Robust Default Correlation for Cost Risk Analysis

Christian Smart, Ph.D., CCEA
Director, Cost Estimating and Analysis
Missile Defense Agency
christian.smart@mda.mil
256-450-4936

Abstract

Correlation is an important consideration in cost risk analysis. Exclusion of correlation from cost risk analysis results in the de facto assumption that all risks are independent. The assumption of independence leads to significant underestimation of total risk. However, figuring out the correct correlation values between work breakdown structures elements is challenging. For instance, it is difficult to estimate the exact correlation value between the structures and thermal protection subsystems in a cost risk estimate.

In order to circumvent these issues a default correlation value is often used. A commonly used value, attributed to Dr. Steve Book, is 20%. There is some empirical evidence to support this value. The basis for the 20% is discussed, and the supporting empirical evidence is presented. However, this evidence is limited, and the default value is sensitive to error in the assumption. For example, for a 100-element work breakdown structure, if the true correlation is 60%, the risk as measured by the total standard deviation of the estimate has been underestimated by 40%. A new approach to default correlation is presented that minimizes the expected value of the absolute error when the assumed correlation is not equal to the actual correlation. This approach is robust in the sense that the error is minimized. Depending upon the underlying assumption, this value varies significantly. The pros and cons of each assumption are discussed, and a new recommended default value is proposed. The derivation of the values is presented in detail.

Introduction

In 1996, Don Mackenzie wrote, "One of the more difficult chores in cost risk analysis is establishing appropriate levels of correlation across the numerous pairs of cost elements in the system cost model." (Mackenzie 1996). Seventeen years later, the cost analysis profession is still struggling with this issue. This paper is an attempt at making progress at addressing this issue.

Correlation in cost between two events is the tendency for those costs to move in tandem. It can be positive when there is a tendency for one Work Breakdown Structure (WBS) element's cost to increase when another WBS element's cost increases. It can also be negative, which means there is a tendency for one WBS element's cost to decrease whenever another WBS element's cost increases, and vice versa.

Correlation is still often ignored in cost risk analysis. However, WBS elements are not independent. which is the underlying assumption when correlation is ignored. The analyst who ignores correlation is implicitly assuming all WBS elements are independent, which is not the case.WBS elements can directly influence one another. For example, if the diameter of a missile increases, additional thermal coatings will necessarily be required. Thus an increase in structures cost can directly lead to an increase in thermal control cost. Also, there are underlying common cause factors for cost growth. A budget constraint that leads to an increase in schedule will affect all WBS elements equally.

Consider two random variables, ***X*** and ***Y***. The mean of ***X***, ***E(X)***, is denoted by μ_x, and similarly, the mean of ***Y***, ***E(Y)***, is denoted by μ_y. The variance of ***X***, ***Var(X)***, is denoted by σ_X^2, and similarly, the variance of ***Y***, ***Var(Y)***, is denoted by σ_Y^2.

The variance of ***X*** and ***Y*** are equal to:

$$Var(X) = Cov(X,X) = E(X^2) - [E(X)]^2$$

$$Var(Y) = Cov(Y,Y) = E(Y^2) - [E(Y)]^2$$

Correlation, denoted by the Greek letter ρ (“rho”), is defined by:

$$\rho_{XY} = Corr(X,Y) = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E(XY) - \mu_X\mu_Y}{\sigma_X\sigma_Y}$$

The mean, or expected value, determines the central tendency. Variance is a measure of uncertainty about the central tendency. Correlation has a significant impact on the variance. For *n* WBS elements, the mean and the variance of the total cost are defined by:

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu_i$$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=}^{n} \sigma_i^2 + 2\sum_{j=2}^{n}\sum_{i=1}^{j-1} \rho_{ij}\sigma_i\sigma_j$$

As the variance of total cost is a function of the variances of the individual WBS elements *and* the correlation between them, it is impossible to avoid making a choice about correlation. As in the song “Free Will” by the rock band Rush, “Even if you choose not to decide you still have made a choice.” So the estimator who ignores correlation is making a choice about correlation: the wrong choice, since assuming complete independence will lead to underestimation of total, aggregate risk. See Figure 1 for a chart of how much the total standard deviation will be underestimated when correlation is assumed to be zero. In Figure

1, *n* represents the number of WBS elements. For example, for a 100-element WBS, if the actual correlation is 20%, but it is assumed to be zero between all elements, then the total standard deviation will be underestimated by approximately 80%. Note that the amount of underestimation increases with the size of the WBS.

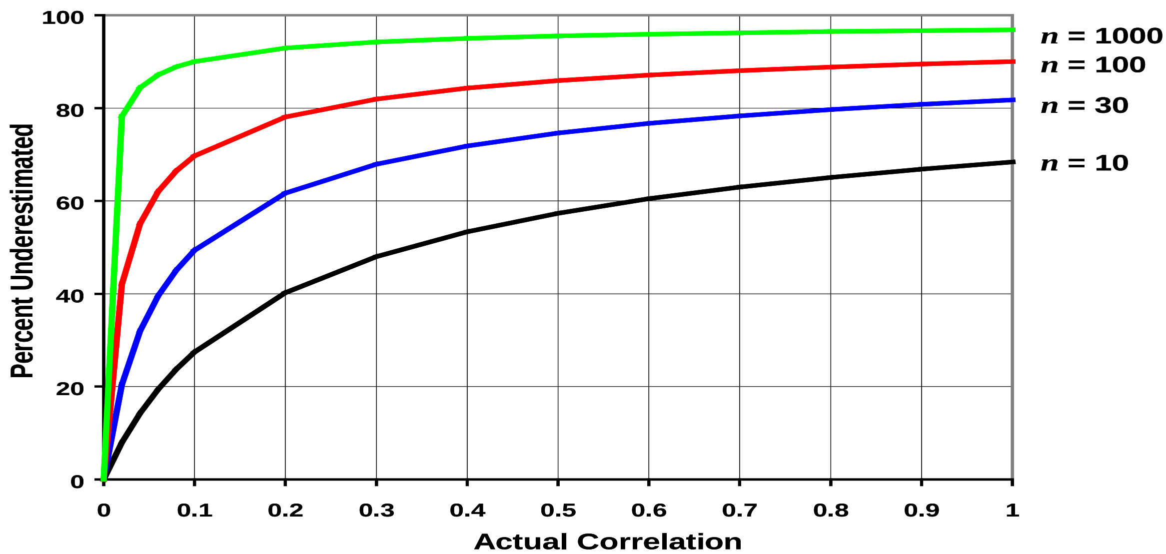

Figure 1. The Impact of Correlation on Risk (Book 1999).

Functional correlation is correlation that is implicit when one cost estimate is a function of another cost estimate. For example, system engineering cost is often modeled as a function of hardware cost. In such a case, when simulation is used to measure and aggregate risk, the variation in hardware cost naturally results in a functional correlation with system engineering cost. In this case, correlation is handled without assigning correlation values and no correlation between hardware cost elements and system engineering needs to be assigned. However, unless structures and thermal control are modeled as functions of a single underlying phenomenon, then correlation between WBS elements needs to be explicitly modeled.

Notice in the graph in Figure 1 there is an apparent knee in the curve around 20%. Above 20% correlation the consequence of assuming less correlation begins to dwindle. This graph is the basis for assuming 20-30% for default correlation for elements between which there is no functional correlation. Book (Book 1999) recommends 20% as a default correlation value because of this. However, the graph in Figure 1 does not tell us how much the total standard deviation is underestimated because correlation is assumed to be 20%, but is actually 60%, for example.

For example, for a 100-element WBS, if the correlation is assumed to be 20% but is actually 60%, the total standard deviation is underestimated by 40%. See Figure 2 (Book 1999).

