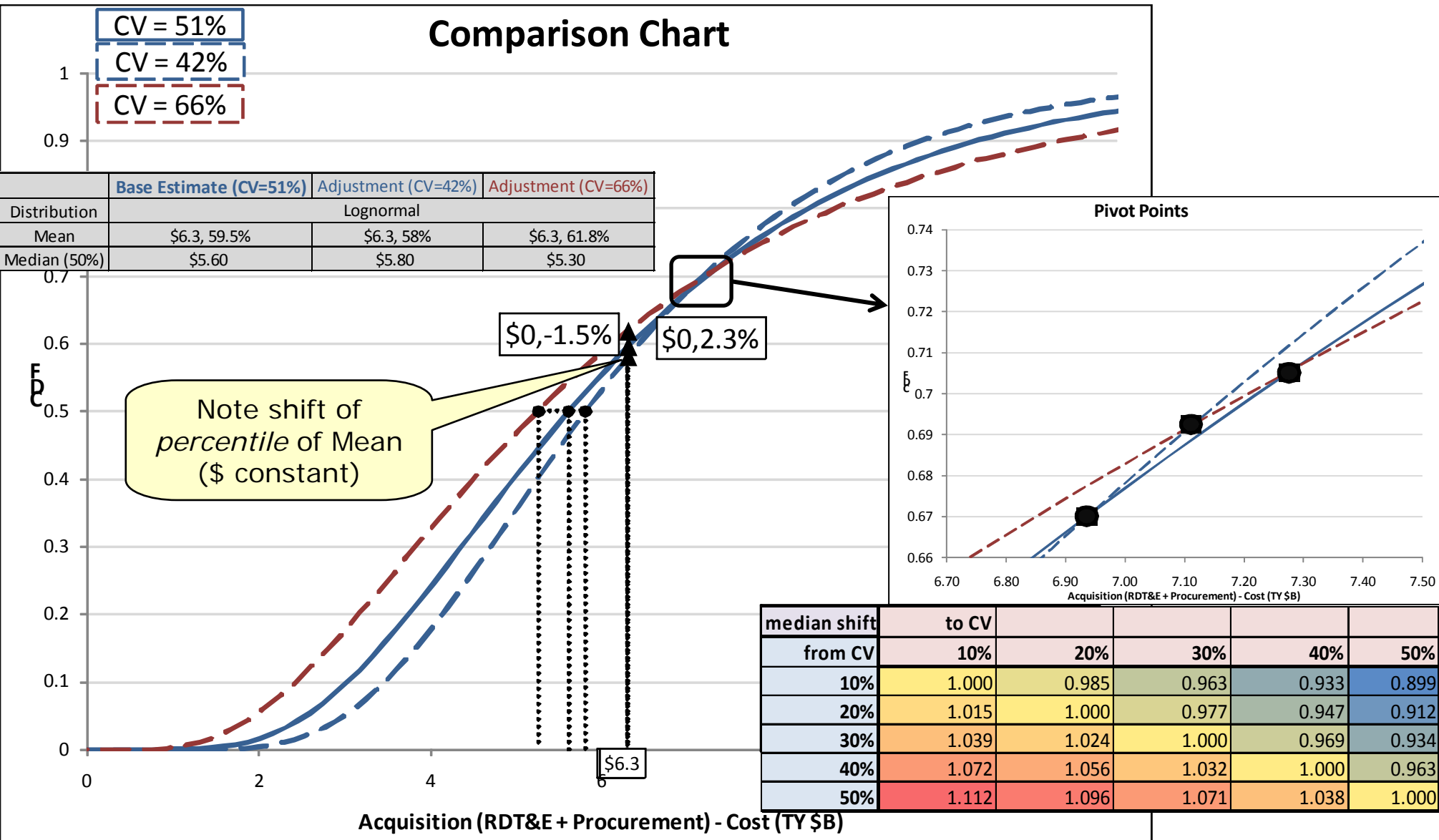
Confidence Bands



Cost values not displayed because of business sensitivity

Case 1: Maintain \$ of Mean

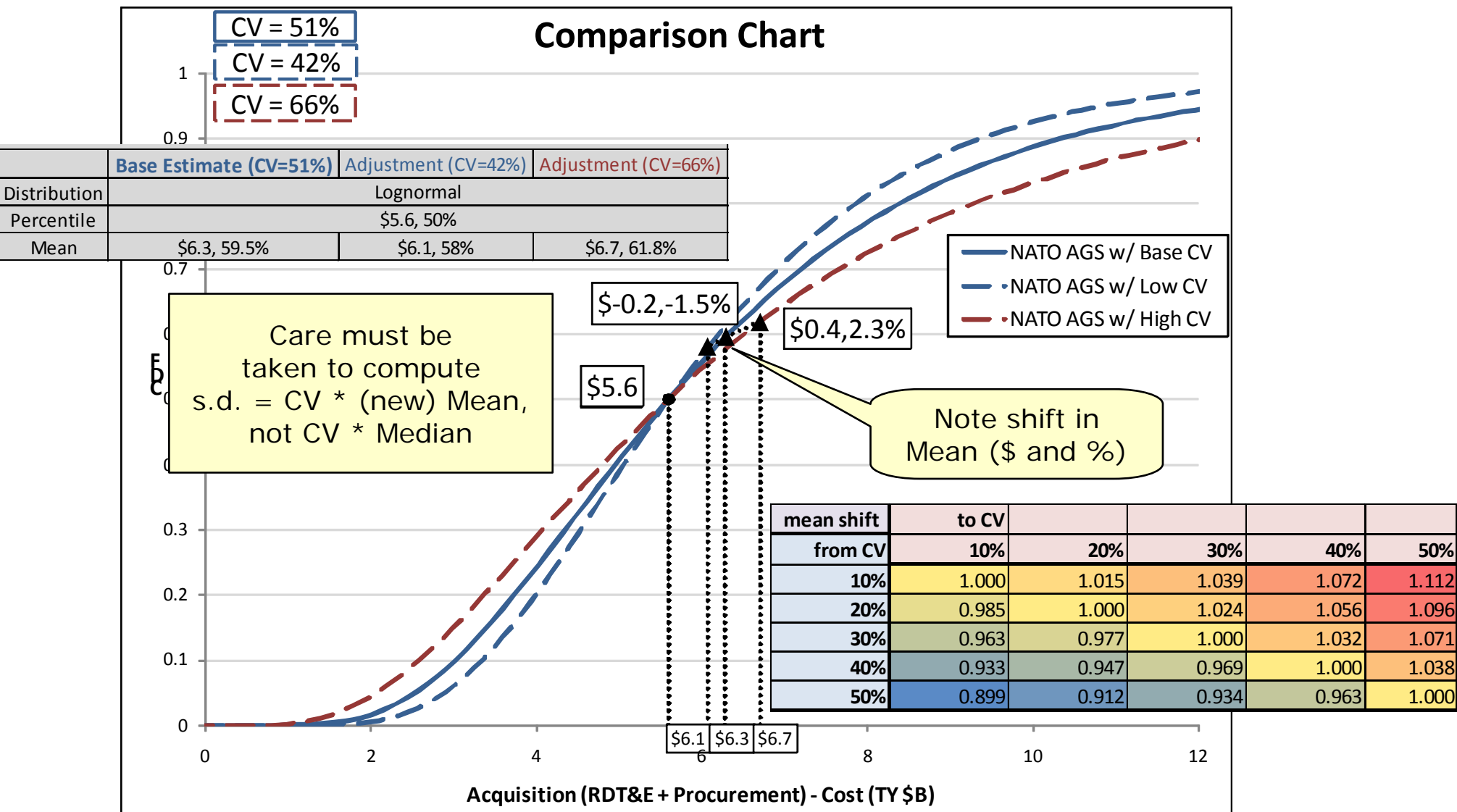
Created chart by applying Mean = \$6.3 and CV = 51%, 42%, & 66%



