The Portfolio Effect And The Free Lunch

Christian Smart, Ph.D.
Senior Parametric Cost Analyst
Science Applications International Corporation
christian.b.smart@saic.com
256-971-6583

Abstract

The portfolio effect is the reduction of risk achieved by funding multiple projects that are not perfectly correlated with one another. It is relied upon in setting confidence level policy for programs that consist of multiple projects. The idea of a portfolio effect has its roots in modern finance as pioneered by Nobel Memorial Prize winner Harry Markowitz. However, in two prior ISPA-SCEA conference presentations, “The Portfolio Reconsidered” in 2007 and “The Fractal Geometry of Cost Risk” in 2008, the author has demonstrated that the portfolio effect is more myth than fact. Additional cost growth data have been collected for an updated study. The number of data points for cost growth considered has increased from 40 in the previous study to 112. Data for schedule growth is also presented, and the distribution of schedule growth is discussed. Tail behavior for cost growth is discussed and the cost growth data are shown to closely follow a lognormal distribution with a high coefficient of variation. The portfolio effect for cost is still found to be minimal, at best. The concept of cost overrun insurance is introduced as one method for effectively implementing a true portfolio effect. The theoretical cost of this insurance, based on the equivalence principle, is found to be significant. Thus in order to achieve a portfolio effect, one must pay for it, in accordance with the famous principle that “there ain’t no such thing as a free lunch!”

Introduction

Consider a portfolio of stocks, such as those that might be held in a diversified stock account. Some of the stocks held may be may be spread across sectors such as energy, foods, and finance, to name just a few. These industries’ fortunes are not perfectly correlated with each another. The stock price for an oil company may increase at the same time as the stock price of a financial company is dropping. While some of these stocks may be quite risky, the combination of these stocks held in a common portfolio is less risky than the most volatile individual member because of this lack of perfect correlation. This effective reduction in risk by holding non-perfectly correlated assets is called diversification. The idea is a simple and appealing one, and was pioneered decades ago by economist Harry Markowitz, who was awarded a Nobel Memorial Prize for his research on the subject.

More recently, Tim Anderson (Ref. 1), among others, has applied the idea of a portfolio effect to the establishment of cost risk reserves for defense and NASA programs and projects. Individual projects within a larger program, such as the Ares I launch vehicle project that is part of the larger Constellation program, can be viewed as risky assets, not all of which are perfectly correlated with each other. Thus in order to achieve a specified confidence level for Constellation, it should be possible to fund each individual project, such as Ares, at a lower level. This strategy is currently being implemented, with a 65%

confidence level requirement for Constellation, but only a 50% confidence level requirement for Ares I. The portfolio effect is widely relied upon by NASA policy makers in reducing the level of funding among multiple projects that are part of a single program across the agency. Recent policy guidance, for example, relies upon achieving 70% program-level confidence by funding projects at the 50th percentile (with the exception of Constellation).

