



Diagnosing the Top-Level CV: An Alternative Approach

Daniel J. Andelin
United States Department of the Navy

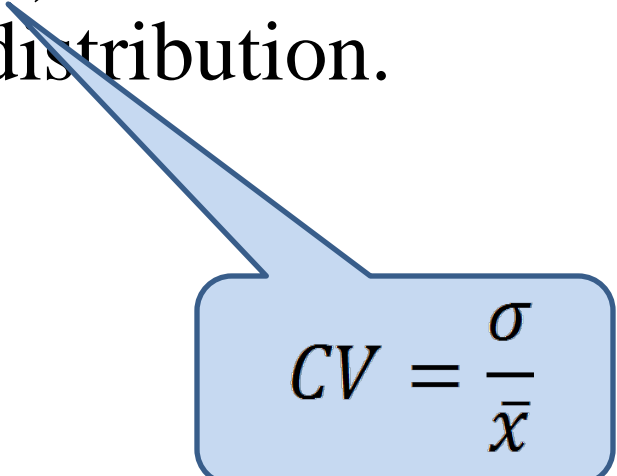
SCEA/ISPA Joint Annual Conference
June 2012

Outline

- Background
- Two Questions
- Weighted Average CV
- Partial CV Representation
 - “Zero-Correlation” Term
 - Correlation Effect
- Two Questions Revisited
- Diagnosing the CV
 - A Mental Path
 - Correlation
 - Partial CV
- Conclusions

Background

- “The point estimate is always wrong.”
- Cost estimates are uncertain positions represented by probability densities.
- Coefficient of Variation (CV) is a useful measure of the spread of a distribution.
 - Relative measure
 - Intuitive
 - Easily comparable


$$CV = \frac{\sigma}{\bar{x}}$$

Background

- The Air Force Cost Risk and Uncertainty Handbook (AF CRUH) contains rule-of-thumb guidance for expected CVs by commodity and life-cycle phase.
 - Examples:
 - “0.35-0.45 typical for space systems and software intensive projects”
 - “0.25-0.35 typical for aircraft and similar complexity hardware”
 - Note: These examples given at “early stages” of the program.
- Comparison to these and other standards allows the analyst to determine whether an appropriate amount of uncertainty is built into the cost estimate.

See Air Force Cost Analysis Agency (AFCAA) Cost Risk and Uncertainty Handbook (AF CRUH), April 2007, 26-27

Two Questions

- When the CV falls outside the expected range, it begs two questions:
 1. “Why is my CV so low (or high)?”
 - What are the contributing factors to the CV?
 - What drives it up or down?
 2. “What can (or should) I do to fix it?”
 - Where can I add additional or remove unwanted uncertainty?
 - What are the limits on my CV (i.e. how high or low can it go)?

Two Questions

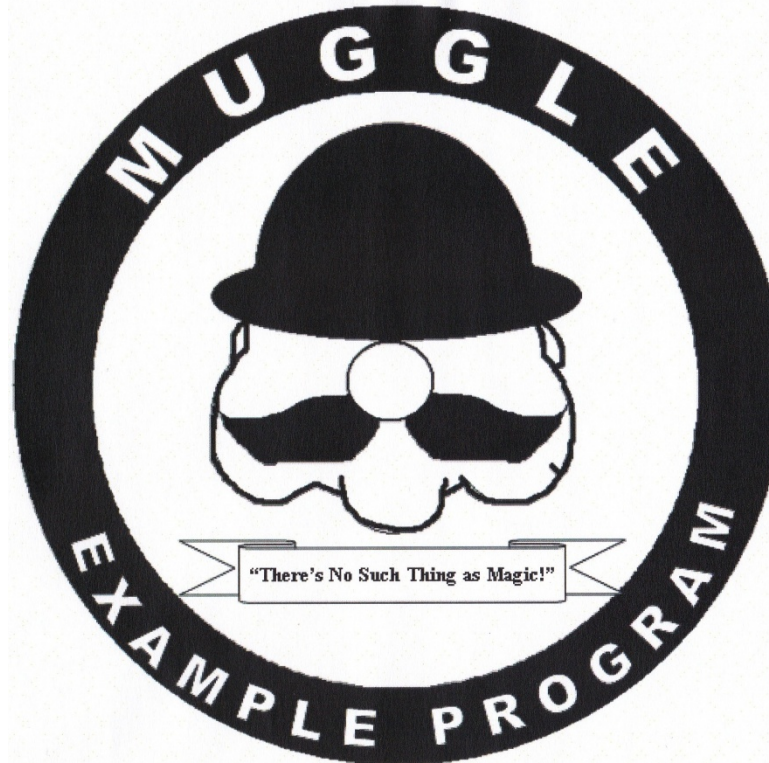
- These two questions can be answered by a formal risk analysis:
 - Variance analysis
 - Sensitivity tests (Tornado charts)
 - Trial and error
- Though worthwhile, these options are time consuming in both computation and analysis of results.
- What if a quick diagnostic is needed (e.g. during a brief)?

Weighted Average CV

- When a diagnostic of the CV is needed on the fly, it is tempting to approximate by taking a weighted average of the CV of the lower level elements.
- This approach is:
 - Straight forward
 - Intuitive
 - Easy
 - **And will give you the wrong answer!**

Example Program

- To illustrate, consider a hypothetical DoD Acquisition Program entering E&MD Phase.
- **Made-Up, Generalized, Good-Looking Example**

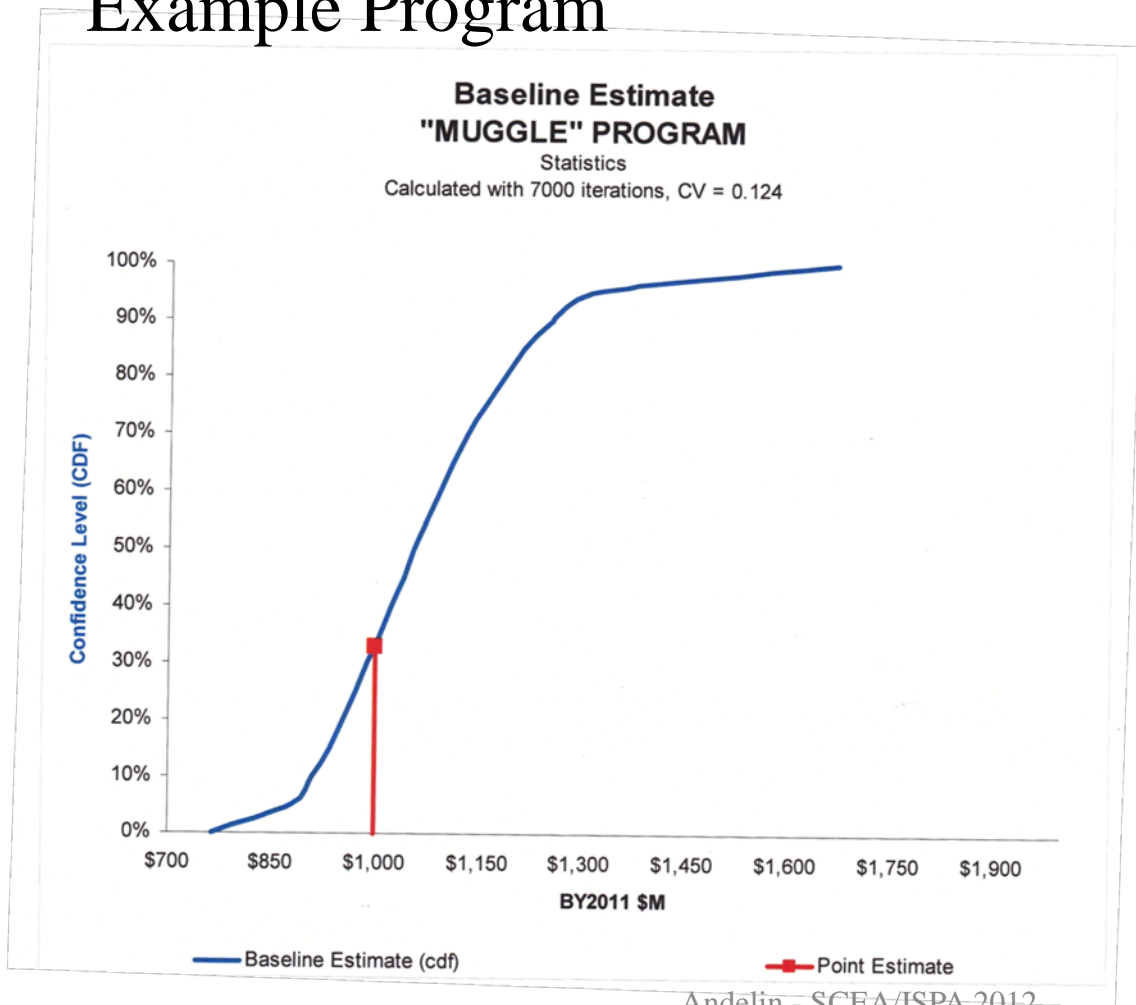


Apologies to J.K. Rowling



Example

- The following is the CDF (S-curve) for the MUGGLE Example Program



- PE = \$1.0B
- CV = 0.124
 - Unexpectedly low for the Program's risk, complexity, etc.
- Senior SYSCOM analyst asks:
 - “Why is your CV so low?”
 - “What can we do to fix it?”