Cost values notional – correspond to tick marks on previous graph

Case 2: Maintain \$ of Median

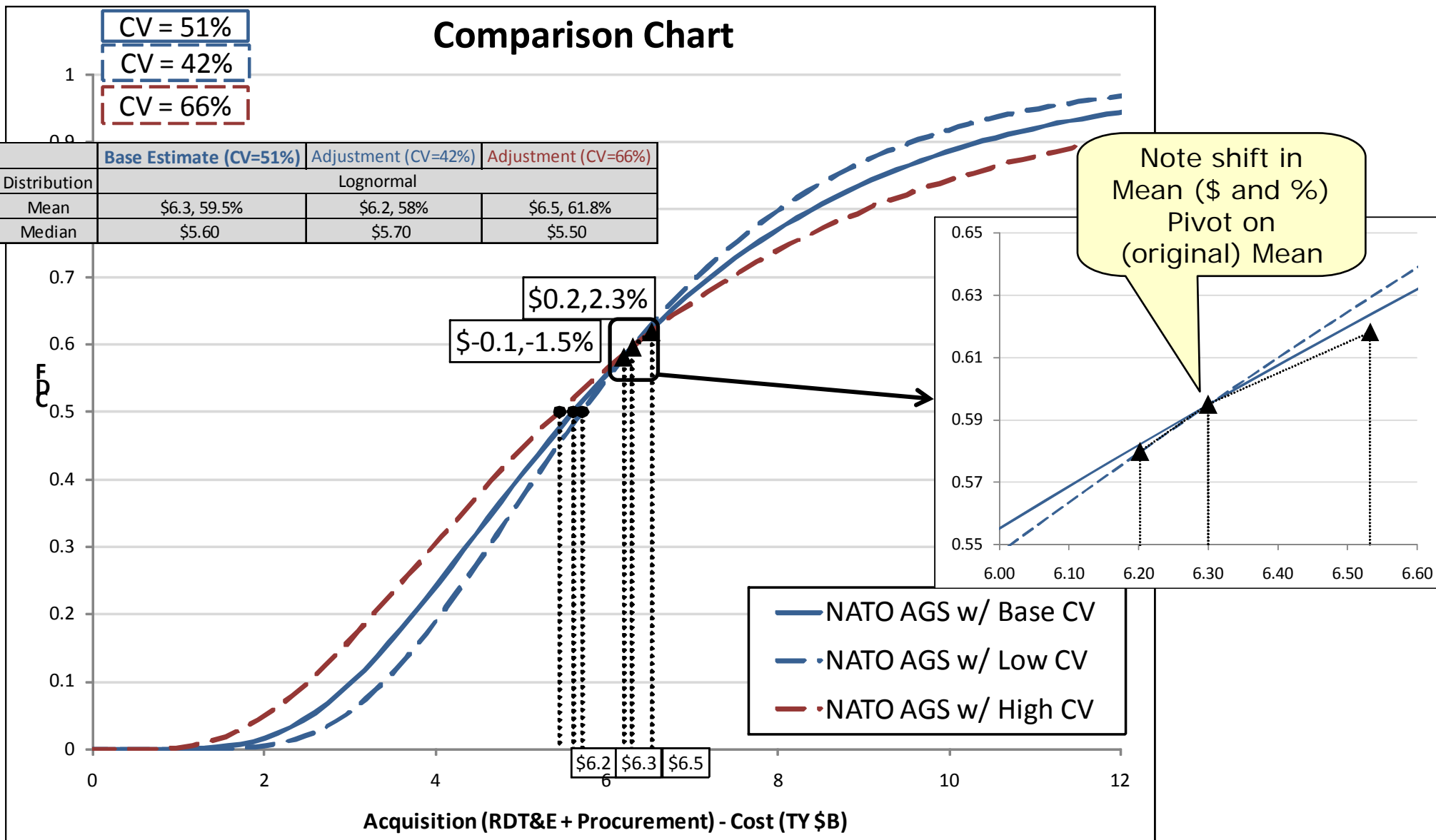
Created chart by **applying CV=51%,42%, & 66% and percentile=(\$5.6,50%)**
 Same chart as shown on previous slide