However, as pointed out by the author in two previous ISPA-SCEA presentations (Refs. 2 and 3), this much-relied upon effect is more apparent than real. Individual projects are riskier than we traditionally model them. The amount of cost risk for a project is often characterized by the variance or standard deviation. This will vary based on the size of the project. For example, the standard deviation that results from the cost risk analysis of an earth-orbiting satellite will likely be measured in millions or tens of millions of dollars, while that derived from a cost risk analysis for a launch vehicle development program will likely be measured in billions of dollars, a difference of an order of magnitude or more. A way to compare risk across projects when the absolute magnitude differs greatly is to examine the coefficient of variation, which is defined as the ratio of the standard deviation to the mean. If, for a launch vehicle cost risk analysis, the standard deviation is $2 billion and the mean is $10 billion, the coefficient of variation is 20%. Suppose that the mean and standard deviation derived from a cost risk analysis of an earth orbiting satellite is $100 million and $30 million, respectively. Then the coefficient of variation for the earth orbiting satellite cost risk is 30%. Even though the dollar amounts differ by a large amount, and the launch vehicle development has much more absolute risk, when comparing the relative amount of risk, the satellite development program has more variation about the mean. As discussed by the author in a previous paper (Ref. 2), the typical coefficient of variation as seen in typical cost risk analyses are not as high as would be expected from cost growth history. In addition, as discussed by the author in a follow-up paper (Ref. 3), traditional statistical methods were found to under-represent the risk of extreme cost growth events that have occurred in the past. Another key aspect of the portfolio effect is correlation. Correlation is necessary for achieving the reduction in risk promised by diversification. However, as everyone with a stock account or 401K that is invested in mutual funds is all too painfully aware, correlations are not static. A portfolio of stocks does not look very risky when the overall stock market is rising because some of the correlations are negative. But when a financial crises occurs like the one currently enveloping markets around the world, all the stocks in a portfolio fall together. Just as the correlation between stock investments can dramatically change from a negative value to perfect positive correlation in a market free-fall, the correlations between individual projects cannot be relied upon to remain a constant value over the development life-cycle. For example, history indicates (Ref. 4) that the correlation between launch vehicle stages is highest for a liquid rocket engine and its associated liquid rocket stage, with less correlation among other elements. However, if budgets are constrained for all the launch vehicle elements, cost will likely increase for all elements across the board. Hence the correlation will be much higher than expected. Treating correlation as a static variable based upon historical average experience will also lend support to a portfolio effect that does not exist. Thus it is important that correlation be treated as dynamic by updating it on a regular basis.

For the current study, additional cost growth data have been collected, including information for more recent missions, more than doubling the number of data points from 40 to 112. Also, schedule growth data for 48 missions have been collected. The results of the updated analysis lead to slightly different outcomes – cost risk is shown to closely follow a lognormal distribution, as opposed to a power law based on a smaller 40-mission cost growth database in a previous paper on this subject (Ref. 3). But the bottom line is found to be the same – the portfolio effect remains a chimera of our limited imagination. However, the effect can be implemented if NASA is willing to pay for it. The concept of cost-overrun insurance as a means to help alleviate issues with extreme cost growth is introduced as one means to implement a true portfolio effect.

Cost Growth and Cost Risk

Cost growth is the amount by which a program exceeds its initial budget. It is typically expressed as a percentage. For example, a program initially expected to cost $100 million at the beginning of the program, but which actually costs $150 million by the end of the program's development, is said to have experienced 50% cost growth.

Cost growth has been shown to be an endemic and universal phenomenon for space program development efforts. Studies by Shaffer, the U.S. Government Accountability Office, and Smart (Refs. 5-8) have shown that over three-quarters of all NASA programs experience cost growth, with average cost growth ranging to 35% and higher. These studies have also shown that many programs experience growth in excess of 100%!

For the current study, additional data were collected from several sources (Refs. 9-13). These studies include work done by Claude Freaner and Brian Rutkowski of NASA HQ to update and expand Matt Shaffer's initial 40-mission study. Much of these recent data points are focused on NASA's Science Mission Directorate, so additional data sources were selected to include a broader swath of missions, including launch vehicle stages. The expanded database includes 112 data points.

For these 112 missions, the minimum cost growth was -25.2% for Super Light Weight Tank (SLWT), an upgrade for the Shuttle Program from a more traditional aluminum structure to aluminum-lithium. The negative number means that costs under ran their initial budget by approximately 25%. Contrary to popular belief, missions occasionally come in under budget. For the current study, 14 such missions experienced under runs, which is12.5% of the missions studied. Only two of the missions hit their budget target spot on. Nine of the missions were within 5% of the initial budget, and 19 within 10% (either above or below).

The maximum cost growth among the missions studied was 385% for the Hubble Space Telescope and Space Telescope Assembly, which suffered from several sources of traditional cost growth, including funding constraints, launch vehicle delays, and under-estimation of the time and resources necessary to develop the requisite technology.