Example

- The intuitive approach—examine the weighted average of the lower level CVs to see where the large contributors are:

Costs in BY2011 \$M				
WBS	Point Estimate	Mean	Std Dev	CV
"MUGGLE" PROGRAM	\$ 1,000.0	\$ 1,072.2	\$ 133.2	0.1242
Prime Mission Product	\$ 500.0	\$ 546.2	\$ 108.0	0.1977
System Engineering	\$ 150.0	\$ 150.8	\$ 15.2	0.1009
Program Management	\$ 150.0	\$ 150.8	\$ 15.2	0.1009
System Test & Evaluation	\$ 100.0	\$ 120.4	\$ 26.0	0.2161
Training	\$ 9.0	\$ 7.2	\$ 2.3	0.3270
Data	\$ 16.0	\$ 20.0	\$ 7.2	0.3618
Support Equipment	\$ 75.0	\$ 76.9	\$ 26.8	0.3489



Example

- Back-of-the-envelope analysis:
 - Prime Mission Product (PMP) is the largest WBS element, representing 50% of the Point Estimate (PE).
 - *This is likely an influential element!*
 - PMP has a CV of ~0.2
 - *Shouldn't that bring the total CV up?*
 - *Maybe the other elements are dragging it down?*

Costs in BY2011 \$M			
WBS	Point Estimate	% of PE	CV
"MUGGLE" PROGRAM	\$ 1,000.0		0.1242
Prime Mission Product	\$ 500.0	50.0%	0.1977
System Engineering	\$ 150.0	15.0%	0.1009
Program Management	\$ 150.0	15.0%	0.1009
System Test & Evaluation	\$ 100.0	10.0%	0.2161
Training	\$ 9.0	0.9%	0.3270
Data	\$ 16.0	1.6%	0.3618
Support Equipment	\$ 75.0	7.5%	0.3489



Example

- System Engineering (SE) and Program Management (PM) together make up 30% of the PE.
- Both SE and PM have a lower CV—0.1009
 - *Is that enough to bring the total CV down to 0.124???*

Costs in BY2011 \$M			
WBS	Point Estimate	% of PE	CV
"MUGGLE" PROGRAM	\$ 1,000.0		0.1242
Prime Mission Product	\$ 500.0	50.0%	0.1977
System Engineering	\$ 150.0	15.0%	0.1009
Program Management	\$ 150.0	15.0%	0.1009
System Test & Evaluation	\$ 100.0	10.0%	0.2161
Training	\$ 9.0	0.9%	0.3270
Data	\$ 16.0	1.6%	0.3618
Support Equipment	\$ 75.0	7.5%	0.3489



Example

- System Test & Evaluation (STE) makes up 10% of the PE and has a CV of ~0.22.
- The rest of the WBS elements (Training, Data, and Support Equipment) together make up the remaining 10% of the PE and have a composite CV of roughly 0.35.
 - *How do these elements affect the CV?*

Costs in BY2011 \$M			
WBS	Point Estimate	% of PE	CV
"MUGGLE" PROGRAM	\$ 1,000.0		0.1242
Prime Mission Product	\$ 500.0	50.0%	0.1977
System Engineering	\$ 150.0	15.0%	0.1009
Program Management	\$ 150.0	15.0%	0.1009
System Test & Evaluation	\$ 100.0	10.0%	0.2161
Training	\$ 9.0	0.9%	0.3270
Data	\$ 16.0	1.6%	0.3618
Support Equipment	\$ 75.0	7.5%	0.3489



Example

- Weighted Average

$$50\% * 0.2 + 30\% * 0.1 + 10\% * 0.22 + 10\% * 0.35$$

$$= (0.1) + (0.03) + (0.022) + (0.035) \rightarrow 0.187$$

Costs in BY2011 \$M			
WBS	Point Estimate	% of PE	CV
"MUGGLE" PROGRAM	\$ 1,000.0		0.1242
Prime Mission Product	\$ 500.0	50.0%	0.1977
System Engineering	\$ 150.0	30.0%	0.1009
Program Management	\$ 150.0		
System Test & Evaluation	\$ 100.0	10.0%	0.2161
Training	\$ 9.0		
Data	\$ 16.0	10.0%	0.35
Support Equipment	\$ 75.0		

This is close to what we might expect for a weighted average, but it does not reflect reality!

Weighted Average CV

- The weighted average CV ($CV_{w,ave}$) is not a good approximation of the total CV for two fundamental reasons:
 1. The CV is dependent on the standard deviation, which sums **quadratically**, not linearly.
 - The standard deviation of the sum is not equal to the sum of the standard deviations.
 - This has a narrowing effect on the sum.
 2. The weighted average ignores the effect of correlation between WBS elements.
 - This has a widening effect on the parent distribution.
- These effects mask each other and obscure the nature of the CV when viewed together.

Weighted Average CV

- Mathematically unsound as an approximation for a top-level CV.
- **HOWEVER**, it *can* be used as a starting point for quick diagnostics
 - Adjustments must be made.
- The terms that, when summed, equal $CV_{w.ave}$ can be called the *partial coefficient of variation* (pCV) of each child WBS element.
- The parent (system level) CV can be expressed in terms of the pCV.
 - Clearly shows the proper weighted average CV (added in quadrature) and correlation terms separately.
 - Leads to a better understanding of the true contributors to the overall CV of an estimate.

Notation

- Before delving into any sort of mathematical representation, it is worthwhile to clearly define the notation to be used.
- This presentation adopts the following conventions:
 - Parent WBS element: X
 - Child WBS elements: x_j, x_k
 - # of children to a given parent: n
 - Weighting factor: p
- In addition, a variable which is generally given its own symbol (e.g. standard deviation), but which pertains to a particular WBS element (e.g.) is denoted by a single subscript:

σ_X (st. dev. of X)

CV_k (CV of x_k)

The Partial CV (pCV)

- We coin the term “partial coefficient of variation” to refer to the weighted child-level CVs, the sum of which make up the weighted average CV.

$$CV_{w.ave} = \sum_{j=1}^n (p_j \cdot CV_j) = \sum_{j=1}^n p_j CV_j$$

- It is important to be clear on how to calculate the weighting factor p_j .
- Since the mean is the only point on a probability distribution that sums linearly, it is logical to use the relative means as weights.

$$p_j = \frac{\bar{x}_j}{\bar{X}}$$

- The partial CV (pCV_j) of an element x_j is defined as its coefficient of variation (CV_j) times its weighting factor (p_j).

$$pCV_j = p_j \cdot CV_j$$

The Partial CV (pCV)

- The sum of the pCV is the *weighted average CV*.
 - Except for the most special of circumstances, this is *not* equal to the true CV of the parent distribution.

$$\sum_{j=1}^n pCV_j = CV_{w.ave} \neq CV_X$$

- However, the true CV can be written in terms of the pCV.

Note: The acronym pCV can be singular *or* plural (as in partial coefficients of variation).