Cost values notional – correspond to tick marks on previous graph

Case 3: CDF Passes Through Coordinates of Base Estimate Mean

Created chart by **applying CV=51%,42%, & 66% and percentile=(\$6.3,59.5%)**



Cost values notional – correspond to tick marks on previous graph

S-Curve vs. Stair Step

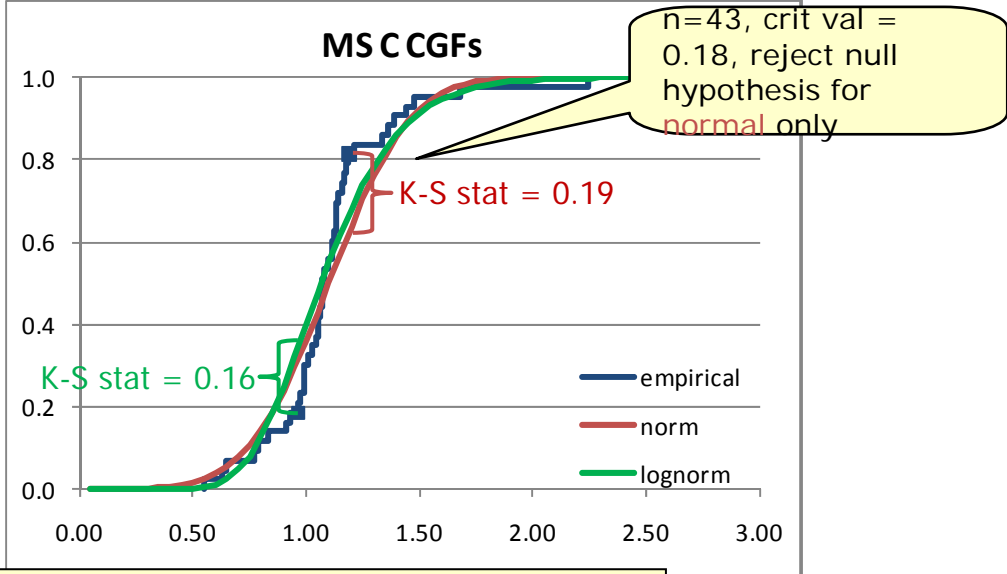
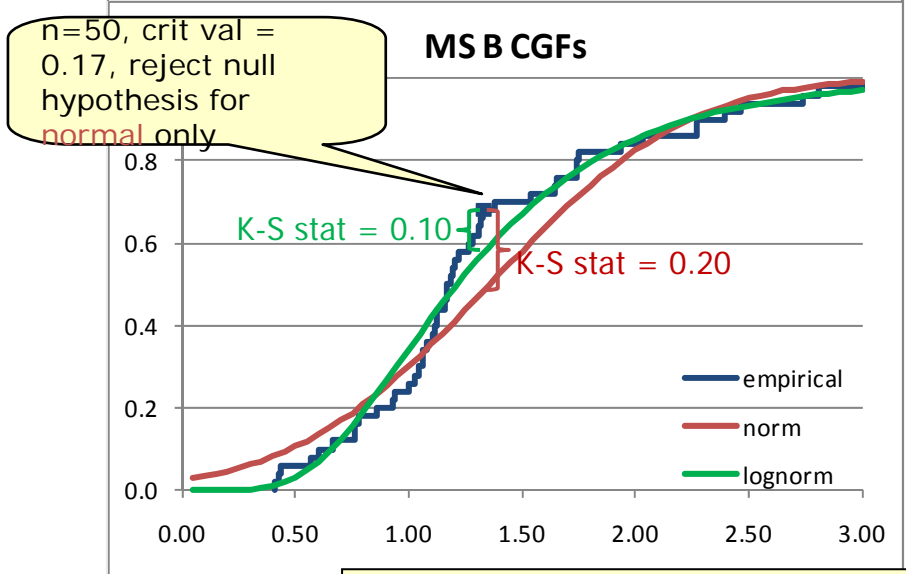


The aforementioned methods preserve the “shape” of the baseline distribution and ignore the “shape” of historical (cross-program) risk [Peril #3B]

- Parametric approach
 - Normal or lognormal distribution with historical CV
 - Distributions diverge as CV increases
- Non-parametric approach
 - Empirical distribution of CGFs (non-parametric)
 - Does not circumvent Peril #2, that distribution of CGFs may not be the right thing to look at in the first place

For $\alpha=0.10$, $n>40$,
K-S critical value is approx.⁸ $\frac{1.22}{\sqrt{n+\sqrt{n/10}}}$

• Comparison graphs for MS B (CV = 51%) and MS C (CV = 26%), **TY\$ Fisher**



8. *Practical Nonparametric Statistics* (3rd ed.), W.J. Conover, John Wiley and Sons, Inc., 1999. Table A13 Quantiles of the Kolmogorov Test Statistic.