A range from -25% on the low side to over 350% on the high end is a wide range. The average cost growth for all missions was 53.0%, with median growth equal to 32.1%. The difference between the mean and median indicates a high degree of positive skew in the data, with most missions experiencing relatively small amounts of cost growth (half experienced growth less than 33%), with some missions experiencing extreme amounts of cost growth, such as Hubble and others. Overall, seventeen missions had cost growth in excess of 100%, which means cost more than doubled. While representing only about 15% of the cost growth data we will see that growth of this severity, while not common, occurs often enough to offset any hoped-for portfolio effect. See Figure 1 for a graphical summary of these data.

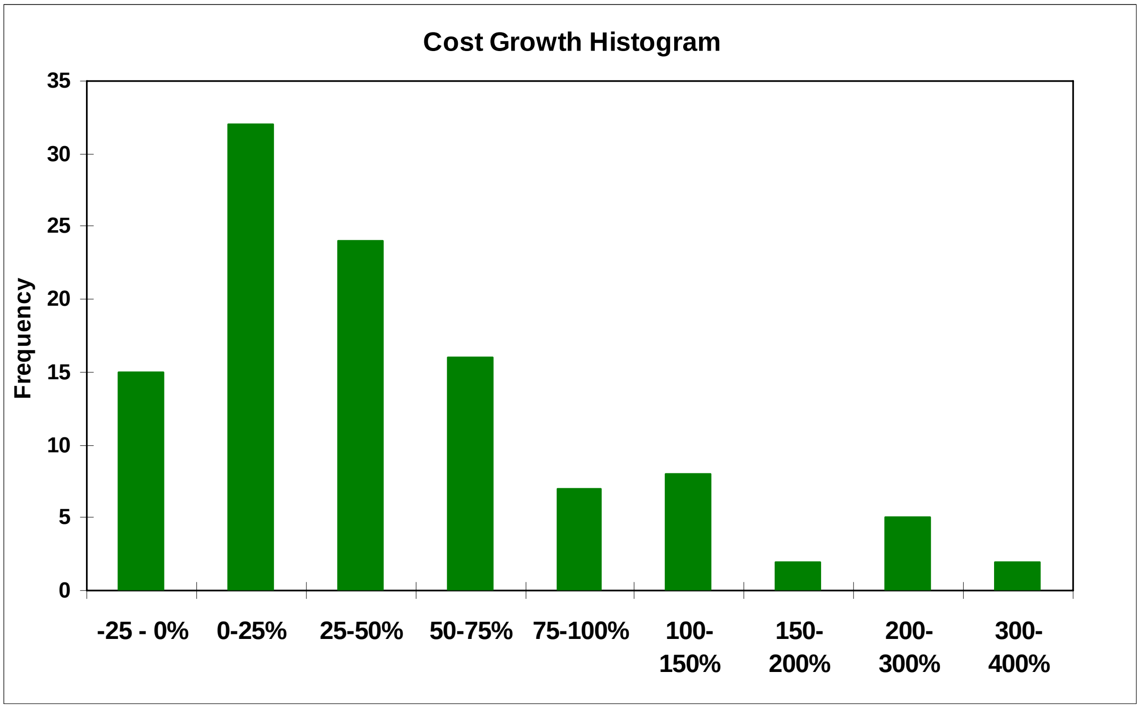

Figure 1. Graphical Summary of NASA Cost Growth.

Note that the data in Figure 1 has positive skew with a heavy right tail. The normal distribution, which is symmetric, and has thin tails, obviously does not fit these data, as is demonstrated in detail by author in a previous paper (Ref. 3). One of the most common distributions used for analyzing data that are, like cost, bounded below by zero and which have positive skew is the lognormal distribution. It was demonstrated previously by the author (Ref. 3) that the lognormal distribution tended to under-represent the right tail of the cost growth data. With additional missions added to the database, we re-examine which distributions best represent cost growth, including tail behavior. In order to examine the lognormal fit, the cost growth data are shifted so that the minimum growth amount, which is a negative number, becomes positive. We look at the ratio of final cost to actual cost, which ranges from .748 (representing the 25.5% under-run) to 485.1 (representing the 385.1% overrun), and subtract 0.7. Thus the data are transformed via

$$\frac{\textit{Final Cost}}{\textit{Initial Cost}} - 0.7 \ .$$

To fit the distributions, the Anderson-Darling, chi-square, and Kolmogorov-Smirnov statistics were calculated for each distribution. The statistics for the top four, as ranked by the Anderson-Darling statistic are displayed in Table 1.