The pCV Representation

- To arrive at the partial CV representation of CV_X , start with the following expression for the variance of X :

Given in Alfred Smith, "Effective Use of Cost Reports" (SCEA/ISPA 2011) and derived independently by this author.

$$\sigma_X^2 = \sum_{j=1}^n \sum_{k=1}^n r_{j,k} \sigma_j \sigma_k$$

- From the definition of the CV:

$$(CV_X \cdot \bar{X})^2 = \sum_{j=1}^n \sum_{k=1}^n r_{j,k} (CV_j \cdot \bar{x}_j) (CV_k \cdot \bar{x}_k)$$

- Divide through by \bar{X}^2 :

$$(CV_X)^2 = \sum_{j=1}^n \sum_{k=1}^n r_{j,k} \left(\frac{\bar{x}_j}{\bar{X}} \cdot CV_j \right) \left(\frac{\bar{x}_k}{\bar{X}} \cdot CV_k \right) = \sum_{j=1}^n \sum_{k=1}^n r_{j,k} (p_j \cdot CV_j) (p_k \cdot CV_k)$$


$$p_j = \frac{\bar{x}_j}{\bar{X}}$$

$$CV = \frac{\sigma}{\bar{x}}$$

The pCV Representation

- From the previous result:

$$(CV_X)^2 = \sum_{j=1}^n \sum_{k=1}^n r_{j,k} \left(\frac{\bar{x}_j}{\bar{X}} \cdot CV_j \right) \left(\frac{\bar{x}_k}{\bar{X}} \cdot CV_k \right) = \sum_{j=1}^n \sum_{k=1}^n r_{j,k} (p_j \cdot CV_j) (p_k \cdot CV_k)$$



$$CV_X = \sqrt{\sum_{j=1}^n \sum_{k=1}^n r_{j,k} (pCV_j) (pCV_k)}$$

- This is the partial CV representation of the coefficient of variation of a parent WBS (X).*

The pCV Representation: Vector Form

$$CV_X = \sqrt{\sum_{j=1}^n \sum_{k=1}^n r_{j,k} (pCV_j)(pCV_k)} = \sqrt{(\overrightarrow{pCV}_T \cdot \hat{R} \cdot \overrightarrow{pCV})}$$

- For the math enthusiast, notice that the expression under the radical is the inner product of the pCV vector with itself, after being acted upon by the fully populated correlation matrix R .
- Equivalently, it is the result of the Hermitian operator \hat{R} acting on the pCV vector.
- This becomes useful in a couple special cases.

A Hermitian matrix is equal to its own conjugate transpose. (See <http://mathworld.wolfram.com/HermitianOperator.html>)

The pCV Representation

- The utility of the pCV representation is made clear by separating the terms for which $j = k$ from the cross terms.

$$CV_X = \sqrt{\sum_{j=1}^n (pCV_j)^2 + \sum_{j=1}^n \sum_{\substack{k=1 \\ k \neq j}}^n r_{j,k} (pCV_j)(pCV_k)}$$

- The first set corresponds to the diagonal entries of the correlation matrix, which are all equal to one.

$R =$

	pCV_1	pCV_2	pCV_3	...	pCV_n	
Row: j	1	$r_{1,2}$	$r_{1,3}$...	$r_{1,n}$	pCV_1
	$r_{2,1}$	1	$r_{2,3}$...	$r_{2,n}$	pCV_2
	$r_{3,1}$	$r_{3,2}$	1	...	$r_{3,n}$	pCV_3

Column: k	$r_{n,1}$	$r_{n,2}$	$r_{n,3}$...	1	pCV_n

Andelin - SCEA/ISPA 2012

The pCV Representation

- The utility of the pCV representation is made clear by separating the terms for which $j = k$ from the cross terms.

$$CV_X = \sqrt{\sum_{j=1}^n (pCV_j)^2 + 2 \sum_{j=1}^{n-1} \sum_{k=j+1}^n r_{j,k} (pCV_j)(pCV_k)}$$

- The first set corresponds to the diagonal entries of the correlation matrix, which are all equal to one.
- Symmetry in the off-diagonal entries ($r_{j,k} = r_{k,j}$) leads to a factor of two in the cross terms.

$R =$

	pCV_1	pCV_2	pCV_3	...	pCV_n	
Row: j	1	$r_{1,2}$	$r_{1,3}$...	$r_{1,n}$	pCV_1
	$r_{2,1}$	1	$r_{2,3}$...	$r_{2,n}$	pCV_2
	$r_{3,1}$	$r_{3,2}$	1	...	$r_{3,n}$	pCV_3

Column: k	$r_{n,1}$	$r_{n,2}$	$r_{n,3}$...	1	pCV_n

The pCV Representation

- The utility of the pCV representation is made clear by separating the terms for which $j = k$ from the cross terms.

$$CV_X = \sqrt{\underbrace{\sum_{j=1}^n (pCV_j)^2}_{\text{Zero-Correlation Term}} + 2 \underbrace{\sum_{j=1}^{n-1} \sum_{k=j+1}^n r_{j,k} (pCV_j)(pCV_k)}_{\text{Correlation Effect Term}}}$$

- The first set corresponds to the diagonal entries of the correlation matrix, which are all equal to one.
 - Symmetry in the off-diagonal entries ($r_{j,k} = r_{k,j}$) leads to a factor of two in the cross terms.
- This form illustrates the two components that make up the CV:
 - “Zero-correlation” term, equal to the sum (in quadrature) of the pCV.
 - “Correlation effect” term that accounts for correlation between WBS elements
- Note that because of the radical, the CV is not equal to the sum of these terms.**

“Zero-Correlation” Term

- If the correlation between different WBS elements were zero (i.e. if all WBS elements were truly independent), the expression for the true CV would reduce to the form shown below:
 - Call this the “zero-correlation CV.”

$$CV_{zero\ corr} = \sqrt{\sum_{j=1}^n (pCV_j)^2}$$

- The zero-correlation CV is equal to the root-sum-square of the pCV.
- While similar to the intuitive weighted average CV, the two are clearly not equal. However, $CV_{w.ave}$ can be used as a starting point to get to $CV_{zero\ corr}$, provided an adjustment is made:

$$CV_{zero\ corr} = f \cdot CV_{w.ave}$$

“Zero-Correlation” Term- f Factor

$$CV_{zero\ corr} = f \cdot CV_{w.ave}$$

- f is defined as the ratio of the zero-correlation CV to the weighted average CV.
 - Easy to calculate with spreadsheet software (e.g. Excel) or even a scientific calculator.
- Because $CV_{w.ave}$ is a linear sum, f can be distributed throughout its terms.
 - Thus, $CV_{zero\ corr}$ can be thought of as the adjusted $CV_{w.ave}$ OR as the sum of the adjusted pCV.
 - Adjustment in both cases made via multiplication by f .

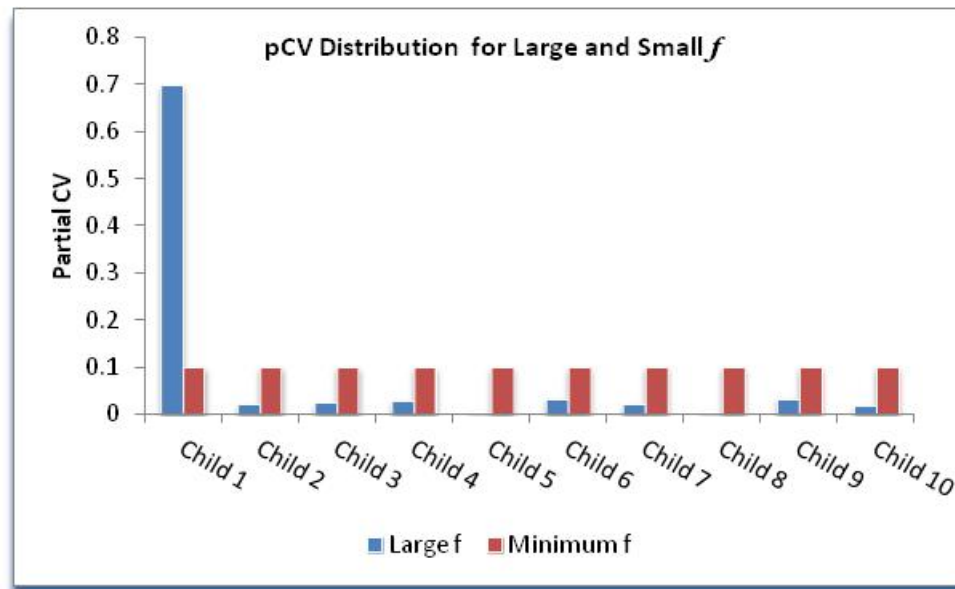
“Zero-Correlation” Term- f Factor

$$CV_{zero\ corr} = f \cdot CV_{w.ave}$$

- $CV_{w.ave}$ is much easier to calculate mentally than $CV_{zero\ corr}$.
- When a quick, back-of-the-envelope calculation is needed for $CV_{zero\ corr}$, start with the intuitive $CV_{w.ave}$ and multiply by f .
- Though not simple to mentally calculate f , rules of thumb can help approximate.

“Zero-Correlation” Term- f Factor

- f depends on the number of WBS elements n and on how *evenly* the pCV are distributed.
 - Minimum when pCV are all equal
 - Maximum when one pCV dominates the others



- The proof is beyond the scope of this presentation, but f is bounded from $n^{-1/2}$ to 1.