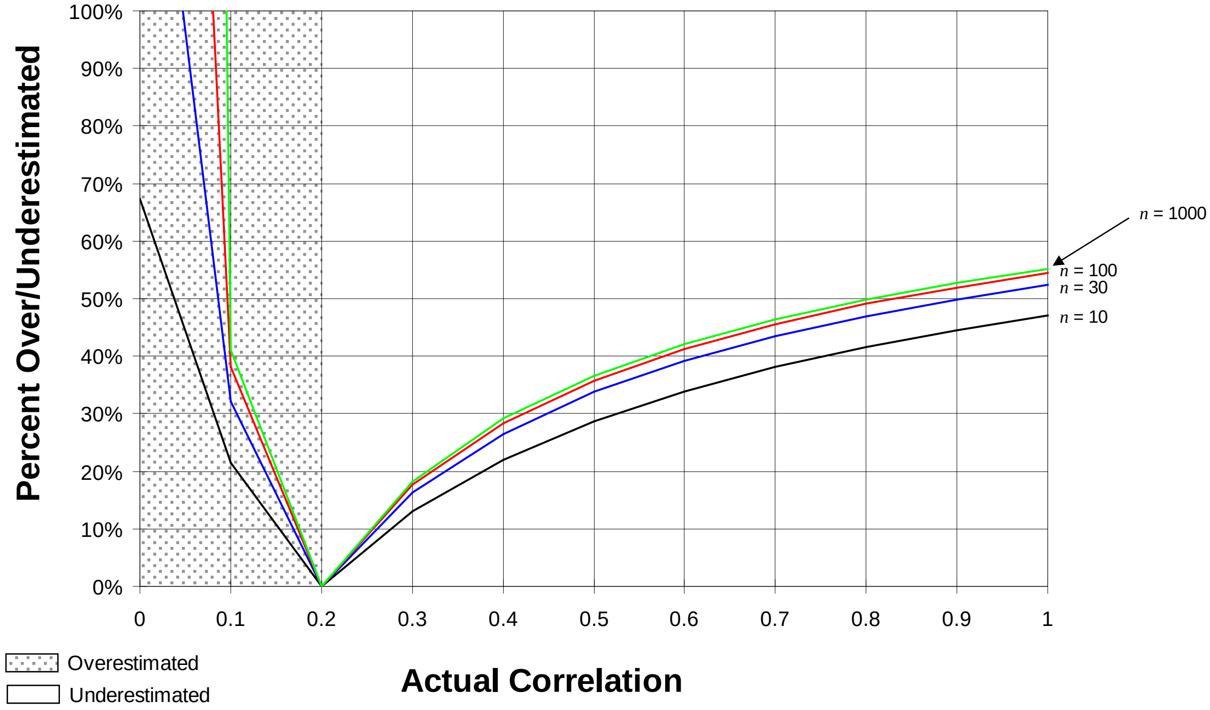

Figure 2. Percent Over/Underestimated When Correlation Assumed to be 0.2 Instead of ρ (Book, 1999).

The knee in the curve approach can still lead to significant underestimation of correlation. We turn to a different approach that seeks to minimize the amount of error in estimation, both over and under.

Robust Approach

A more robust approach to assigning correlations would be to use the value that results in the least amount of error. This approach is robust in the sense that without solid evidence to assign a correlation value, it minimizes the amount by which the total standard deviation is underestimated or overestimated due to the correlation assumption. This robust default measure of correlation would be a value for correlation that would minimize the error when the assumed correlation differs from the actual underlying correlation.

The percent error in total standard deviation between the assumed and actual correlation values, denoting the assumed correlation by ρ_1 and the actual correlation by ρ_2 is

$$\varepsilon = \frac{\sqrt{n}\sigma\sqrt{1+(n-1)\rho_2} - \sqrt{n}\sigma\sqrt{1+(n-1)\rho_1}}{\sqrt{n}\sigma\sqrt{1+(n-1)\rho_2}} = \frac{\sqrt{1+(n-1)\rho_2} - \sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}}$$

Note that ε is the percentage difference from the actual correlation to the assumed correlation. That is, this percentage measures how far the assumed value is from the actual correlation.

We assume that the level of default correlation ranges between 0 and 1. This is a reasonable assumption since on average the level of correlation between WBS elements should be a positive number, and correlation is always less than or equal to 1. We assume no other knowledge about the level of actual correlation; in other words, the distribution of the actual correlation values is uniform on the interval (0,1).

Considering ε to be a function of ρ_2, the actual correlation, the metric we are interested in measuring is the expected value of the absolute error, or $|\varepsilon|$, which can be expressed as

$$E(|\varepsilon|) = \int_0^1 |\varepsilon| f(\rho_2) d\rho_2 = \int_0^1 |\varepsilon| d\rho_2,$$

since $f(\rho_2) = 1$. This expression can be calculated as

$$\int_0^{\rho_1} \frac{\sqrt{1+(n-1)\rho_1} - \sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_2}} d\rho_2 + \int_{\rho_1}^1 \frac{\sqrt{1+(n-1)\rho_2} - \sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}} d\rho_2$$

Solving this integral, it is found that

$$\int_0^{\rho_1} \frac{\sqrt{1+(n-1)\rho_1} - \sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_2}} d\rho_2 + \int_{\rho_1}^1 \frac{\sqrt{1+(n-1)\rho_2} - \sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}} d\rho_2 =$$

$$\int_0^{\rho_1} \left(\frac{\sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}} - 1 \right) d\rho_2 + \int_{\rho_1}^1 \left(1 - \frac{\sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}} \right) d\rho_2$$

Noting that $\sqrt{1+(n-1)\rho_1}$ is a constant, and setting $c = \sqrt{1+(n-1)\rho_1}$, the integral simplifies to

$$\int_0^{\rho_1} \left(c(1+(n-1)\rho_2)^{-\frac{1}{2}} - 1 \right) d\rho_2 + \int_{\rho_1}^1 \left(1 - c(1+(n-1)\rho_2)^{-\frac{1}{2}} \right) d\rho_2 =$$

$$\left[\frac{2c}{n-1}(1+(n-1)\rho_2)^{\frac{1}{2}} - \rho_2 \right]_0^{\rho_1} + \left[\rho_2 - \frac{2c}{n-1}(1+(n-1)\rho_2)^{\frac{1}{2}} \right]_{\rho_1}^1 =$$

$$\frac{2c}{n-1}(1+(n-1)\rho_1)^{\frac{1}{2}} - \rho_1 - \frac{2c}{n-1} + 1 - \frac{2c}{n-1}(n)^{\frac{1}{2}} - \rho_1 + \frac{2c}{n-1}(1+(n-1)\rho_1)^{\frac{1}{2}} =$$

$$\frac{4c}{n-1}\left(1+(n-1)\rho_1\right)^{\frac{1}{2}} - 2\rho_1 - \frac{2c}{n-1}\left(1+\sqrt{n}\right)+1$$

Substituting for $c = \sqrt{1+(n-1)\rho_1}$ yields

$$\frac{4}{n-1}\left(1+(n-1)\rho_1\right) - 2\rho_1 - \frac{2}{n-1}\sqrt{1+(n-1)\rho_1}\left(1+\sqrt{n}\right)+1 =$$

$$2\rho_1 + \frac{4}{n-1} - \frac{2}{n-1}\sqrt{1+(n-1)\rho_1}\left(1+\sqrt{n}\right)+1$$

This final expression is a function of $\boldsymbol{n}$, the number of WBS elements, and ρ_1, the assumed correlation:

(Equation 1) $$f(\rho_1, n) = 2\rho_1 + \frac{4}{n-1} - \frac{2}{n-1}\sqrt{1+(n-1)\rho_1}\left(1+\sqrt{n}\right)+1$$

For $\boldsymbol{n = 100}$, the percentage differences with the column headings representing the assumed correlation, and the row headings representing the actual correlation, is shown in Table 1.