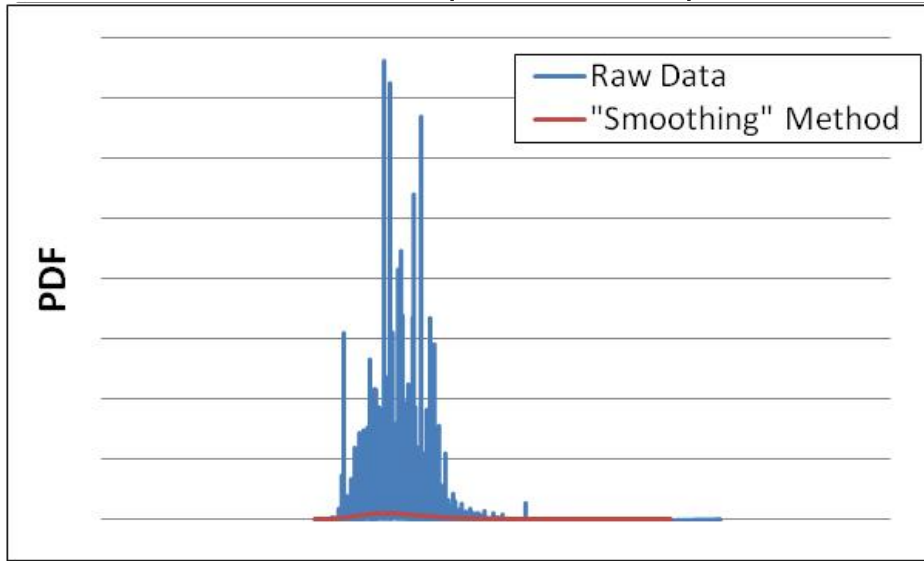
Graphical Sampling

- Graphical sampling approach
 - Define min and max x -values for graph
 - Empirical case: based on actual extreme values
 - Parametric case: $-3\sigma / +3\sigma$ for Normal; $-2\sigma / +4\sigma$ (unit space) for Lognormal
 - Subdivide into number of equal intervals (say, 200)
 - Use each “fencepost” to look up a single adjacent x -value from the empirical data set
 - If many trials and x is dense, several x -values will be skipped, avoiding PDF spikes for x -values very close together (constant probability divided by very small Δx)
 - If few trials and x is sparse, x -values may be repeated several times, but no matter
- Reduces size, increases speed of model
 - Only need an order of magnitude fewer points for graphing
 - Summary statistics performed on original data set (trials)
 - Order of operations: historical adjustments can be applied to sampled values to again save memory and computation

Graphical Sampling Illustrated

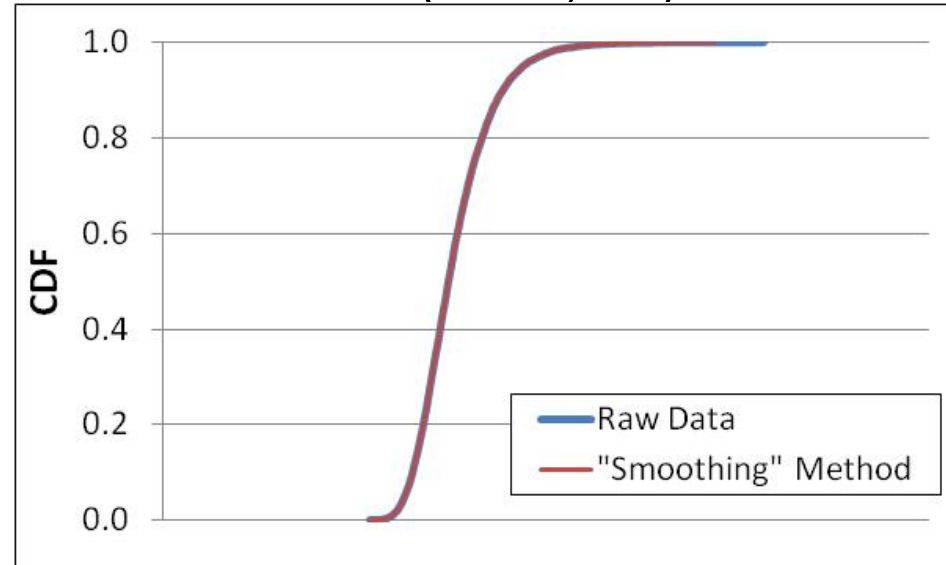
PDF

- Dense x-data ($n = 10,000$)