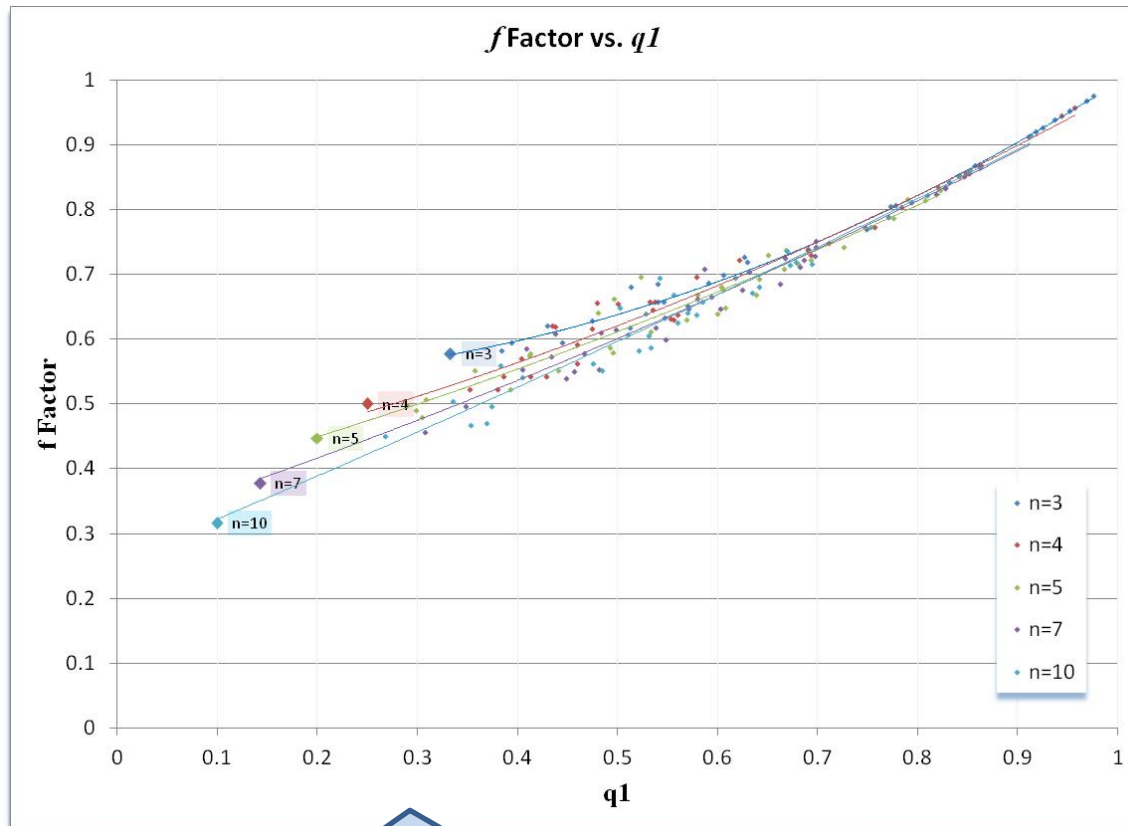
“Zero-Correlation” Term- f Factor

- It is helpful to consider the behavior of f as the leading pCV varies.
- For consistency and comparability, normalize the pCV.
- Define q to be the ratio of the pCV to the sum of all pCV:

$$q \equiv \frac{pCV}{\sum_j^n (pCV_j)}$$

- Define q_1 to be the largest value of q (the largest *normalized* pCV) for a given set of child WBS elements.

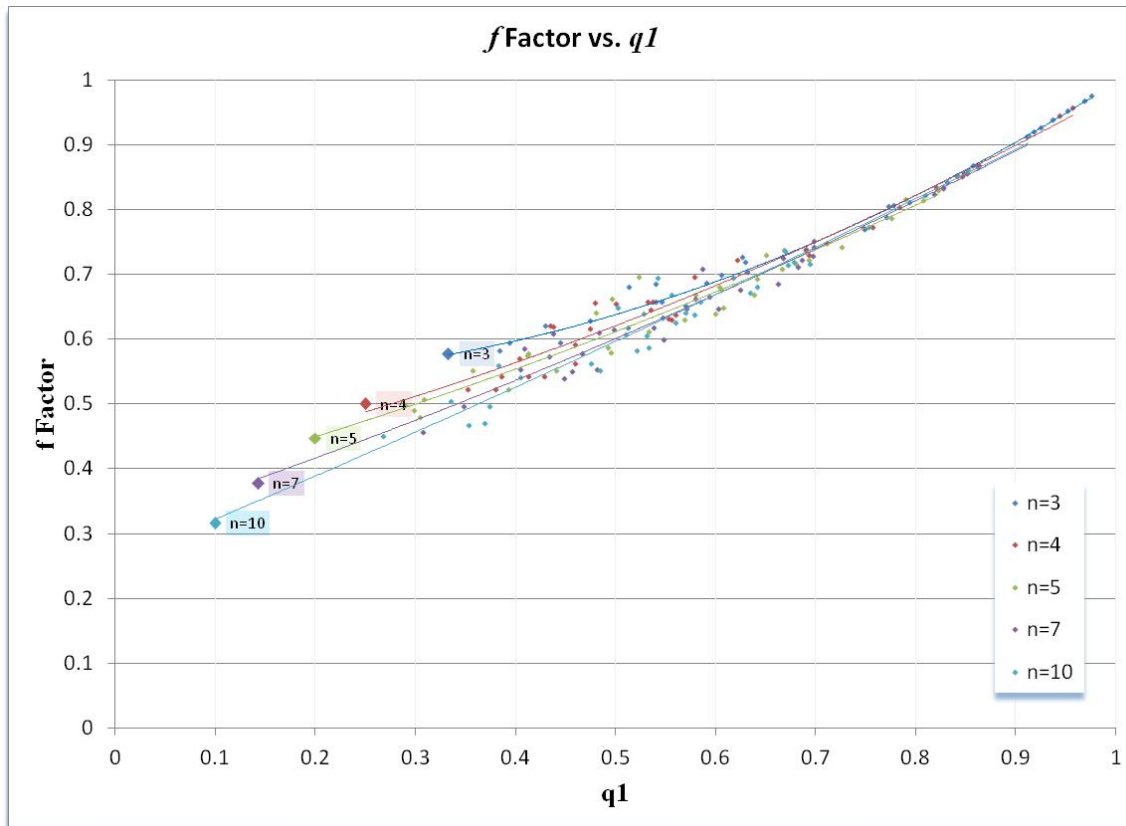
“Zero-Correlation” Term- f Factor



Toy data, consisting of random samples of q , engineered so that for each set of data, the sum of q values is equal to one.

- The graph to the left shows the behavior of f vs. q_1 for various values of n .
- For small q_1 , the data are somewhat scattered.
 - Influenced by n and varying distributions of non-leading q .
- For large q_1 , the data converge and follow a linear trend.
 - Influence of q_1 dominates.
 - $f \sim q_1$

f Factor: Rules of Thumb



- If q_1 is:

> 0.70 :

$$f \sim q_1$$

$0.40 - 0.70$:

$f \sim \frac{1}{2}$ for low q_1 and/or high n

$f \sim \frac{2}{3}$ for med q_1 and/or med n

$f \sim \frac{3}{4}$ for high q_1 and/or low n

< 0.40 :

$$f \sim n^{-1/2}$$



Example

- To mentally calculate the zero-correlation CV, start with the pCV.

Costs in BY2011 \$M					
WBS	Point Estimate	Mean	CV	Weight (p)	pCV
				$p = \text{Mean}/\text{Total Mean}$	$pCV = p * CV$
"MUGGLE" PROGRAM	\$ 1,000.0	\$ 1,072.2	0.1242	100.0%	N/A
Prime Mission Product	\$ 500.0	\$ 546.2	0.1977	50.9%	0.10
System Engineering	\$ 150.0	\$ 150.8	0.1009	14.1%	0.01
Program Management	\$ 150.0	\$ 150.8	0.1009	14.1%	0.01
System Test & Evaluation	\$ 100.0	\$ 120.4	0.2161	11.2%	0.02
Training	\$ 9.0	\$ 7.2	0.3270	0.7%	0.00
Data	\$ 16.0	\$ 20.0	0.3618	1.9%	0.01
Support Equipment	\$ 75.0	\$ 76.9	0.3489	7.2%	0.03

Weighted Average CV is sum of pCV ~ **0.18**



Example

- The leading pCV is Prime Mission Product.
 - A little more than 50% of the total pCV.
- Since $n = 6$ and $q_1 \sim 0.5$ are both fairly middle-range values, approximate $f \sim 2/3$.
- $CV_{zero\ corr} = f * CV_{w.ave} = (2/3)*(0.18) \sim \mathbf{0.12}$