Actual Correlation	Assumed Correlation										
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
0%	0.0%	230.2%	356.1%	454.1%	537.2%	610.6%	677.2%	738.5%	795.5%	849.2%	900.0%
10%	69.7%	0.0%	38.1%	67.8%	93.0%	115.2%	135.4%	154.0%	171.3%	187.5%	202.9%
20%	78.1%	27.6%	0.0%	21.5%	39.7%	55.8%	70.4%	83.8%	96.4%	108.1%	119.3%
30%	82.0%	40.4%	17.7%	0.0%	15.0%	28.3%	40.3%	51.3%	61.6%	71.3%	80.5%
40%	84.3%	48.2%	28.4%	13.0%	0.0%	11.5%	22.0%	31.6%	40.5%	49.0%	56.9%
50%	85.9%	53.5%	35.8%	22.0%	10.3%	0.0%	9.4%	18.0%	26.0%	33.6%	40.7%
60%	87.1%	57.5%	41.3%	28.7%	18.0%	8.6%	0.0%	7.9%	15.2%	22.1%	28.7%
70%	88.1%	60.6%	45.6%	33.9%	24.0%	15.2%	7.3%	0.0%	6.8%	13.2%	19.3%
80%	88.8%	63.1%	49.1%	38.1%	28.8%	20.6%	13.2%	6.4%	0.0%	6.0%	11.7%
90%	89.5%	65.2%	52.0%	41.6%	32.9%	25.1%	18.1%	11.7%	5.7%	0.0%	5.4%
100%	90.0%	67.0%	54.4%	44.6%	36.3%	28.9%	22.3%	16.2%	10.4%	5.1%	0.0%

Table 1. Absolute Percentage Error from the Actual Correlation to the Assumed Correlation.

Figure 3 displays the assumed correlations and the average absolute percentage errors. A visual inspection shows that the minimum value is in the 10-20% range.

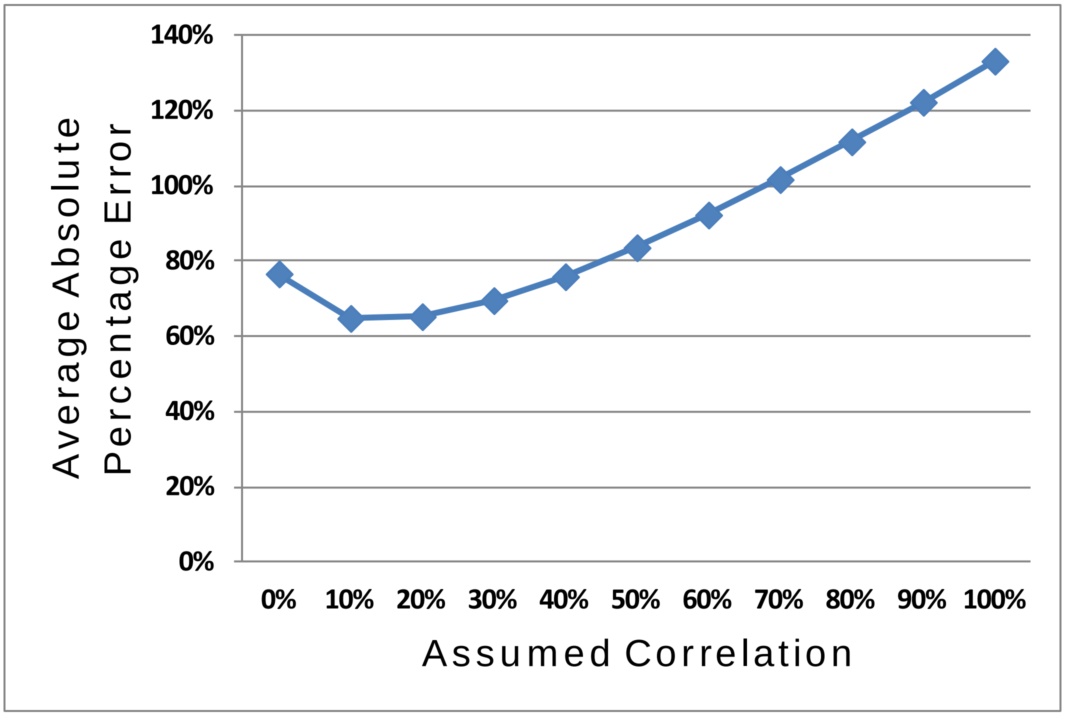

Figure 3. Total Error from Equation 1 for the case *n=100*.

To find a solution for the value of the correlation that minimizes the expected absolute error, take the partial derivative of function f in Equation 1 with respect to ρ_1:

$$\frac{\partial f(\rho_1, n)}{\partial \rho_1} = 2 - \left(1 + (n-1)\rho_1\right)^{-\frac{1}{2}}\left(1+\sqrt{n}\right)$$

Setting this equal to zero, and solving for ρ_1 yields:

$$\left(1 + (n-1)\rho_1\right)^{-\frac{1}{2}}\left(1+\sqrt{n}\right) = 2$$

$$\left(1 + (n-1)\rho_1\right)^{-\frac{1}{2}} = \frac{2}{\left(1+\sqrt{n}\right)}$$

$$\sqrt{1 + (n-1)\rho_1} = \frac{1+\sqrt{n}}{2}$$

$$1 + (n-1)\rho_1 = \frac{\left(1+\sqrt{n}\right)^2}{4}$$

$$(n-1)\rho_1 = \frac{\left(1+\sqrt{n}\right)^2}{4} - 1$$

(Equation 2) $$\rho_1 = \frac{\left(1+\sqrt{n}\right)^2 - 4}{4(n-1)}$$

The second derivative is

$$\frac{\partial^2 f(\rho_1, n)}{\partial \rho_1^2} = \frac{1}{2}(n-1)(1+(n-1)\rho_1)^{-\frac{3}{2}}(1+\sqrt{n})$$

which is positive as long as $n \geq 1$, confirming the minimum.

Table 2 shows values of ρ_1 from Equation 2 for various values of n:

n	ρ_1
10	37.0%
30	32.7%
100	29.5%
1,000	26.5%
10,000	25.5%

Table 2. Values of ρ_1 from Equation 2 for Various Values of *n*.

The limit of ρ_1 as $n \to \infty$ can be calculated using L'Hospital's Rule:

$$\lim_{n\to\infty} \frac{(1+\sqrt{n})^2 - 4}{4(n-1)} = \lim_{n\to\infty} \frac{1+2\sqrt{n}+n-4}{4n-4} = \lim_{n\to\infty} \frac{n+2\sqrt{n}-3}{4n-4} = \lim_{n\to\infty} \frac{1+\frac{1}{\sqrt{n}}}{4} = \frac{1}{4}$$

Thus, the single recommended value from this approach is 25%, close to the 20% recommended by Book (Book 1999) and widely used in cost analysis. However, the expected value of the absolute error is minimized by this value because of the large penalty assigned when overestimating actual correlations near zero. For example, let $n = 100$ and assume the correlation is 40%. The absolute percentage error when the actual correlation is equal to zero is

$$\frac{\sqrt{1+(n-1)\rho_1} - \sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_2}} = \frac{\sqrt{1+99 \cdot 0.40} - 1}{1} = 537\%$$

while the absolute percentage error when the actual correlation is equal to 80% is

$$\frac{\sqrt{1+(n-1)\rho_2} - \sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}} = \frac{\sqrt{1+99 \cdot 0.80} - \sqrt{1+99 \cdot 0.40}}{\sqrt{1+99 \cdot 0.80}} = 29\%.$$

This type of skewness is due to the actual correlation being in the denominator. For example if the assumed correlation is 40%, but the actual is 50%, we have underestimated the actual correlation by

$$1-\frac{.4}{.5}=20\%$$

However, if the assumed correlation is 40% but the actual correlation is 30%, then we have overestimated the actual correlation by

$$\frac{.4}{.3}-1=33\%$$

My opinion is that the penalty should not differ greatly whether you are overestimating or underestimating. The biggest penalty is for overestimating zero correlation, and we know that the correlation is not zero. If we assume 50% for the correlation value but the true value is 0, then we have overestimated the total variance by over 600%, while if the true correlation is 1 we have underestimate the total standard deviation by less than 30%.