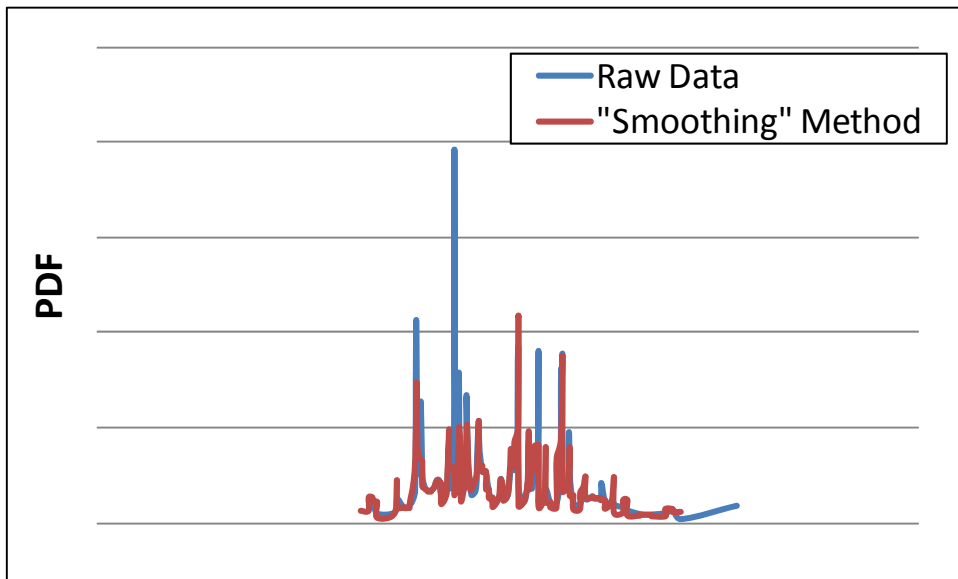


CDF

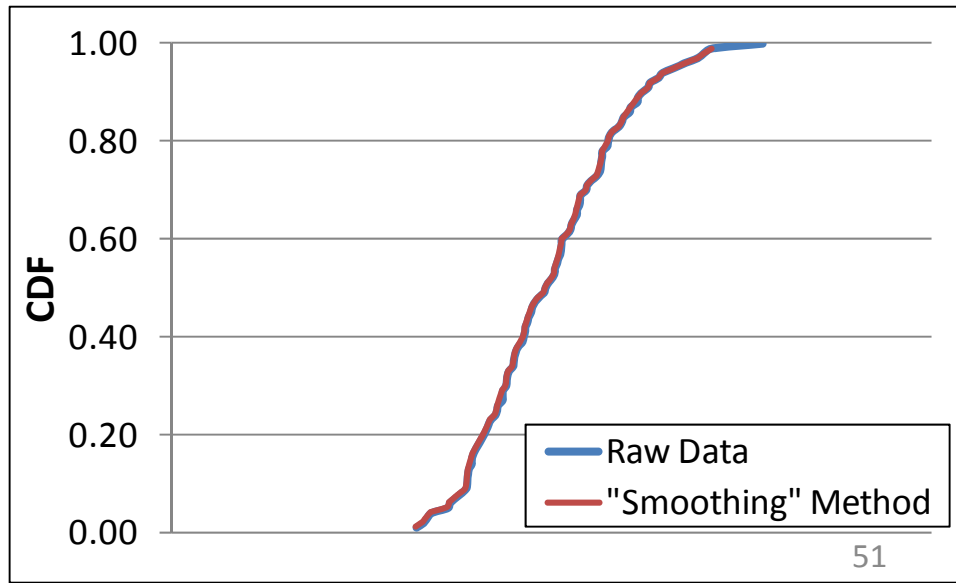
- Dense x-data ($n = 10,000$)



- Sparse x-data ($n = 100$)



- Sparse x-data ($n = 100$)



Open Questions and Future Work

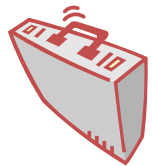
- Additional SAR data
 - Development / Procurement breakouts for DON programs
 - Army and Air Force program data
- SAR Database update plan with USD(AT&L) ARA and OSD CAPE
- Integration of SAR Database and S-Curve Tool
 - Provide greater flexibility (“live” queries) for Historical Adjustments
- Implementation of era (secular decline), size effects, confidence intervals in S-Curve Tool
- Inflation^{9,10} – not nearly as straightforward as we thought!
 - Contributions of inflation *and* time-phasing to difference in TY\$ and BY\$ factors
- DON Guidance for Cost and Acquisition Communities
 - Probability of Program Success (PoPS) update, e.g.
- Toward a common Cost Growth taxonomy (SAR, PARCA, et al.)
 - Decomposition of Risk and Uncertainty into components

Might be one reason analysts traditionally jump straight to CGFs!

9. “The Use of Inflation Indexes in the Department of Defense,” Stanley Horowitz, DoDCAS, 2011.

10. “Then Year Estimates: Is There A Better Way?” John McCrillis, DoDCAS, 2011.

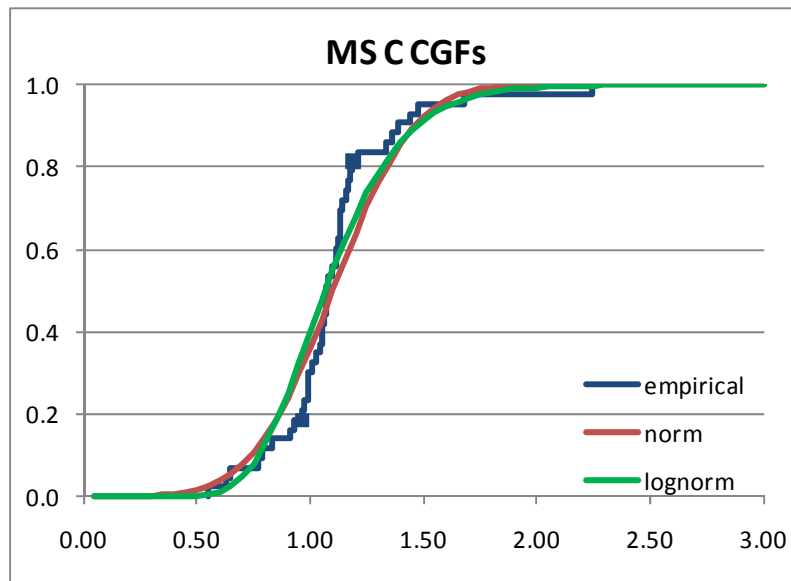
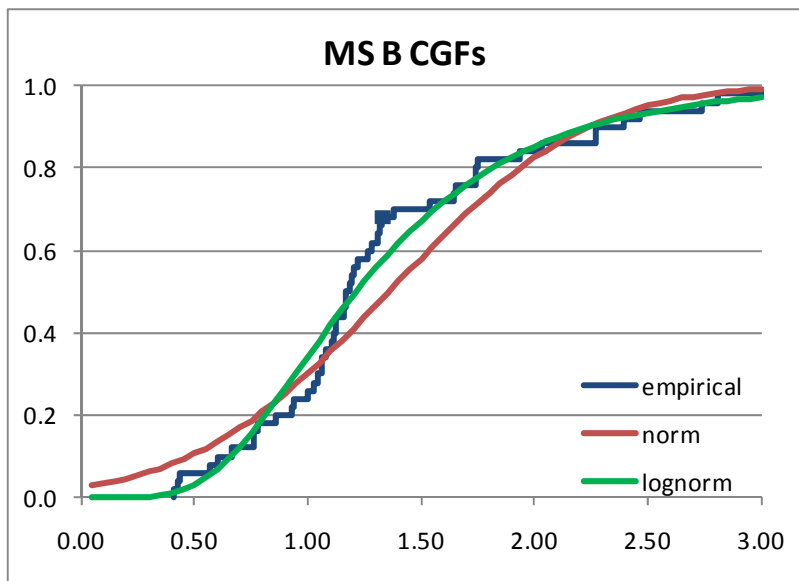
Perils of Portability Summary