Costs in BY2011 \$M	
WBS	pCV
	pCV = p*CV
"MUGGLE" PROGRAM	N/A
Prime Mission Product	0.10
System Engineering	0.01
Program Management	0.01
System Test & Evaluation	0.02
Training	0.00
Data	0.01
Support Equipment	0.03

Correlation-Effect Term

$$CV_X = \sqrt{\sum_{j=1}^n (pCV_j)^2 + 2 \sum_{j=1}^{n-1} \sum_{k=j+1}^n r_{j,k} (pCV_j)(pCV_k)}$$

- The zero-correlation case is theoretical only.
 - All real estimates contain some degree of correlation between WBS elements.
 - Effects of correlation must be added to the zero-correlation CV.
- It is not easy to mentally calculate the correlation-effect term.
 - Many entries
 - Correlation matrix not typically simple (sometimes not even immediately available)
 - Does not add linearly
 - No easy rules-of-thumb

Correlation-Effect Term

$$CV_X = \sqrt{\sum_{j=1}^n (pCV_j)^2 + 2 \sum_{j=1}^{n-1} \sum_{k=j+1}^n r_{j,k} (pCV_j)(pCV_k)}$$

- **GOOD NEWS!** We don't have to calculate this term!
 - Since CV_X is typically known, simply subtract the calculated zero-correlation CV from the known true coefficient of variation (CV_X).
 - The difference between the true CV and the zero-correlation CV is the additional CV that you get because of the effects of correlation.

$$CV_X - CV_{zero\ corr} = \Delta CV_{corr}$$

- This is not equal to the circled term, but it doesn't have to be!

Correlation Effect Term: Perfect Correlation Case

- There is an interesting result for the perfect correlation case.
 - $r_{j,k} = 1$ for all j,k
- Recall that the square of the true CV can be given by

$$\overline{pCV}_T \cdot \hat{R} \cdot \overline{pCV}$$



pCV_1	pCV_2	pCV_3	...	pCV_n						
					1	$r_{1,2}$	$r_{1,3}$...	$r_{1,n}$	pCV_1
					$r_{2,1}$	1	$r_{2,3}$...	$r_{2,n}$	pCV_2
					$r_{3,1}$	$r_{3,2}$	1	...	$r_{3,n}$	pCV_3
				
					$r_{n,1}$	$r_{n,2}$	$r_{n,3}$...	1	pCV_n

Correlation Effect Term: Perfect Correlation Case

- If all entries in the correlation matrix are equal to one...
...the product of the matrix multiplication is equal to the square of the sum of the pCV vector entries!

$$\overrightarrow{pCV}_T \cdot \hat{R} \cdot \overrightarrow{pCV}$$



pCV_1	pCV_2	pCV_3	\dots	pCV_n	pCV_1	pCV_2	pCV_3	\dots	pCV_n
					$(pCV_1 + pCV_2 + pCV_3 + \dots + pCV_n)^*$				
					$(pCV_1 + pCV_2 + pCV_3 + \dots + pCV_n) =$				
					$\left(\sum_{j=1}^n pCV_j \right)^2$				
					$pCV_1 + pCV_2 + pCV_3 + \dots + pCV_n$				
					$pCV_1 + pCV_2 + pCV_3 + \dots + pCV_n$				
					$pCV_1 + pCV_2 + pCV_3 + \dots + pCV_n$				
					$pCV_1 + pCV_2 + pCV_3 + \dots + pCV_n$				

Correlation Effect Term: Perfect Correlation Case

- In other words, the perfect correlation case is that one special case for which $CV_{w.ave}$ is the true CV!
- Of course, this is just as unrealistic as the zero-correlation case, but it is useful to note, because it puts an upper limit on the additional CV that can be gained by adding correlation to the child WBS elements.
 - $CV_{w.ave}$ is the highest CV that can be obtained, given a set of pCV .

CAUTION: If additional correlation is added to the model *inputs*, this will change the pCV of the child WBS elements. So, by applying correlation at level lower than the immediate children, the parent level CV can actually be pushed higher than shown here.

Two Questions Revisited

- With the pCV representation in hand, we can now revisit the two questions discussed earlier:
 1. “Why is my CV so low (or high)?”
 2. “What can (or should) I do to fix it?”

Diagnosing the CV: A Mental Path

- Based on the previous discussion, there is a logical mental path that can be followed to understand what drives (mathematically) the true CV:

1. Start with the weighted average CV

$CV_{w.ave}$

2. Multiply by f to account for summing in quadrature

$CV_{zero\ corr}$

3. Add the effects of correlation

CV_X



Example

- Step 1: Calculate the weighted average CV
 - Calculate pCV for each WBS element and sum

Costs in BY2011 \$M					
WBS	Point Estimate	Mean	CV	Weight (p)	pCV
				p = Mean/Total Mean	pCV = p*CV
"MUGGLE" PROGRAM	\$ 1,000.0	\$ 1,072.2	0.1242	100.0%	N/A
Prime Mission Product	\$ 500.0	\$ 546.2	0.1977	50.9%	0.1007
System Engineering	\$ 150.0	\$ 150.8	0.1009	14.1%	0.0142
Program Management	\$ 150.0	\$ 150.8	0.1009	14.1%	0.0142
System Test & Evaluation	\$ 100.0	\$ 120.4	0.2161	11.2%	0.0243
Training	\$ 9.0	\$ 7.2	0.3270	0.7%	0.0022
Data	\$ 16.0	\$ 20.0	0.3618	1.9%	0.0068
Support Equipment	\$ 75.0	\$ 76.9	0.3489	7.2%	0.0250

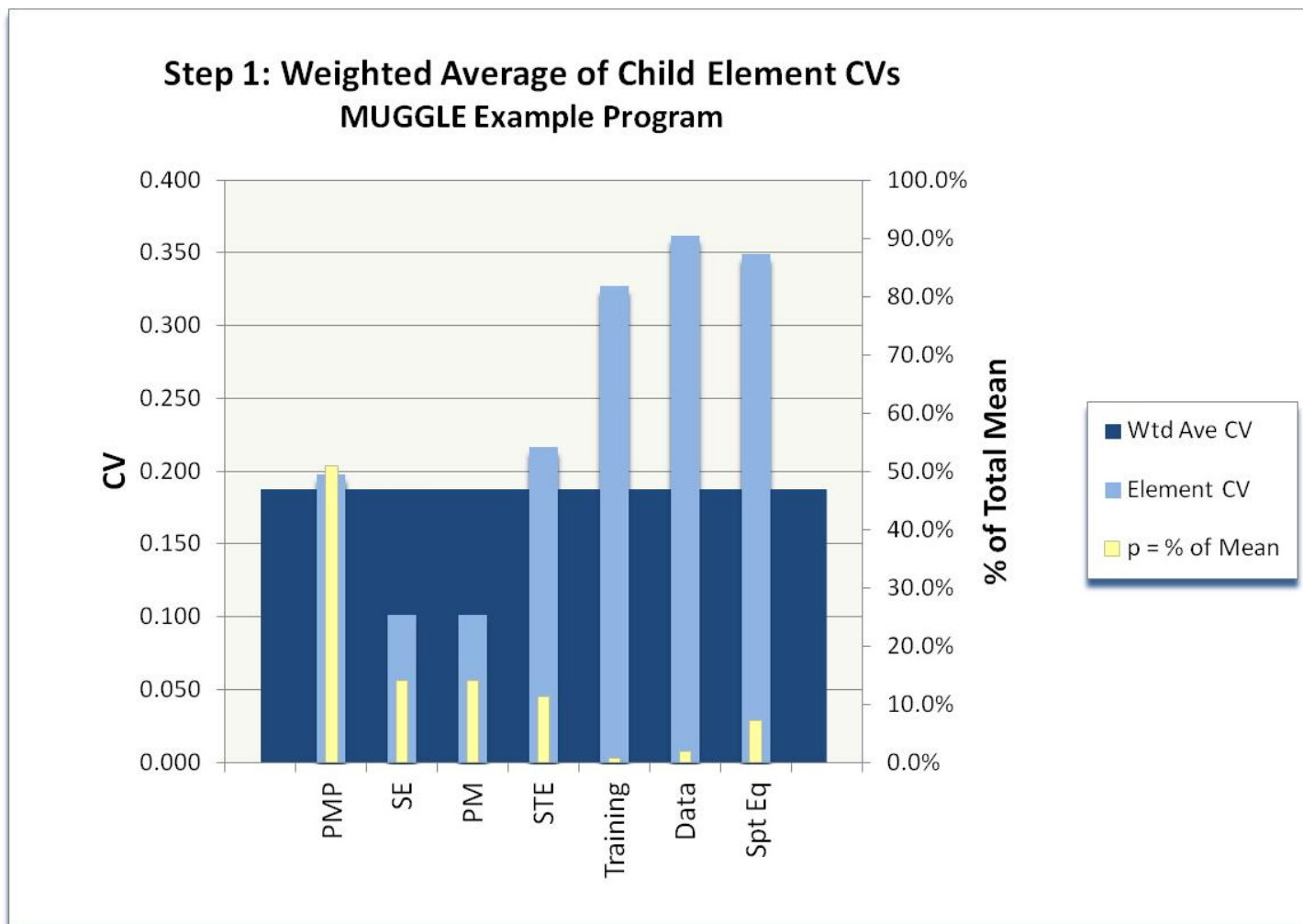
$$B \quad \Sigma(pCV) = 0.1873$$

$$CV_{w.ave} = 0.1873$$

$$- CV_{w.ave} = \Sigma(pCV) = 0.187$$



Example: Graphical Representation of Step 1





Example

- Step 2: Calculate the zero-correlation CV
 - Use f to adjust for summation in quadrature