Distribution	**Anderson-Darling**	**Chi-Square**	**Kolmogorov-Smirnov**
Lognormal	**0.5407**	**5**	**0.0656**
Gamma	**1.2712**	**16.7857**	**0.0906**
Max. Extreme Value	**2.1452**	**13.7857**	**0.1085**
Weibull	**2.4850**	**25.1429**	**0.1184**

Table 1. Comparison of Best-Fitting Distributions for Cost-Growth Data.

Each of these tests can be thought of as a measure of deviation away from a perfect fit for the data. Thus for all three, a smaller test-statistic value indicates a better fit. These three tests focus on slightly different aspects of a distribution's fit. Anderson-Darling is focused on the fit at the tails of the distribution, Kolmogorov-Smirnov measures the maximum difference between the data and the fitted distribution, and chi-square is a sum of squares of deviation measure.

Note that the lognormal distribution is the best-fitting distribution according to all three tests. Even though the lognormal has the best rank according to each test, that does not mean we should unequivocally accept the lognormal distribution as a good representative of the underlying data. When it comes to statistics, we can never positively prove a hypothesis such as "the cost growth data fit a lognormal distribution." We can however dis-prove hypotheses with data. Thus the best we can hope to do in distribution fitting is to fail to reject a given hypothesis. For each test, a critical value is determined based on the degrees of freedom of the data. The Anderson-Darling critical value is unique for each distribution. For example, the critical value for a lognormal at the 5% significance level is 0.752, and the lognormal test-statistic is below that amount by a comfortable margin at 0.5407. The Weibull test-statistic on the other hand is 2.4850, well above the 0.757 critical value at the 5% significance level.

The Kolmogorov-Smirnov critical value at the 5% significance level is $\frac{1.36}{\sqrt{N}} = \frac{1.36}{\sqrt{112}} \approx 0.129$. The critical value at the 20% level is approximately ***0.101***. Thus we fail to reject the lognormal and gamma hypotheses at either of these levels.

For the chi-square test, the critical value given the number of degrees of freedom is much higher than the chi-square test statistics shown in the table up to a 10% significance level, so we fail to reject any of the top four distributions by the chi-square criteria.

Thus we fail to reject the lognormal for all three goodness-of-fit tests. The fit according to the chi-square criterion is particularly good. Figure 2 shows a graphical comparison of the empirical cost growth distribution and the lognormal fit.