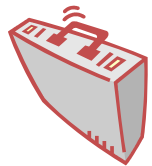
Perils of fit (#1A/1B and #3B)

- Differences in empirical and fit distributions of CGFs
- Mean/median ratio affects portability of CV
- Fit distributions overestimate proportion of programs that underrun

		Peril #1A -- Mean/Median ratio			Peril #1B -- Percentile of 1.0		
		empirical	normal	lognormal	empirical	normal	lognormal
MS B	BY	1.11	1.00	1.06	26.0%	29.7%	38.7%
	TY	1.16	1.00	1.12	24.0%	30.1%	39.5%
MS C	BY	1.02	1.00	1.02	23.3%	29.4%	37.5%
	TY	1.01	1.00	1.03	30.2%	36.1%	43.3%



Perils of Portability Summary

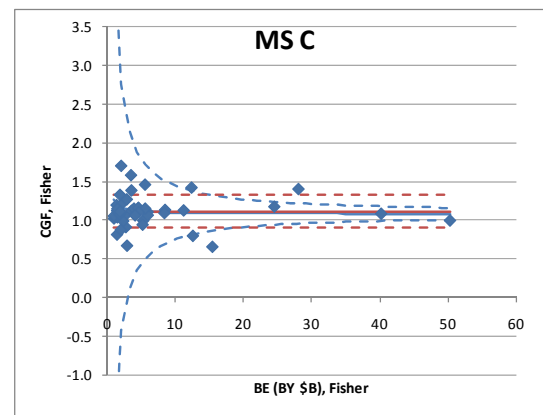
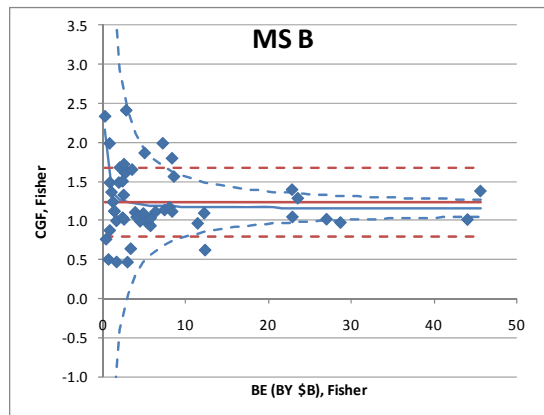
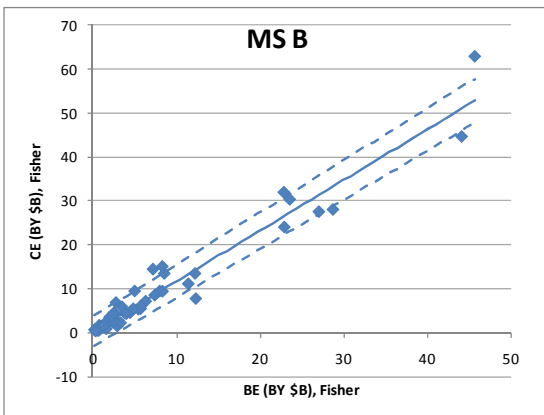


Perils of application (#2A/2B and #3A)

- Adjusting CV causes a median shift in the lognormal when “pivoting” at the mean (and vice versa)
- Applying CGF causes an increase in standard deviation (even for fixed CV)

			CGFs			MLE		
			small (MIDS-LVT)	medium (EA-18G)	large (DDG 51)	small (MIDS-LVT)	medium (EA-18G)	large (DDG 51)
MS B	BY	CGF	1.23	1.23	1.23	1.35	1.18	1.16
		CV	35.4%	35.4%	35.4%	290.8%	44.8%	18.3%
	TY	CGF	1.36	1.36	1.36	1.80	1.21	1.13
		CV	50.6%	50.6%	50.6%	255.8%	56.6%	26.5%
MS C	BY	CGF	1.11	1.11	1.11	1.17	1.10	1.09
		CV	18.8%	18.8%	18.8%	182.4%	41.5%	13.1%
	TY	CGF	1.10	1.10	1.10	1.05	1.11	1.12
		CV	25.5%	25.5%	25.5%	170.6%	43.6%	14.0%