A simple way to solve this issue is to change the denominator. If the denominator is the assumed correlation, then whether the actual is 10 percentage greater or 10 percentage points less than the assumed correlation of 40%, the percent error is equal to

$$\frac{.1}{.4}=25\%$$

Using this approach the percent difference is very similar to the previous approach, the only difference is the denominator:

$$\frac{\sqrt{n}\sigma\sqrt{1+(n-1)\rho_2}-\sqrt{n}\sigma\sqrt{1+(n-1)\rho_1}}{\sqrt{n}\sigma\sqrt{1+(n-1)\rho_1}}=\frac{\sqrt{1+(n-1)\rho_2}-\sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_1}}$$

where again ρ_1 is equal to the assumed correlation and ρ_2 is the actual correlation.

In this case, the expected value of the absolute error is given by

$$\int_0^{\rho_1}\frac{\sqrt{1+(n-1)\rho_1}-\sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_1}}d\rho_2+\int_{\rho_1}^{1}\frac{\sqrt{1+(n-1)\rho_2}-\sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_1}}d\rho_2=$$

$$\int_0^{\rho_1}\left(1-\frac{\sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_1}}\right)d\rho_2+\int_{\rho_1}^{1}\left(\frac{\sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_1}}-1\right)d\rho_2$$

Noting that $\sqrt{1+(n-1)\rho_1}$ is a constant, and setting $c=(1+(n-1)\rho_1)^{-\frac{1}{2}}$, the integral simplifies to

$$\int_0^{\rho_1}\left(1-c(1+(n-1)\rho_2)^{\frac{1}{2}}\right)d\rho_2+\int_{\rho_1}^{1}\left(c(1+(n-1)\rho_2)^{\frac{1}{2}}-1\right)d\rho_2=$$

$$\left[\rho_2 - \frac{2c}{3(n-1)}(1+(n-1)\rho_2)^{\frac{3}{2}}\right]_0^{\rho_1} + \left[\frac{2c}{3(n-1)}(1+(n-1)\rho_2)^{\frac{3}{2}} - \rho_2\right]_{\rho_1}^{1} =$$

$$\rho_1 - \frac{2c}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}} + \frac{2c}{3(n-1)} + \frac{2c}{3(n-1)}(n)^{\frac{3}{2}} - 1 - \frac{2c}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}} + \rho_1 =$$

$$2\rho_1 - \frac{4c}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}} + \frac{2c}{3(n-1)}\left(1+n^{\frac{3}{2}}\right) - 1$$

Substituting for $c = (1+(n-1)\rho_1)^{-\frac{1}{2}}$ yields

$$2\rho_1 - \frac{4}{3(n-1)}(1+(n-1)\rho_1) + \frac{2}{3(n-1)}\left(1+n^{\frac{3}{2}}\right)(1+(n-1)\rho_1)^{-\frac{1}{2}} - 1$$

This final expression is a function of ***n***, the number of WBS elements, and ρ_1, the assumed correlation:

(Equation 3) $$g(\rho_1, n) = 2\rho_1 - \frac{4}{3(n-1)}(1+(n-1)\rho_1) + \frac{2}{3(n-1)}\left(1+n^{\frac{3}{2}}\right)(1+(n-1)\rho_1)^{-\frac{1}{2}} - 1$$

For ***n* = *100***, the percentage differences with the column headings representing the assumed correlation, and the row headings representing the actual correlation, is shown in Table 3.

	Assumed Correlation										
Actual Correlation	**0%**	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
0%	0.0%	69.7%	78.1%	82.0%	84.3%	85.9%	87.1%	88.1%	88.8%	89.5%	90.0%
10%	230.2%	0.0%	27.6%	40.4%	48.2%	53.5%	57.5%	60.6%	63.1%	65.2%	67.0%
20%	356.1%	38.1%	0.0%	17.7%	28.4%	35.8%	41.3%	45.6%	49.1%	52.0%	54.4%
30%	454.1%	67.8%	21.5%	0.0%	13.0%	22.0%	28.7%	33.9%	38.1%	41.6%	44.6%
40%	537.2%	93.0%	39.7%	15.0%	0.0%	10.3%	18.0%	24.0%	28.8%	32.9%	36.3%
50%	610.6%	115.2%	55.8%	28.3%	11.5%	0.0%	8.6%	15.2%	20.6%	25.1%	28.9%
60%	677.2%	135.4%	70.4%	40.3%	22.0%	9.4%	0.0%	7.3%	13.2%	18.1%	22.3%
70%	738.5%	154.0%	83.8%	51.3%	31.6%	18.0%	7.9%	0.0%	6.4%	11.7%	16.2%
80%	795.5%	171.3%	96.4%	61.6%	40.5%	26.0%	15.2%	6.8%	0.0%	5.7%	10.4%
90%	849.2%	187.5%	108.1%	71.3%	49.0%	33.6%	22.1%	13.2%	6.0%	0.0%	5.1%
100%	900.0%	202.9%	119.3%	80.5%	56.9%	40.7%	28.7%	19.3%	11.7%	5.4%	0.0%

Table 3. Absolute Percentage Error from the Assumed Correlation to the Actual Correlation.

Figure 4 displays the assumed correlations and the absolute percentage errors. A visual inspection shows that the minimum value is at approximately 50-60%. The error associated with zero is off the chart, and is so high relative to the rest that it is not shown on the graph. Thus in this approach, the error associated with assuming zero correlation is very high.

To find a solution for the value of the correlation that minimizes the expected value of the absolute error, take the partial derivative of function ***g*** in Equation 3 with respect to ρ_1:

$$\frac{\partial g(\rho_1, n)}{\partial \rho_1} = \frac{2}{3} - \frac{1}{3}\left(1 + n^{\frac{3}{2}}\right)\left(1 + (n-1)\rho_1\right)^{-\frac{3}{2}}$$

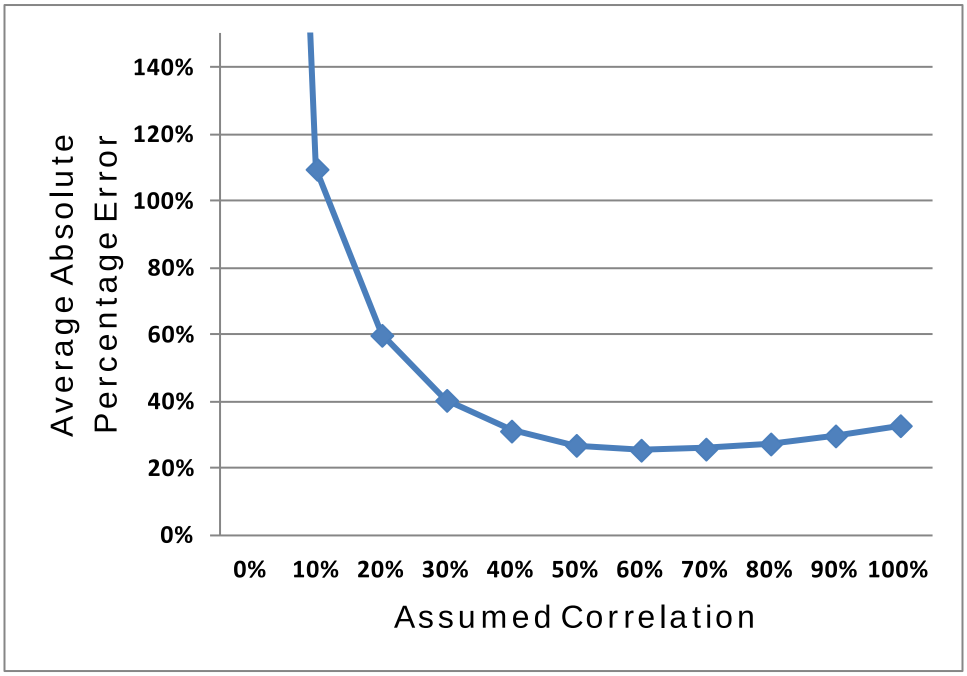

Figure 4. Error from Equation 3 for the case *n=100*.