Costs in BY2011 \$M								
WBS	Point Estimate	Mean	CV	Weight (p)	pCV	Sq. pCV	Adj. CV	Adj. pCV
				p = Mean/Total Mean	pCV = p*CV	pCV^2	Adj CV = f*CV	Adj pCV = f*pCV
"MUGGLE" PROGRAM	\$ 1,000.0	\$ 1,072.2	0.1242	100.0%	N/A	N/A	N/A	N/A
Prime Mission Product	\$ 500.0	\$ 546.2	0.1977	50.9%	0.1007	1.01E-02	0.1147	0.0584
System Engineering	\$ 150.0	\$ 150.8	0.1009	14.1%	0.0142	2.01E-04	0.0585	0.0082
Program Management	\$ 150.0	\$ 150.8	0.1009	14.1%	0.0142	2.01E-04	0.0585	0.0082
System Test & Evaluation	\$ 100.0	\$ 120.4	0.2161	11.2%	0.0243	5.89E-04	0.1254	0.0141
Training	\$ 9.0	\$ 7.2	0.3270	0.7%	0.0022	4.77E-06	0.1897	0.0013
Data	\$ 16.0	\$ 20.0	0.3618	1.9%	0.0068	4.57E-05	0.2099	0.0039
Support Equipment	\$ 75.0	\$ 76.9	0.3489	7.2%	0.0250	6.26E-04	0.2024	0.0145

B	$\Sigma(pCV) =$	0.1873
A	RSS (pCV) =	0.1087
	$f = (A/B) =$	0.58

$\Sigma(\text{Adj. pCV}) =$	0.1087
Adj. CV _{w.ave} =	0.1087

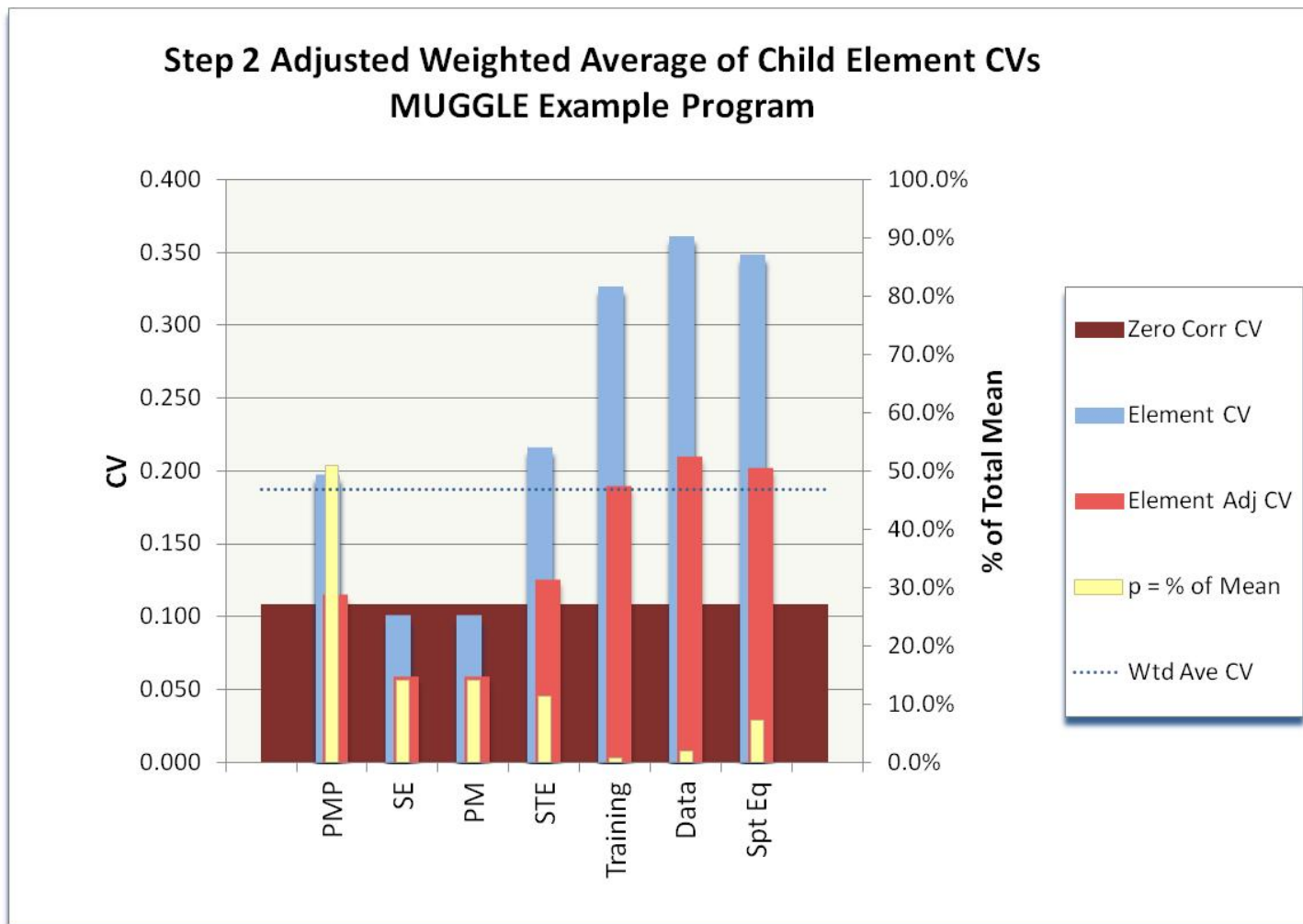
CV _{w.ave} =	0.1873
CV _{zero corr} =	0.1087

$$- CV_{zero\ corr} = f * CV_{w.ave} = \Sigma(f * pCV) = 0.1087$$

- Note that whether it is calculated as the product of f and $CV_{w.ave}$ or as the sum of the adjusted pCV, the result is the same.



Example: Graphical Representation of Step 2





Example

- Step 3: Add (or solve for) correlation effect
 - Additional CV due to correlation is equal to the difference between the true CV and the zero-correlation CV.

Costs in BY2011 \$M								
WBS	Point Estimate	Mean	CV	Weight (p)	pCV	Sq. pCV	Adj. CV	Adj. pCV
				$p = \text{Mean}/\text{Total Mean}$	$pCV = p * CV$	pCV^2	$\text{Adj CV} = f * CV$	$\text{Adj pCV} = f * pCV$
"MUGGLE" PROGRAM	\$ 1,000.0	\$ 1,072.2	0.1242	100.0%	N/A	N/A	N/A	N/A
Prime Mission Product	\$ 500.0	\$ 546.2	0.1977	50.9%	0.1007	1.01E-02	0.1147	0.0584
System Engineering	\$ 150.0	\$ 150.8	0.1009	14.1%	0.0142	2.01E-04	0.0585	0.0082
Program Management	\$ 150.0	\$ 150.8	0.1009	14.1%	0.0142	2.01E-04	0.0585	0.0082
System Test & Evaluation	\$ 100.0	\$ 120.4	0.2161	11.2%	0.0243	5.89E-04	0.1254	0.0141
Training	\$ 9.0	\$ 7.2	0.3270	0.7%	0.0022	4.77E-06	0.1897	0.0013
Data	\$ 16.0	\$ 20.0	0.3618	1.9%	0.0068	4.57E-05	0.2099	0.0039
Support Equipment	\$ 75.0	\$ 76.9	0.3489	7.2%	0.0250	6.26E-04	0.2024	0.0145

B	$\Sigma(pCV) =$	0.1873	$\Sigma(\text{Adj. pCV}) =$	0.1087
A	RSS (pCV) =	0.1087	Adj. CV _{w.ave} =	0.1087
	$f = (A/B) =$	0.58		

CV _{w.ave} =	0.1873
CV _{zero corr} =	0.1087
True CV =	0.1242
Corr. Effect =	0.0155

This small amount of correlation may be why the CV is so low!