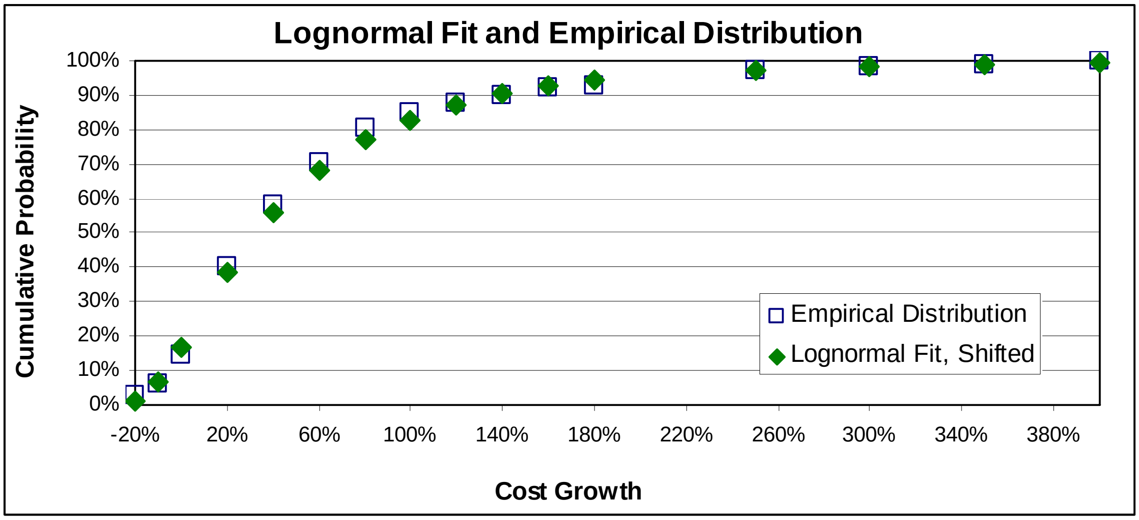

Figure 2. Graphical Comparison of Empirical Cost Growth Data and Fitted Lognormal Distribution.

In a 2008 presentation to ISPA-SCEA (Ref. 3), empirical evidence was provided that cost growth follows a power law for a 40-mission database. Mathematician Benoit Mandlebrot (Refs. 14,15) has shown that numerous natural and man-made phenomena follow these types of laws, which says that as size or magnitude increases, the impact decreases according to a power equation, which has the form

$$Y = aX^{b}.$$

These types of laws indicate that events of large magnitude are much more likely than you would expect if you modeled the phenomenon using a normal or lognormal distribution. How well does a power law fit the expanded data set? See Figure 3 for a

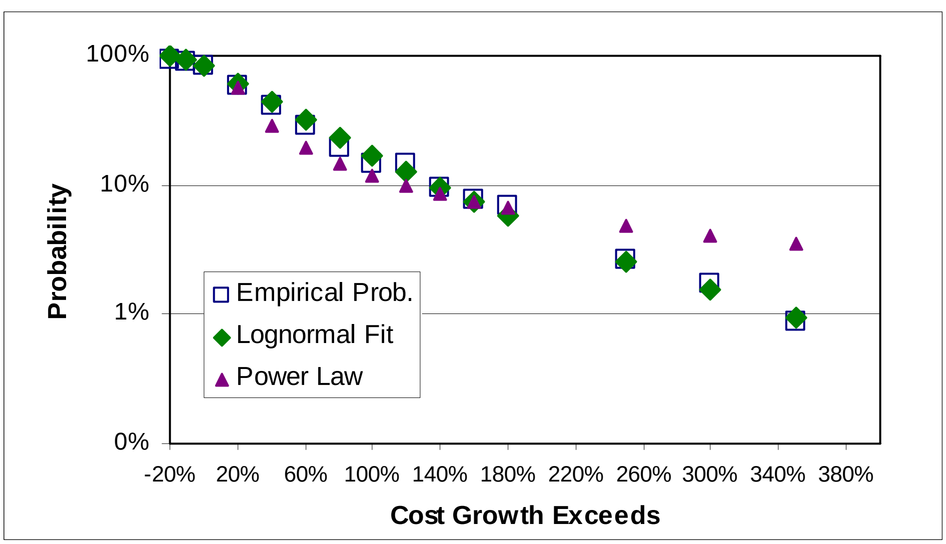

Figure 3. Graphical Comparison of the Empirical Cost Growth Data, the Lognormal Fit, and the Best-Fit Power Law.

graphical comparison. Note that in Figure 3 the y-axis is the probability that cost growth exceeds the amount specified on the x-axis, so it is basically the cumulative function turned upside down. It is seen from this graph that the expanded database changes the conclusions reached in the previous study. Cost growth is found to closely follow a lognormal distribution, including the right tail, while the power law fit to the data overestimates the likelihood of extreme cost growth.

Cost risk is the probability that an estimate will exceed a specified amount, such as $100 million or $150 million. Cost growth and cost risk are thus intrinsically related. Historical cost growth provides an excellent means for determining the overall level of risk for cost estimates. For example, if 95% of past programs have experienced less than 100% growth, we should expect that the ratio of actual cost to the initial estimate should be less than 100% with 95% confidence. Thus cost growth is the impact of cost risk in action. Because of uncertainty in historical data, cost models, program parameters, etc., the term "cost risk" is redundant. We will see how these data can be applied to cost risk when we look at the portfolio effect.