Setting this equal to zero, and solving for ρ_1 yields:

$$\frac{1}{3}\left(1 + n^{\frac{3}{2}}\right)\left(1 + (n-1)\rho_1\right)^{-\frac{3}{2}} = \frac{2}{3}$$

$$\left(1 + (n-1)\rho_1\right)^{-\frac{3}{2}} = \frac{2}{\left(1 + n^{\frac{3}{2}}\right)}$$

$$\left(1 + (n-1)\rho_1\right)^{\frac{3}{2}} = \frac{1 + n^{\frac{3}{2}}}{2}$$

$$1 + (n-1)\rho_1 = \left(\frac{1 + n^{\frac{3}{2}}}{2}\right)^{\frac{2}{3}}$$

$$(n-1)\rho_1 = \left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{\frac{2}{3}} - 1$$

$$\rho_1 = \frac{\left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{\frac{2}{3}} - 1}{n-1}$$

The second derivative is

$$\frac{\partial^2 g(\rho_1, n)}{\partial \rho_1^2} = \frac{1}{2}(n-1)(1+(n-1)\rho_1)^{-\frac{5}{2}}\left(1+n^{\frac{3}{2}}\right)$$

which is positive as long as $n \geq 1$, confirming the minimum.

Table 4 displays values of ρ_1 from Equation 2 for various values of n:

n	ρ_1
10	60.4%
30	62.0%
100	62.7%
1,000	63.0%
10,000	63.0%

Table 4. Values of ρ_1 from Equation 2 for Various Values of n.

The limit of ρ_1 as $n \to \infty$ can be calculated using L'Hospital's Rule:

$$\lim_{n\to\infty} \frac{\left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{\frac{2}{3}} - 1}{(n-1)} = \lim_{n\to\infty} \frac{\frac{2}{3}\left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{-\frac{1}{3}} \frac{3}{4} n^{\frac{1}{2}}}{1} = \lim_{n\to\infty} \frac{\frac{1}{2} n^{\frac{1}{2}}}{\left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{\frac{1}{3}}} = \lim_{n\to\infty} \frac{\frac{1}{4} n^{-\frac{1}{2}}}{\frac{1}{3}\left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{-\frac{2}{3}} \frac{3}{4} n^{\frac{1}{2}}}$$

$$= \lim_{n\to\infty} \frac{\left(\frac{1+n^{\frac{3}{2}}}{2}\right)^{\frac{2}{3}}}{n} = \lim_{n\to\infty} \frac{\left(\frac{n^{\frac{3}{2}}}{2}\right)^{\frac{2}{3}}}{n} = \left(\frac{1}{2}\right)^{\frac{2}{3}} \approx 63\%$$

Thus, the single recommended value from this approach is 63%. This is much larger than the 25% value using the other approach, or the 20% rule of thumb widely used in practice.

See Table 5 for the percentage increase in total standard deviation that results from increasing the default 20% to 63% for all WBS elements. The increase is at least 50%, and can result in as much as a 77% increase.

n	% Increase in σ
10	54.3%
30	68.3%
100	74.5%
1,000	77.2%
10,000	77.5%

Table 5. The Percentage Increase in Total Standard Deviation Due to Changing Default Correlation from 20% to 63%, as a Function of the Number of WBS Elements.

The change from 20% to 63% has a significant impact on cost risk analysis. This is all due to a change in the denominator in the percentage difference that is considered. What about other metrics? Do they result in default values significantly larger than 20%? Consider for example, the absolute difference in total standard deviation, viz.,

$$\sqrt{n}\sigma\sqrt{1+(n-1)\rho_2} - \sqrt{n}\sigma\sqrt{1+(n-1)\rho_1}$$

Since σ is a constant, this is equivalent to minimizing

$$\sqrt{n}\sqrt{1+(n-1)\rho_2} - \sqrt{n}\sqrt{1+(n-1)\rho_1}$$

The expected value of the absolute difference is equal to

$$\int_0^{\rho_1}\left(\sqrt{n}\sqrt{1+(n-1)\rho_1} - \sqrt{n}\sqrt{1+(n-1)\rho_2}\right)d\rho_2 - \int_{\rho_1}^{1}\left(\sqrt{n}\sqrt{1+(n-1)\rho_2} - \sqrt{n}\sqrt{1+(n-1)\rho_1}\right)d\rho_2$$

$$= \sqrt{n}\left(\int_0^{\rho_1}\left(\sqrt{1+(n-1)\rho_1} - \sqrt{1+(n-1)\rho_2}\right)d\rho_2 - \int_{\rho_1}^{1}\left(\sqrt{1+(n-1)\rho_2} - \sqrt{1+(n-1)\rho_1}\right)d\rho_2\right)$$

Noting that $\sqrt{1+(n-1)\rho_1}$ is a constant, and setting $c=\sqrt{1+(n-1)\rho_1}$, the integral simplifies to

$$\sqrt{n}\left(\int_0^{\rho_1}\left(c-\sqrt{1+(n-1)\rho_1}\right)d\rho_2-\int_{\rho_1}^{1}\left(\sqrt{1+(n-1)\rho_2}-c\right)d\rho_2\right)$$

$$=\sqrt{n}\left(\left[c\rho_2-\frac{2}{3(n-1)}(1+(n-1)\rho_2)^{\frac{3}{2}}\right]_0^{\rho_1}+\left[\frac{2}{3(n-1)}(1+(n-1)\rho_2)^{\frac{3}{2}}-c\rho_2\right]_{\rho_1}^{1}\right)$$

$$=\sqrt{n}\left(c\rho_1-\frac{2}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}}+\frac{2}{3(n-1)}+\frac{2}{3(n-1)}n^{\frac{3}{2}}-c-\frac{2}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}}+c\rho_1\right)$$

$$=\sqrt{n}\left(2c\rho_1-\frac{4}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}}+\frac{2}{3(n-1)}\left(1+n^{\frac{3}{2}}\right)-c\right)$$

$$=\sqrt{n}\left((2\rho_1-1)(1+(n-1)\rho_1)^{\frac{1}{2}}-\frac{4}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}}+\frac{2}{3(n-1)}\left(1+n^{\frac{3}{2}}\right)\right)$$

This final expression is a function of n, the number of WBS elements, and ρ_1, the assumed correlation:

$$h(\rho_1,n)=\sqrt{n}\left((2\rho_1-1)(1+(n-1)\rho_1)^{\frac{1}{2}}-\frac{4}{3(n-1)}(1+(n-1)\rho_1)^{\frac{3}{2}}+\frac{2}{3(n-1)}\left(1+n^{\frac{3}{2}}\right)\right)$$

Differentiating with respect to ρ_1 yields

$$\frac{\partial h(\rho_1,n)}{\partial\rho_1}=\sqrt{n}\left(2(1+(n-1)\rho_1)^{\frac{1}{2}}+\frac{(2\rho_1-1)(n-1)}{2}(1+(n-1)\rho_1)^{-\frac{1}{2}}-2(1+(n-1)\rho_1)^{\frac{1}{2}}\right)$$

Simplifying,

$$\frac{\partial h(\rho_1,n)}{\partial\rho_1}=\sqrt{n}\left(\frac{(2\rho_1-1)(n-1)}{2}(1+(n-1)\rho_1)^{-\frac{1}{2}}\right)$$

Setting this derivative equal to zero and solving, we find

$$\sqrt{n}\left(\frac{(2\rho_1-1)(n-1)}{2}(1+(n-1)\rho_1)^{-\frac{1}{2}}\right)=0$$

Which implies that

$$2\rho_1 - 1 = 0$$

which means that

$$\rho_1 = \frac{1}{2}$$

Note that the second derivative with respect to ρ_1 is

$$\frac{\partial^2 h(\rho_1, n)}{\partial \rho_1^2} = \sqrt{n}\left((n-1)(1+(n-1)\rho_1)^{-\frac{1}{2}} - \frac{(n-1)^2(2\rho_1 - 1)}{4}(1+(n-1)\rho_1)^{-\frac{3}{2}} \right)$$

Substituting $\rho_1 = 0.5$, the second derivative is equal to

$$\sqrt{n}\left((n-1)(1+(n-1)0.5)^{-\frac{1}{2}} - 0 \right)$$

$$= \sqrt{n}\left((n-1)(1+(n-1)0.5)^{-\frac{1}{2}} \right)$$

$$= \sqrt{n}\left((n-1)(0.5+0.5n-0.5)^{-\frac{1}{2}} \right)$$

$$= \sqrt{n}\left((n-1)(0.5n+0.5)^{-\frac{1}{2}} \right)$$

$$= \sqrt{n}\left((n-1)\left(\frac{n+1}{2}\right)^{-\frac{1}{2}} \right)$$

which is greater than zero since $n \geq 1$. Thus 50% is a minimum value for the expression.