- The presence of correlation increases the CV by 0.0155



Example:

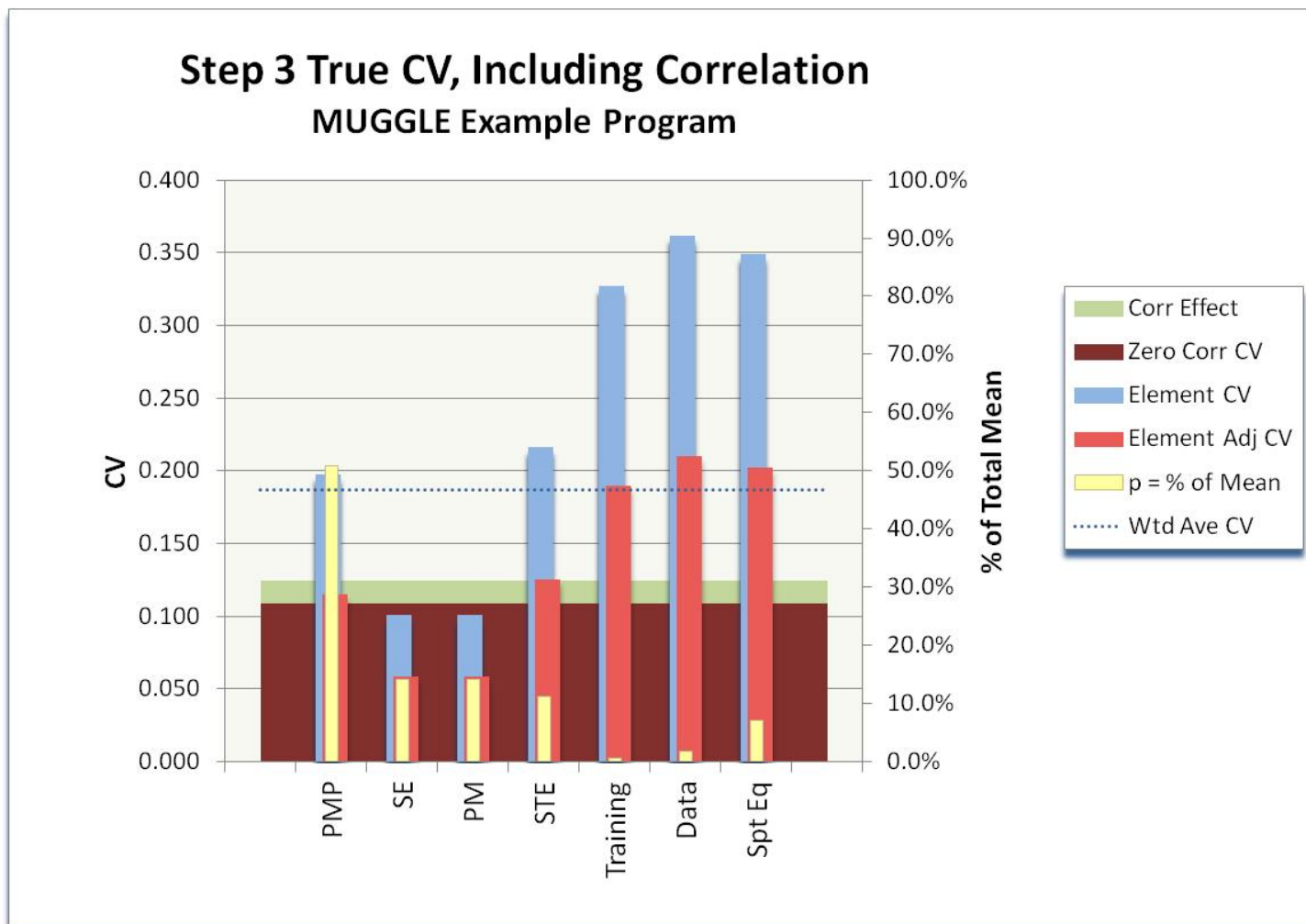
MUGGLE Correlation Matrix

- A glance at the correlation matrix for the estimate reveals that little to no correlation exists between most child elements.

	PMP	SE	PM	STE	Training	Data	Spt Eq
PMP	1	0.0	0.0	0.0	0.0	0.5	0.6
SE		1	0.0	0.0	0.0	0.0	0.0
PM			1	0.0	0.0	0.0	0.0
STE				1	0.0	0.1	0.0
Training					1	0.0	0.0
Data						1	0.3
Spt Eq							1



Example: Graphical Representation of Step 3



Two Questions Revisited: “What can we do?”

- Once the contributors to the CV are clearly identified, the question remains of what, if anything, should be done.
 - Ask yourself:
 - Do the distributions on the model inputs adequately describe what is known or uncertain about those parameters?
 - Has uncertainty been appropriately applied to the cost estimating relationships (CERs)?
 - Has correlation been given due diligence?
 - If the answer to all these questions is affirmative, it may not be appropriate to artificially alter the CV.
 - If it is believed that the CV reflects a deficiency in the application of uncertainty, there are a few quick exercises that can help resolve that deficiency.

Diagnosing the CV: Correlation

- First, consider the effects of correlation.
- The zero-correlation CV and the weighted average CV bound the effects of correlation on the child elements.
 - If the CV is high, $CV_{\text{zero corr}}$ is as low as it can go by adjusting correlation on the child elements.
 - If the CV is low, $CV_{\text{w.ave}}$ is as high as it can go by adjusting correlation on the child elements.

Assumes net positive correlation between WBS elements.

CAUTION: Remember the CV can go higher or lower if correlation is adjusted on the lower level inputs!