The Portfolio Effect and the Free Lunch

One facet of cost risk is that budgets are not set in isolation. Rather, budgets are set in the context of multiple ongoing missions. Thus a portfolio of missions is often considered in practice. In this case, it has been suggested that due to diversification across a suite of missions it is possible to achieve a high level of confidence in the overall budget while setting budgets for individual missions at a lower level (Ref. 1). This draws on ideas in economics, such as modern portfolio theory as expounded by Nobel laureate Harry Markowitz. See Table 2 for an example. In Table 2 there are ten mutually independent normal distributions. In this particular case it is possible to achieve 80% confidence for the full portfolio of ten missions while budgeting each individual mission at the 61% confidence level. This is a significant savings. However, it was shown in the author's

Project	μ	σ	61% Confidence Level
Project 1	$1,696	$539	$1,846
Project 2	$1,481	$404	$1,594
Project 3	$1,395	$435	$1,516
Project 4	$874	$288	$954
Project 5	$840	$219	$901
Project 6	$1,449	$371	$1,552
Project 7	$1,638	$537	$1,788
Project 8	$1,031	$259	$1,103
Project 9	$1,271	$323	$1,361
Project 10	$1,937	$602	$2,105
Total	$13,612	$1,317	$14,720

80th Percentile At Overall Level (annotation pointing to Total $14,720)

Table 2. Example of the Portfolio Effect for 10 Mutually Indendepent Normal Distributions.

previous studies on this subject that when cost growth follows a power law, the portfolio effect savings are so small they are almost nonexistent (Refs. 2, 3). For the updated database of 112 missions, cost growth, and hence cost risk, are best represented by a lognormal distribution. Since the emphasis in studying the portfolio effect is on the confidence level, we use the fact that for the cost growth data, the mean is at the 66th percentile. For a lognormal distribution this implies that

$$e^{\mu+0.5\sigma^2} = e^{\mu+0.4125\sigma},$$

and hence $\sigma = 0.825$. Since for a lognormal distribution, $\sigma = \sqrt{ln\left(1+CV^2\right)}$, this implies that ***CV= 0.99***, where ***CV*** is the coefficient of variation (ratio of the standard deviation to the mean). For each project in Table 1, the standard deviations based on empirical data can be derived by multiplying each mean by the coefficient of variation. Summing these 10 projects with the same means as in Table 2, assuming lognormal distributions, with coefficients of variation derived from empirical data, was performed using Monte Carlo simulation. The results are displayed in Table 3.

Project	μ	σ	74.7% Confidence Level
Project 1	$1,696	1679.04	$2,086
Project 2	$1,481	1466.19	$1,821
Project 3	$1,395	1381.05	$1,716
Project 4	$874	865.26	$1,075
Project 5	$840	831.6	$1,033
Project 6	$1,449	1434.51	$1,782
Project 7	$1,638	1621.62	$2,014
Project 8	$1,031	1020.69	$1,268
Project 9	$1,271	1258.29	$1,563
Project 10	$1,937	1917.63	$2,382
Total	$13,612		$16,740

80th Percentile At Overall Level

Table 3. The Portfolio Effect Based on Empirical Data.

Note that the portfolio effect based on empirical data offers much less savings. The effect in Table 3 also assumes independence, while in practice most projects are correlated. When 20% correlation is assumed between each of the 10 projects, each project must be funded at the 77% confidence level in order to achieve 80% confidence at the portfolio level. Thus, again, the portfolio effect is found to be a myth. Continued over reliance by policy makers on this nonexistent safety net will only mean continued issues with cost growth in excess of budgets and reserves.