We can also re-look at the initial expression, the percentage difference from the actual correlation to the assumed correlation. If we eliminate the 0% case, which we know is incorrect, and the 100% case, which is also an incorrect assumption, then the average absolute percentage error is 40%. See Figure 5.

In looking at the expected absolute percentage error for this case, we calculate the same integral as in the first case, but change the limits

$$\int_{0.1}^{\rho_1} \frac{\sqrt{1+(n-1)\rho_1} - \sqrt{1+(n-1)\rho_2}}{\sqrt{1+(n-1)\rho_2}} d\rho_2 + \int_{\rho_1}^{0.9} \frac{\sqrt{1+(n-1)\rho_2} - \sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}} d\rho_2 =$$

$$\int_{0.1}^{\rho_1}\left(\frac{\sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}}-1\right)d\rho_2+\int_{\rho_1}^{0.9}\left(1-\frac{\sqrt{1+(n-1)\rho_1}}{\sqrt{1+(n-1)\rho_2}}\right)d\rho_2$$

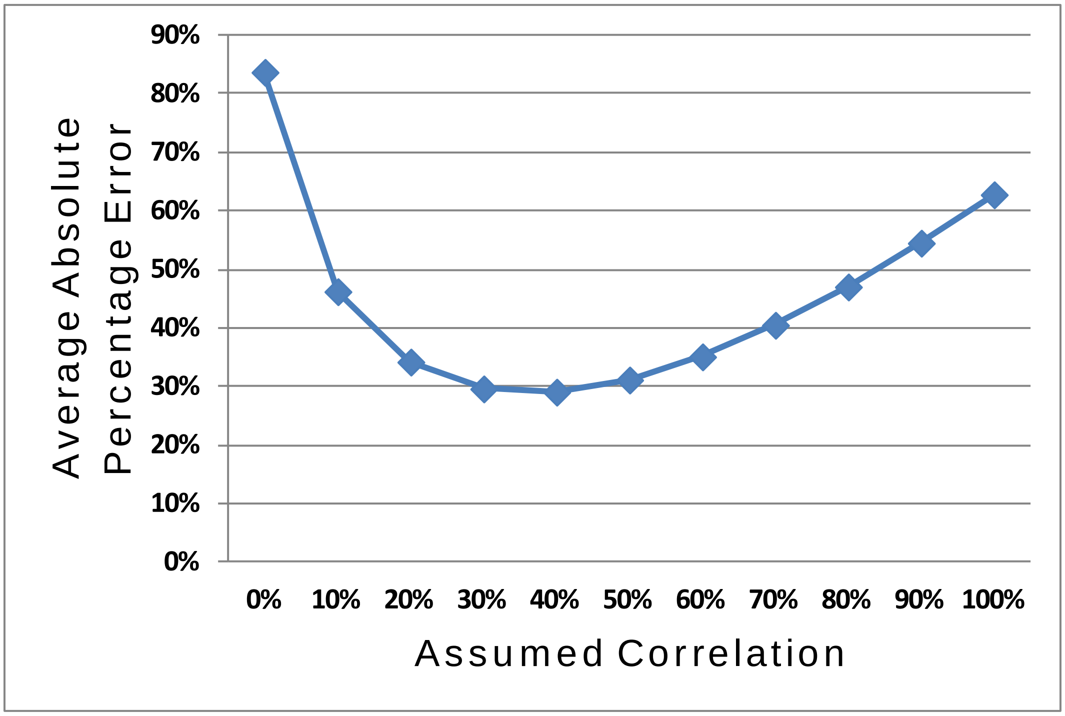

Figure 5. Average Absolute Percentage Error for the Percentage Error from the Actual Case, with 0% and 100% Correlation Not Considered.

Noting that $\sqrt{1+(n-1)\rho_1}$ is a constant, and setting $c=\sqrt{1+(n-1)\rho_1}$, the integral simplifies to

$$\int_{0.1}^{\rho_1}\left(c\left(1+(n-1)\rho_2\right)^{-\frac{1}{2}}-1\right)d\rho_2+\int_{\rho_1}^{0.9}\left(1-c\left(1+(n-1)\rho_2\right)^{-\frac{1}{2}}\right)d\rho_2=$$

$$\left[\frac{2c}{n-1}\left(1+(n-1)\rho_2\right)^{\frac{1}{2}}-\rho_2\right]_{0.1}^{\rho_1}+\left[\rho_2-\frac{2c}{n-1}\left(1+(n-1)\rho_2\right)^{\frac{1}{2}}\right]_{\rho_1}^{0.9}=$$

$$\left[\frac{2c}{n-1}\left(1+(n-1)\rho_2\right)^{\frac{1}{2}}-\rho_2\right]_{0.1}^{\rho_1}+\left[\rho_2-\frac{2c}{n-1}\left(1+(n-1)\rho_2\right)^{\frac{1}{2}}\right]_{\rho_1}^{0.9}=$$

$$\frac{2c^2}{n-1}-\rho_1-\frac{2c^2}{n-1}(0.1n+0.9)^{\frac{1}{2}}+0.1+0.9-\frac{2c}{n-1}(0.9n+0.1)^{\frac{1}{2}}-\rho_1+\frac{2c^2}{n-1}=$$

$$\frac{4c^2}{n-1}-2\rho_1+1-\frac{2c^2}{n-1}\left((0.1n+0.9)^{\frac{1}{2}}+(0.9n+0.1)^{\frac{1}{2}}\right)$$

As before this final expression is a function of $\boldsymbol{n}$, the number of WBS elements, and ρ_1, the assumed correlation:

(Equation 4) $$f_1(\rho_1, n) = \frac{4c^2}{n-1} - 2\rho_1 + 1 - \frac{2c^2}{n-1}\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+0.1)^{\frac{1}{2}}\right)$$

To find a solution for the value of the correlation that minimizes the expected value of the absolute error, take the partial derivative of function f_1 in Equation 4 with respect to ρ_1:

$$\frac{\partial f(\rho_1, n)}{\partial \rho_1} = 2 - (1+(n-1)\rho_1)^{-\frac{1}{2}}\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right)$$

Setting this equal to zero, and solving for ρ_1 yields

$$(1+(n-1)\rho_1)^{-\frac{1}{2}}\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right) = 2$$

$$(1+(n-1)\rho_1)^{-\frac{1}{2}} = \frac{2}{(0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}}$$

$$\sqrt{1+(n-1)\rho_1} = \frac{(0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}}{2}$$

$$1+(n-1)\rho_1 = \frac{\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right)^2}{4}$$

$$(n-1)\rho_1 = \frac{\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right)^2}{4} - 1$$

$$\rho_1 = \frac{\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right)^2 - 4}{4(n-1)}$$

The second derivative is

$$\frac{\partial^2 f(\rho_1, n)}{\partial \rho_1^2} = \frac{1}{2}(n-1)(1+(n-1)\rho_1)^{-\frac{3}{2}}\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right)$$

which is greater than zero for values of n greater than zero, confirming the minimum.

Note that the limit is $\lim_{n\to\infty} \frac{\left((0.1n+0.9)^{\frac{1}{2}} + (0.9n+.1)^{\frac{1}{2}}\right)^2 - 4}{4(n-1)} =$

$$\lim_{n\to\infty}\frac{\left(\left(0.1^{\frac{1}{2}}+0.9^{\frac{1}{2}}\right)n^{\frac{1}{2}}\right)^{2}}{4n}=$$

$$\lim_{n\to\infty}\frac{\left(0.1^{\frac{1}{2}}+0.9^{\frac{1}{2}}\right)n}{4n}=\frac{\left(0.1^{\frac{1}{2}}+0.9^{\frac{1}{2}}\right)}{4}=0.40$$

A value of 40% is much higher than the 25% value calculated without constraint.