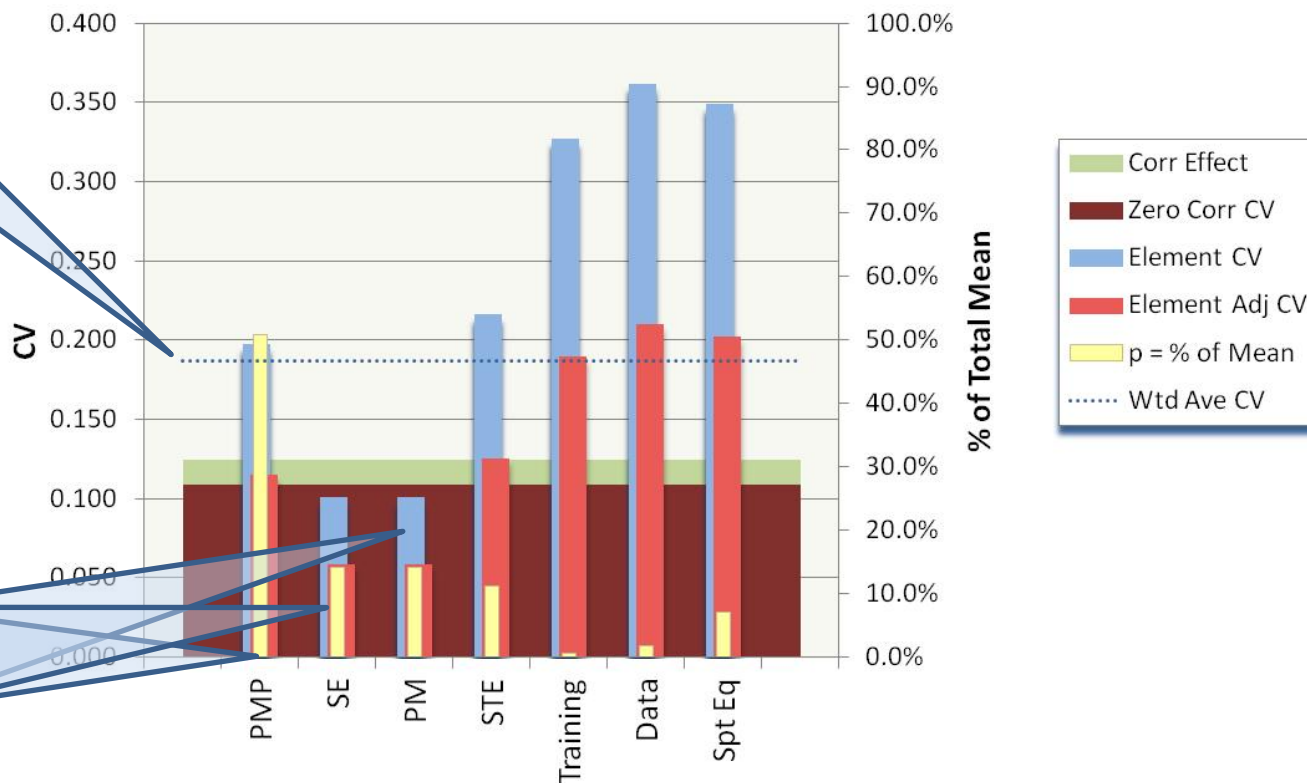
Diagnosing the CV: Partial CV

- If the range bounded by $CV_{\text{zero corr}}$ and $CV_{\text{w.ave}}$ does not fall within the expected range, the uncertainty distributions on the model inputs may need to be adjusted.
- The partial CV is a measure of how much a given WBS element contributes to the total CV.
 - If the CV is high, look for “big ticket” elements with a large pCV.
 - Influential elements that may be pulling the risk too high.
 - If the CV is low, look for elements with a large value of p , but a relatively small pCV.
 - Potentially influential elements that may not have enough uncertainty modeled.
 - Remember that uncertainty is generally applied to the inputs, not on the Level 2 WBS elements.
- These exercises should not replace a formal risk analysis when needed, but may give insight into where to focus efforts.



Example: MUGGLE Redux

Step 3 True CV, Including Correlation
MUGGLE Example Program



Correlation Limit:

This is as high as the CV can go by adjusting correlation alone.

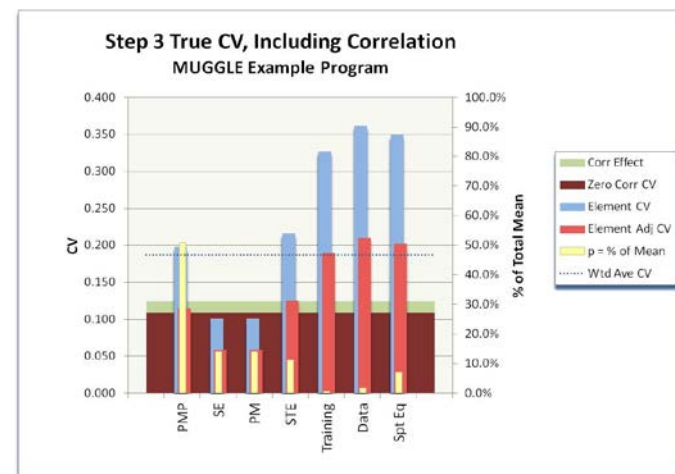
Influential Elements:

Because of the *f* adjustment, these are quite low. By directly increasing the risk on these, the CV could be increased.



Example: MUGGLE Redux

- The highest we could drive the CV by applying additional correlation to the child elements is ~ 0.187 .
 - In reality we can't go that high, because perfect correlation is absurd.
- The value of 0.187 is within the expected range for this program.
 - Complexity
 - Phase
 - Commodity
 - Risk
- Let's examine correlation!





Example: MUGGLE Redux

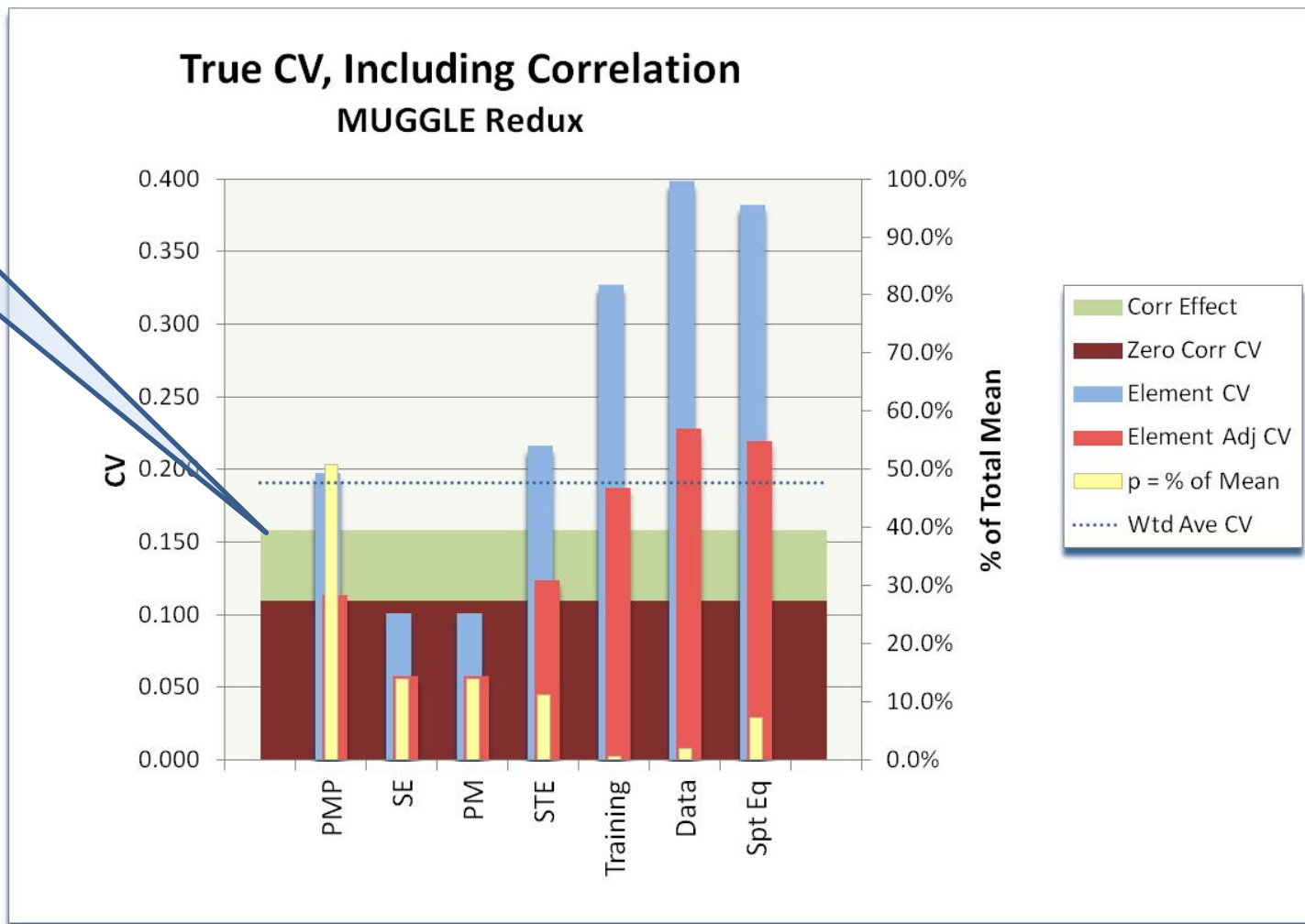
- Apply additional correlation to the WBS elements:

ORIGINAL CORRELATION MATRIX							
	PMP	SE	PM	STE	Training	Data	Spt Eq
PMP	1	0.0	0.0	0.0	0.0	0.5	0.6
SE		1	0.0	0.0	0.0	0.0	0.0
PM			1	0.0	0.0	0.0	0.0
STE				1	0.0	0.1	0.0
Training					1	0.0	0.0
Data						1	0.3
Spt Eq							1

AFTER APPLYING ADDITIONAL CORRELATION							
	PMP	SE	PM	STE	Training	Data	Spt Eq
PMP	1	0.7	0.5	0.5	0.3	0.7	0.7
SE		1	0.4	0.4	0.2	0.5	0.5
PM			1	0.3	0.1	0.3	0.3
STE				1	0.1	0.4	0.4
Training					1	0.2	0.2
Data						1	0.4
Spt Eq							1



Example: MUGGLE Redux



Conclusions

- The coefficient of variation (CV) is a useful metric for measuring the spread of a risk-based estimate.
- The weighted average CV, though not mathematically sound, is an intuitive starting point for a quick diagnostic.
- Define the partial CV (pCV) as the CV times the relative mean.
- The pCV representation leads to a logical path toward understanding the CV:
 1. Start with the weighted average CV.
 2. Multiply by f to get the zero correlation CV.
 3. Add correlation.
- The weighted average CV limits potential increase due to adding correlation to children.

Acknowledgements

- I would like to thank the following individuals for their immensely helpful feedback on this presentation:
 - Janet Vacca-Lebouef
 - Brian Harris
 - Stephanie Anderson
 - Chris Conrad
 - Andy Serfass
 - Ryan Shakely

Questions?