The lack of a portfolio effect should not be surprising. The portfolio effect relies upon diversification, which, as stock pundit Jim Cramer has touted, is the “only free lunch on Wall Street” (Ref. 16). But that statement was made in 2007, before the recent stock market meltdown - no matter how diversified an individual’s stock portfolio most probably lost money, perhaps a significant amount, in the recent collapse. As several people have noted, including science fiction author Robert Heinlein and Nobel Prize-winning economist Milton Friedman "there ain't no such thing as a free lunch" (Refs. 17,18). Or, as the 19th century Hungarian mathematician Janos Bolyai wrote "One must do no violence to nature, nor model it in conformity to any blindly formed chimera" (Refs. 19,20). The lack of a portfolio effect is in keeping with this fundamental common-sense notion. And "investing" in a NASA project is different than buying a stock. When an investor buys a stock, he or she can at most lose the initial investment. When NASA invests on a project, however, the initial investment, or budget, can be exceeded many times over. Rather than being a stock investor, NASA and other agencies are making highly-leveraged speculative bets on projects with multi-year timeframes. This is akin to some of the highly leveraged risk taking that helped to create the current Wall Street crisis.

However, if cost growth could be capped at a specified percentage, such as 100%, then the portfolio effect does offer real savings. If each project is initially funded at the 50th percentile, in keeping with current NASA policy, and cost growth is limited to 100% growth in excess of the median, then achieving 80% confidence for the portfolio of missions only requires funding each mission at the 65.5th percentile, assuming lognormal distributions and 20% correlation between all projects. Table 4 displays the results.

Project	μ	σ	65.5% Confidence Level
Project 1	$1,696	1679.04	$1,676
Project 2	$1,481	1466.19	$1,463
Project 3	$1,395	1381.05	$1,378
Project 4	$874	865.26	$864
Project 5	$840	831.6	$830
Project 6	$1,449	1434.51	$1,432
Project 7	$1,638	1621.62	$1,619
Project 8	$1,031	1020.69	$1,019
Project 9	$1,271	1258.29	$1,256
Project 10	$1,937	1917.63	$1,914
Total	$13,612		$13,451

80th Percentile At Overall Level

Table 4. the Portfolio Effect with Cost Growth Caps.

Note that capping cost growth also significantly reduces the overall 80th percentile, to the point where it is less than the overall mean. Of course, capping cost growth to 100% may be hard to achieve in practice. Even without optimistic initial budgets, there may be external factors that can add to cost growth, such as funding instability and labor strikes.

One possible way to implement a cost growth cap would be to cancel any mission as soon as cost growth exceeds 100%. This would provide managers with the incentive to be realistic in setting initial budgets. However, such a policy is draconian and would in many cases punish missions and project managers for externally-caused cost growth. It would also mean a loss of science. For example had this policy been in place at the time, we wouldn't have the scientific insights produced by the Hubble Space Telescope or Galileo Orbiter.

Another, more practical way to implement a cost growth cap would be to purchase cost growth overrun insurance. Such insurance would have a high deductible, such as 100%, but pay dollar-for-dollar for any cost growth above a set amount. However, this insurance would not be free but would cost a substantial amount. One common way to price insurance is to use the equivalence principle, which simply states that the amount of the premium should equal the time-discounted expected payout (Ref. 21). In equation form this is

$$P.V.\ of\ Premiums = P.V.\ of\ Benefits,$$

where P.V. denotes present value. We can calculate this by looking at the mean payout for the 10 missions used in examples in Table 4, which is the mean of the values above the 100% cost growth cap. The total expected claim amount for this case is $2,854. Discounted to the beginning of the year at the current prime rate of 3.25%, the premium is $2,765, approximately 30% of the total budgets for the 10 projects. This is a large amount of money, but it does remove a great deal of uncertainty from the decision process, and may help NASA achieve its scientific objectives. A portfolio effect can be implemented, but it's not free. After all "there ain't no such thing as a free lunch!"