We've calculated four different default values – 25%, 40%, 50%, and 63% – using the minimum error approach by making slightly different starting assumptions. Which one is the best? I think that a percentage difference approach makes sense. Knowing that the difference between the estimated total standard deviation and the actual total standard deviation is $100 million doesn't tell me much, since it could be large if the standard deviation is $100 million, or relatively small if the total standard deviation is $1 billion. Since it is important to understand the relative difference, minimizing the percentage difference is favored. This eliminates from consideration the 50% value from the absolute error approach. The 25% value is not realistic because of the heavy penalty for overestimating zero correlation. That leaves the 40% from the actual error percentage approach and the 63% from the assumed error percentage approach. The issue with looking at the error relative to the actual correlation is that we don't know the actual correlation. We only know the assumed correlation. The same is true for CER residuals. The Minimum Unbiased Percent Error (MUPE) and the Zero bias Minimum Percent Error (ZMPE) CER methods look at the percentage error from the estimate, not from the actual. We should use the same metric in looking at correlation. Thus, I recommend using a default value for correlation that is equal to 63%.

The value 63% is high, much higher than the commonly cited 20%. This value is recommended if you do not have other information that can be used. Speaking of which, is there any empirical evidence for correlation? There is some, for spacecraft bus costs.

Empirical Evidence on Correlation

There is some limited empirical evidence for spacecraft bus cost estimating relationships to assign correlation at the 20% level. Both the NASA/Air Force Cost Model (NAFCOM) and the Unmanned Spacecraft Cost Model (USCM) are parametric models for estimating spacecraft costs. The models include cost estimating relationships for spacecraft bus and payload subsystems. The subsystems include both hardware and systems WBS elements. The hardware elements include items such as structures, thermal control, attitude determination and control system (ADCS), electric power system (EPS), and reaction control system (RCS). The systems elements include integration, assembly and checkout (IACO); system test operations (STO); ground support equipment (GSE); systems engineering and integration (SEI); program management (PM); and launch operations and orbital support (LOOS).

When CERs are available, a correlation coefficient between each subsystem can be calculated from the residuals between the estimated and the actual costs. The standard deviation for each WBS element in a cost risk analysis is calculated as the standard deviation of the residuals. The correlation is thus calculated as the ratio of the covariance of the residuals to the product of the standard deviations.

The average correlation value for version 2004 of the NASA/Air Force Cost Model (NAFCOM) was 20% for nonrecurring development costs (see Table 6). These correlation values were calculated by correlating the residuals between the CERs, as discussed in the Aerospace Corporation's "Correlation Tutorial" (Covert and Anderson, 2005).

					Design and Development Correlation Matrix							
DD	ADCS	CCDH	EPS	Structures	Thermal	RCS	IACO	STO	GSE	SEI	PM	LOOS
ADCS	1	0.36	0.38	0	0.25	0.09	0.17	0.27	0.06	0.07	0.05	0.19
CCDH		1	0	0.15	0.05	0	-0.16	0.52	0.05	0.04	0.06	0.06
EPS			1	-0.04	-0.04	0.14	-0.17	0.54	0.06	0.06	0.07	0.19
Structures				1	0.32	0.24	-0.01	0.14	0.06	0.06	0.06	0.09
Thermal					1	0.11	0.16	-0.11	0	0.01	0.03	0.25
RCS						1	0.07	-0.03	0.09	0.09	0.11	0.22
IACO							1	-0.26	0.79	0.83	0.83	0.73
STO								1	0.1	0.1	0.1	0.24
GSE									1	1	0.69	1
SEI										1	1	1
PM											1	0.46
LOOS												1

Table 6. Correlations in NAFCOM v 2004 for Development Costs (previously unpublished, developed by Smart, 2004) – Average = 21.3%

The average correlation value for recurring theoretical first unit costs in version 2004 of NAFCOM was 16.8%. This is similar to the 16% value reported in version 7 of the Unmanned Spacecraft Cost Model in Table 3 which includes correlation for both nonrecurring ("NR") and theoretical first unit ("T1") costs. Don Mackenzie and Bonnie Addison, in a 2000 study, found subsystem-level correlations between average unit costs to range between 20% to 40% (Mackenzie and Addison, 2000).

				Flight Unit Correlation Matrix					
FU	ADCS	CCDH	EPS	Structures	Thermal	RCS	IACO	SEI	PM
ADCS	1	0.5	-0.12	-0.26	0.58	0.24	0.16	0.07	0.09
CCDH		1	0.14	0.01	0.13	-0.03	-0.04	-0.05	-0.04
EPS			1	0.08	-0.12	-0.09	0.16	0.12	0.14
Structures				1	0.21	-0.1	0.06	0.11	0.11
Thermal					1	0.1	0.22	0.17	0.17
RCS						1	0.27	0.32	0.33
IACO							1	0.71	0.71
SEI								1	1
PM									1

Table 7. Correlations in NAFCOM v 2004 for Theoretical First Unit Costs (previously unpublished, developed by Smart) – Average = 16.8%

	ADCSNR	AGENR	COMMNR	EPSNR	IATNR	PROGNR	STRCNR	THERNR	TT CNR	ADCST1	AKMT1	COMMT1	EPST1	IATT1	LOOST1	PROGT1	STRCT1	THERT1	TT CT1
ADCSNR	1.000	-0.067	-0.096	-0.035	0.035	0.012	0.413	0.605	0.121	-0.095	0.983	-0.122	0.099	0.564	0.139	0.089	-0.047	-0.057	0.092
AGENR		1.000	-0.028	0.525	-0.079	0.127	0.091	-0.230	-0.125	0.416	0.001	0.085	-0.043	-0.163	-0.189	0.033	0.146	0.151	0.232
COMMNR			1.000	0.888	0.884	0.966	0.762	0.281	0.850	-0.166	0.305	-0.176	0.157	0.368	0.884	-0.158	0.109	0.037	-0.004
EPSNR				1.000	0.265	0.604	0.409	0.003	0.337	0.237	0.011	-0.275	0.076	0.342	0.021	-0.049	0.465	0.123	0.035
IATNR					1.000	0.721	0.615	0.331	0.747	-0.037	0.391	-0.133	-0.028	0.501	0.265	-0.145	0.113	-0.014	-0.189
PROGNR						1.000	0.697	0.222	0.868	-0.065	0.145	-0.191	-0.044	0.444	0.329	-0.191	-0.000	-0.125	0.019
STRCNR							1.000	0.837	0.761	-0.001	0.117	-0.214	-0.113	0.418	0.173	-0.018	0.220	-0.103	0.069
THERNR								1.000	0.077	-0.200	0.662	-0.171	-0.053	0.514	0.102	-0.010	-0.063	-0.165	0.092
TT CNR									1.000	-0.149	0.475	-0.118	-0.071	0.519	0.294	-0.178	-0.111	-0.095	0.022
ADCST1										1.000	-0.100	0.614	0.421	-0.262	-0.354	0.543	0.676	-0.029	0.655
AKMT1											1.000	-0.006	0.292	0.855	0.286	0.176	-0.003	-0.027	0.052
COMMT1												1.000	0.266	-0.454	-0.088	0.777	0.729	0.126	0.391
EPST1													1.000	-0.150	-0.145	0.381	0.388	-0.007	0.520
IATT1														1.000	0.448	-0.144	-0.224	-0.014	-0.320
LOOST1															1.000	-0.336	-0.097	-0.074	-0.169
PROGT1																1.000	0.421	-0.039	0.481
STRCT1																	1.000	-0.175	0.285
THERT1																		1.000	-0.140
TT CT1																			1.000

Table 8. Subsystem bus correlation for the Unmanned Spacecraft Cost Model v 7, average = 16% (Covert and Anderson, 2005)

Conclusion

The 20% value for correlation is often used when there is no information to provide a value. However, we have shown that in the absence of other information that this value is too low. Using a minimum error approach, we calculated values that ranged from 40-63% for default correlation. I recommend using a default value for 63%. This is a high value. It is higher than the empirical evidence on spacecraft bus correlation, which is closer to 20%, or even less. However, this evidence is scant – it is for only one type of hardware (spacecraft), and for only three models for that type of hardware. More research needs to be done but in the mean time, using 63% for default correlation is a reasonable approach. If the 63% value is too high, the downside will be higher percentiles, and higher estimates. As a profession, we don't have a reputation for over estimating. There are no Government Accountability Office reports on how we over estimate projects, or are too pessimistic. The opposite is just the case – we are often criticized for underestimating. Perhaps one reason for that is that we have been under representing correlation in our cost risk estimates. Increasing the default value for cost risk correlation can help us avoid cases where the actual cost is nowhere near our estimated cumulative distribution functions, or

S-curves, like the example in Figure 6, for which the actual cost is more than double the 95th percentile of the cost risk estimate.