Need to examine extremes of size effect



Final Thoughts on Portability

- Portability is a noble goal and not in and of itself evil
 - CER Libraries, Secondary Data, SCEA papers (!), and other information sharing are the embodiment of portability
- Portability can be a slippery slope or a crutch, so we exhort you to:
 - Use the “correct” functional form, which almost always includes a non-zero y-intercept
 - Use the “correct” probability distribution to model uncertainty
 - Retain and use measures of statistical significance and goodness of fit
 - Retain and scrutinize basis for all estimating relationships
 - Integrate data collection, data normalization, data analysis, and estimating, wherever practicable

Bibliography

1. “CERs – Rates, Factors, and Ratios” in Cost Estimating Body of Knowledge (CEBoK), Module 3 Parametric Estimating, SCEA, 2010
2. “ ‘To b or Not to b’: The y-intercept in Cost Estimation,” R.L. Coleman, J.R. Summerville, P.J. Braxton, B.L. Cullis, E.R. Druker, SCEA/ISPA, 2007
3. “Development and Application of CV Benchmarks,” Brian Flynn, Paul Garvey, Peter Braxton, Richard Lee, DoDCAS, 2011
4. Sources of Weapon System Cost Growth (MG-670), Joseph G. Bolten, Robert S. Leonard, Mark V. Arena, Obaid Younossi, Jerry M. Sollinger, RAND Project Air Force, 2008
5. IN 06B “Probability Distributions for Risk Analysis,” Peter Braxton, SCEA/ISPA, 2011
6. <http://en.wikipedia.org/wiki/Percentile>
7. “Confidence Bounds for Normal and Lognormal Distribution Coefficients of Variation,” Steve Verrill, Forest Products Laboratory, Research Paper FPL-RP-609, 2003
8. *Practical Nonparametric Statistics* (3rd ed.), W.J. Conover, John Wiley and Sons, Inc., 1999.
9. “The Use of Inflation Indexes in the Department of Defense,” Stanley Horowitz, DoDCAS, 2011
10. “Then Year Estimates: Is There A Better Way?” John McCrillis, DoDCAS, 2011

Acronyms

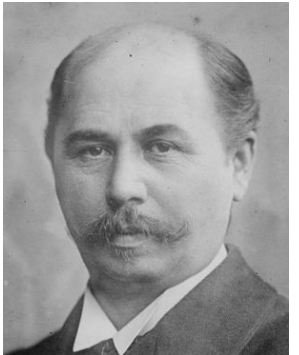
ACAT = Acquisition Category	LPD = Landing, Platform, Dock [ship]
AF = Air Force	MDAP = Major Defense Acquisition Program
AGS = Alliance Ground Surveillance	MIDS-LVT = Multifunctional Information Distribution System - Low Volume Terminal
AIM = Air Intercept Missile	MLE = Maximum Likelihood Estimation
ARA = Acquisition Resources and Analysis	MS = [Acquisition] Milestone
BE = Baseline Estimate	n = number of observations
BY\$ = Base Year dollars	NATO = North Atlantic Treaty Organization
CAIG = Cost Analysis Improvement Group	NIST = National Institute of Standards and Technology
CAPE = Cost Assessment and Program Evaluation	OLS = Ordinary Least Square [Regression]
cdf = cumulative distribution function	OSD = Office of the Secretary of Defense
CE = Current Estimate	PARCA = Performance Assessments and Root Cause Analyses
CEBoK = Cost Estimating Body of Knowledge	PdE = Production Estimate
CER = Cost Estimating Relationship	pdf = probability density function
CG = Guided Missile Cruiser	PE = Planning Estimate
CGF = Cost Growth Factor	POE = Program Office Estimate
CI = Confidence Interval	PoPS = Probability of Program Success
COTS = Commercial Off-The-Shelf	Q Δ = Quantity delta
CV = Coefficient of Variation	R&D = Research and Development
CVN = Nuclear-powered Aircraft Carrier	RDT&E = Research, Development, Test, and Evaluation
DDG = Guided Missile Destroyer	SAR = Selected Acquisition Report
DE = Development Estimate	SCEA = Society of Cost Estimating and Analysis
DoDCAS = Department of Defense Cost Analysis Symposium	s.d. = standard deviation
DON = Department of the Navy	SLOC = Source Lines Of Code
EA = Electronic Attack [aircraft]	TY\$ = Then Year dollars
eSBM = enhanced Scenario-Based Method	USD(AT&L) = Under Secretary of Defense for Acquisition, Technology, and Logistics
F/A = Fighter/Attack [aircraft]	w.l.o.g. = without loss of generality
ICE = Independent Cost Estimate	wrt = with respect to
i.i.d. = independent and identically distributed	
K-S = Kolmogorov-Smirnov [test]	

Laspeyres, Paasche, Fisher – Oh, My!



Ernst Louis Étienne Laspeyres

- (Ernst Louis) Étienne Laspeyres (1834-1913)
 - German economist of Huguenot Gascon (SW France, near Spain) descent
 - 1871 development of eponymous price index
 - Comparison using base period quantities, tends to overstate inflation
 - Wrote such page-turners as *Geschichte der Volkswirtschaftlichen Anschauungen der Niederländer und ihrer Literatur zur Zeit der Republik*



- Hermann Paasche (1851-1925)
 - German statistician and economist, died in Detroit
 - Eponymous price index
 - Comparison using current period quantities, tends to understate inflation
 - Wrote such page-turners as *Wandlungen in der modernen Volkswirtschaft*



- Irving Fisher (1867-1947)
 - American neoclassical economist
 - Eponymous price index
 - Geometric mean of Laspeyres and Paasche indices
 - Wrote the 1922 classic *The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability*