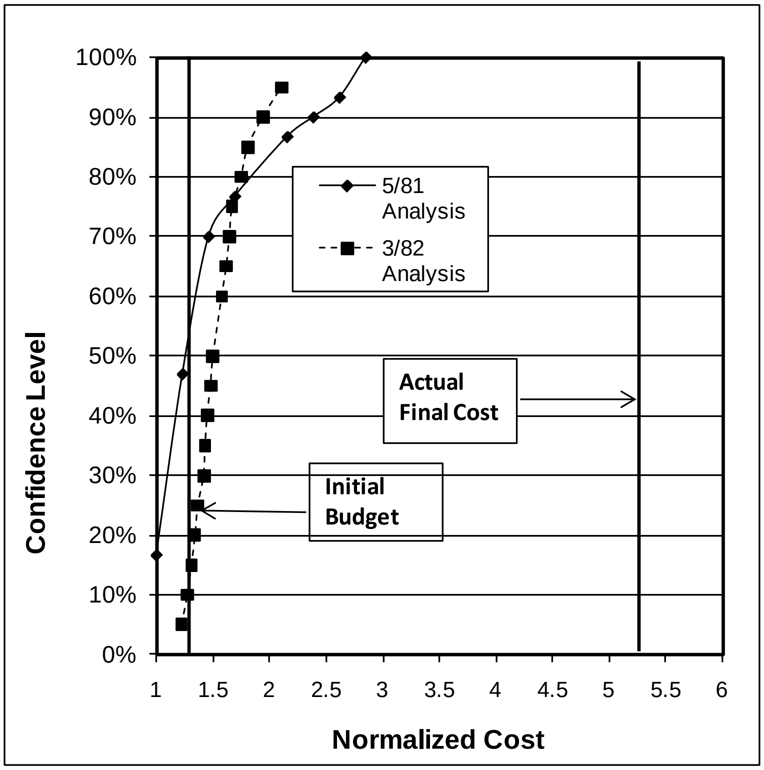

Figure 6. Comparison of S-curves and Final Actuals for the Tethered Satellite System (Source: Smart, 2011)

Afterword: Is Correlation Enough?

Correlation is only one measure of stochastic dependency. Correlation is simply a tendency for two elements to move together. One issue with correlation as a measure of stochastic dependency is the lack of tail dependency. For example, we expect a budget cut to impact all WBS elements, but this cannot be modeled with correlation. For a comparison of correlation and tail-dependency see Figures 7 and 8.

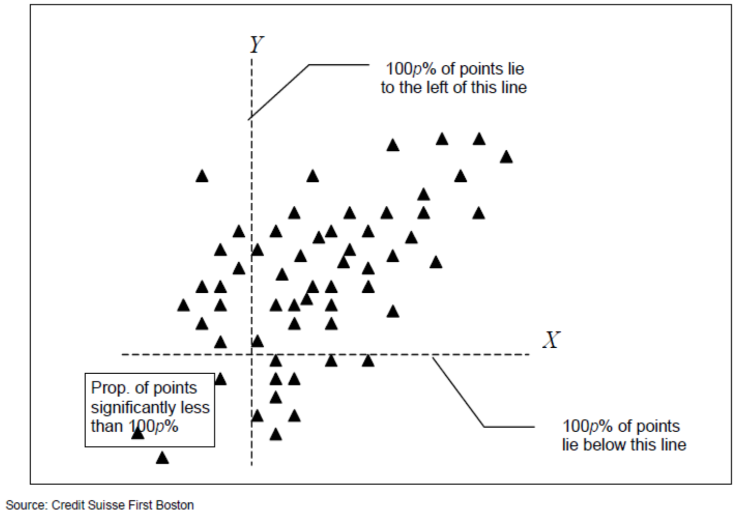

Figure 7. Correlated but Not Tail-Dependent (Source Gupton 1997).

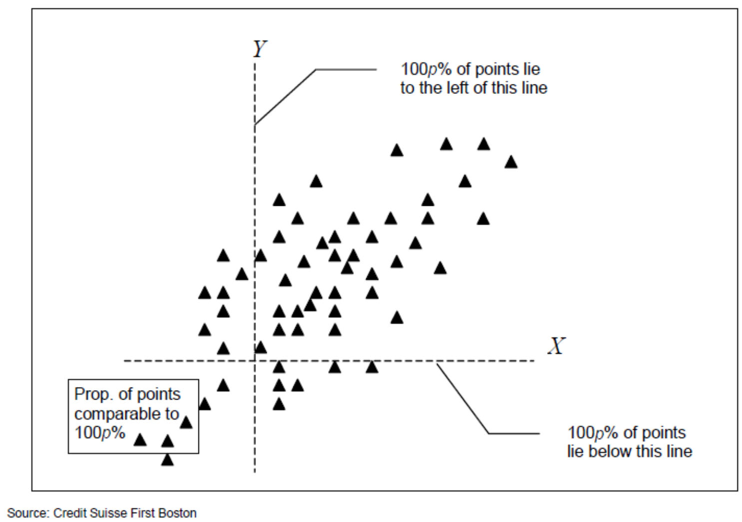

Figure 8. Tail Dependent (Source Gupton 1997).

Clearly, there is more to accurately modeling the real-world behavior of dependency among random variables. However, the commonly used multivariate distributions, such as the normal and the lognormal, do not account for tail dependency. However, there are ways to incorporate any type of dependency structure with any type of marginal distributions. These are called copulas, named after the Italian word for "couple." Copulas can solve many of these issues experienced with joint confidence level (JCL) modeling of cost and schedule. For more on copulas and their application to JCL, see Smart (Smart 2009).

References

1. Book, S.A., "Why Correlation Matters in Cost Estimating," Advanced Training Session, 32nd Annual DOD Cost Analysis Symposium, Williamsburg, VA, 1999.
2. Covert, R. and T. Anderson, "Correlation Tutorial, Rev. H," presented at the Cost Drivers Learning Event, November 2005.
3. Gupton, G.M., C.C. Finger, and M. Bahtia, *CreditMetrics – Technical Document*, 1997, J.P. Morgan, New York, available at: http://www.defaultrisk.com/_pdf6j4/creditmetrics_techdoc.pdf
4. Mackenzie, D., "Cost Variance in Idealized Systems," presented at the International Society of Parametric Analysts Annual Conference, Cannes, France, June, 1996.
5. Mackenzie, D., and B. Addison, "Space System Cost Variance and Estimating Uncertainty," presented at the Space Systems Cost Analysis Group meeting, Seattle, WA, October, 2000.
6. Smart, C., "Average Correlation Values for NAFCOM 2004," unpublished white paper, SAIC, 2004.
7. Smart, C., "Risk Analysis in the NASA/Air Force Cost Model," presented at the Joint Annual ISPA/SCEA Conference, Denver, CO, June, 2005.
8. Smart, C., "Mathematical Techniques for Joint Cost and Schedule Analysis," presented at the 2009 NASA Cost Symposium, Cocoa Beach, FL, May, 2009.
9. Smart, C., "Covered With Oil: Incorporating Realism in Cost Risk Analysis," presented at the 2011 Joint Annual International Society of Parametric Analysts and Society of Cost Estimate and Analysis Conference, Albuquerque, NM, June 2011.