Implied Percentile – Normal

- Without loss of generality (w.l.o.g.), assume mean of one (1.0)
 - Then the standard deviation is equal to the CV!
- We are looking for the percentile of the reciprocal of CGF

$$X \sim N(\mu, \sigma) \Leftrightarrow Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$p = \Phi\left(\frac{(1/CGF) - \mu}{\sigma}\right) = \Phi\left(\frac{(1/CGF) - 1}{CV}\right)$$

Implied Percentile – Lognormal

- Without loss of generality (w.l.o.g.), assume lognormal mean of one (1.0)
- Then the mean and standard deviation of the related normal are:

$$\mu = \ln\left(\frac{1}{\sqrt{1+CV^2}}\right) \quad \sigma = \sqrt{\ln(1+CV^2)}$$

- We are looking for the percentile of the reciprocal of CGF (unit space)

$$X = e^Y \sim \text{LogN}(\mu, \sigma) \Leftrightarrow Y = \ln X \sim N(\mu, \sigma) \Leftrightarrow Z = \frac{Y - \mu}{\sigma} \sim N(0,1)$$

$$p = \Phi\left(\frac{\ln(1/CGF) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln\left(\frac{\sqrt{1+CV^2}}{CGF}\right)}{\sqrt{\ln(1+CV^2)}}\right)$$

Percentile-to-CGF Conversion

- This table shows the implied CGF for a given percentile of the Point Estimate for various CVs of a Normal distribution

NORMAL	Coefficient of Variation (CV)									
p	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0.05	1.09	1.20	1.33	1.49	1.49	1.49	2.36	2.92	3.85	5.63
0.10	1.07	1.15	1.24	1.34	1.34	1.34	1.81	2.05	2.36	2.78
0.15	1.05	1.12	1.18	1.26	1.26	1.26	1.57	1.71	1.87	2.08
0.20	1.04	1.09	1.14	1.20	1.20	1.20	1.42	1.51	1.61	1.73
0.25	1.03	1.07	1.11	1.16	1.16	1.16	1.31	1.37	1.44	1.51
0.30	1.03	1.06	1.09	1.12	1.12	1.12	1.22	1.27	1.31	1.36
0.35	1.02	1.04	1.06	1.08	1.08	1.08	1.16	1.18	1.21	1.24
0.40	1.01	1.03	1.04	1.05	1.05	1.05	1.10	1.11	1.13	1.15
0.45	1.01	1.01	1.02	1.03	1.03	1.03	1.05	1.05	1.06	1.07
0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.55	0.99	0.99	0.98	0.98	0.98	0.98	0.96	0.95	0.95	0.94
0.60	0.99	0.98	0.96	0.95	0.95	0.95	0.92	0.91	0.90	0.89
0.65	0.98	0.96	0.95	0.93	0.93	0.93	0.88	0.87	0.85	0.84
0.70	0.97	0.95	0.93	0.91	0.91	0.91	0.84	0.83	0.81	0.79
0.75	0.97	0.94	0.91	0.88	0.88	0.88	0.81	0.79	0.77	0.75
0.80	0.96	0.92	0.89	0.86	0.86	0.86	0.77	0.75	0.73	0.70
0.85	0.95	0.91	0.87	0.83	0.83	0.83	0.73	0.71	0.68	0.66
0.90	0.94	0.89	0.84	0.80	0.80	0.80	0.69	0.66	0.63	0.61
0.95	0.92	0.86	0.80	0.75	0.75	0.75	0.63	0.60	0.57	0.55

Percentile-to-CGF Conversion

- This table shows the implied CGF for a given percentile of the Point Estimate for various CVs of a Lognormal distribution

LOGNORMAL	Coefficient of Variation (CV)									
p	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0.05	1.09	1.18	1.29	1.41	1.55	1.69	1.85	2.03	2.22	2.43
0.10	1.07	1.14	1.22	1.31	1.41	1.52	1.64	1.76	1.90	2.05
0.15	1.05	1.11	1.18	1.25	1.33	1.42	1.51	1.61	1.71	1.82
0.20	1.04	1.09	1.15	1.20	1.27	1.34	1.41	1.49	1.57	1.66
0.25	1.04	1.07	1.12	1.17	1.22	1.27	1.33	1.40	1.46	1.54
0.30	1.03	1.06	1.09	1.13	1.17	1.22	1.27	1.32	1.37	1.43
0.35	1.02	1.04	1.07	1.10	1.13	1.17	1.21	1.25	1.29	1.34
0.40	1.01	1.03	1.05	1.07	1.10	1.12	1.15	1.19	1.22	1.26
0.45	1.01	1.02	1.03	1.05	1.06	1.08	1.11	1.13	1.16	1.19
0.50	1.00	1.00	1.01	1.02	1.03	1.04	1.06	1.08	1.10	1.12
0.55	0.99	0.99	0.99	0.99	1.00	1.01	1.02	1.03	1.04	1.05
0.60	0.99	0.98	0.97	0.97	0.97	0.97	0.97	0.98	0.98	0.99
0.65	0.98	0.97	0.95	0.94	0.94	0.93	0.93	0.93	0.93	0.93
0.70	0.98	0.95	0.94	0.92	0.91	0.90	0.89	0.88	0.88	0.87
0.75	0.97	0.94	0.91	0.89	0.87	0.86	0.84	0.83	0.82	0.81
0.80	0.96	0.92	0.89	0.86	0.84	0.82	0.80	0.78	0.76	0.75
0.85	0.95	0.91	0.87	0.83	0.80	0.77	0.74	0.72	0.70	0.69
0.90	0.94	0.88	0.84	0.79	0.75	0.72	0.69	0.66	0.63	0.61
0.95	0.92	0.85	0.79	0.74	0.69	0.64	0.61	0.57	0.54	0.51