Conditional Tail Expectation

Nearly all risk management for DoD and NASA agencies seem to focus solely on finding a single percentile to budget against. NASA policy mentions 70% and 50%, for example. However risk management doesn't stop at that point. Even funding at a 70% confidence level means that there is roughly a one-in-three chance of experiencing an overrun. What happens then? Does the project get cancelled? Unless there is an automatic implied policy of canceling any projects that overrun their budget, management faces a quandary. Even though they have applied modern risk management techniques, fully one third of all programs will still get into trouble on a regular basis. NASA and other agencies also need to set policy on a course of action once the 70% mark is exceeded. One useful statistical tool that comes in handy in such cases is conditional tail expectation. This is defined as the amount of cost growth to expect given that cost has exceeded a specified amount, that is

$$E\left[X \mid X > Q_\alpha\right],$$

where Q_α is a specifed quantile. For example, $Q_{0.95}$ is the 95th percentile. For a lognormal distribution, the formula for the conditional tail expectation (CTE) is

$$CTE_\alpha = E\left[X\right] \cdot \frac{1 - \Phi\left(\frac{\ln Q_\alpha - \mu - \sigma^2}{\sigma}\right)}{1 - \alpha},$$

where Φ is the cumulative normal distribution function. For example, for a single project for which cost risk has been modeled as a lognormal distribution with mean equal to \$100 million and standard deviation equal to \$50 million, $\mu = 4.49, \sigma = 0.72$, and the 70th percentile is equal to

$$e^{4.49+z_{0.70}\cdot 0.72} \approx \$114.6 \text{ million.}.$$

Thus, in this instance,

$$CTE_{0.70} = 100 \cdot \frac{1-\Phi\left(\frac{\ln 114.6 - 4.49 - 0.47^2}{0.47}\right)}{1-0.7} \approx \$159.7 \text{ million}.$$

Therefore, given that the 70th percentile has been reached, the expected amount needed to complete the project will be \$160 million, roughly \$46 million above the 70th percentile budget. This is 40% more than the budget. Table 5 displays the additional amount expected to be required if the budget is exceeded. Note that this amount ranges from around 10% if the budget is set at the 90th percentile and the coefficient of variation of the lognormal cost risk distribution is 20%, to 185% if the budget is set at the 30th percentile and the coefficient of variation is 100%.

	Budget Set at						
Coefficient of Variation	30th	40th	50th	60th	70th	80th	90th
20%	23.6%	20.5%	18.0%	15.9%	14.0%	12.2%	10.2%
30%	38.0%	32.7%	28.5%	25.0%	21.9%	19.0%	15.8%
40%	54.1%	46.1%	40.0%	34.9%	30.4%	26.2%	21.6%
50%	72.0%	60.9%	52.4%	45.5%	39.4%	33.7%	27.7%
60%	91.6%	76.8%	65.7%	56.7%	48.8%	41.5%	33.9%
70%	112.7%	93.8%	79.7%	68.3%	58.5%	49.5%	40.1%
80%	136.0%	112.0%	94.4%	80.5%	68.6%	57.6%	46.4%
90%	160.0%	131.0%	110.0%	93.0%	78.8%	65.9%	52.7%
100%	185.0%	151.0%	125.5%	105.8%	89.2%	74.2%	59.0%

Table 5. Additional Amount Expected if Budget Exceeded.

Thus, reserve setting cannot stop with simply setting reserves at a relatively high confidence level. NASA and other agencies should expect to occasionally spend much, much more. For example, history indicates that the coefficient of variation is 100%, so if budgets are set at the 70th percentile, the average additional amount needed when an overrun occurs will be 89%! Thus in practice one third of missions will need an additional 89% above the budget, even when budgets are set at the 70th percentile. This has sobering policy implications.

Schedule Growth and Schedule Risk

Schedule growth has not been studied as much as cost growth, but that doesn't mean it is not a significant problem. For this study we collected schedule growth for 48 missions. Schedule growth information has not been studied as much as cost growth, so less information was available.

On a percentage basis schedule does not grow as much as cost. However, 90% of schedules overrun, on par with the 86% of missions that experience cost growth. The average schedule growth is equal to 25.6%. The maximum was also much less than cost, at 113%. Unlike cost, schedule had no under runs. The schedule growth data are summarized in Figure 4.

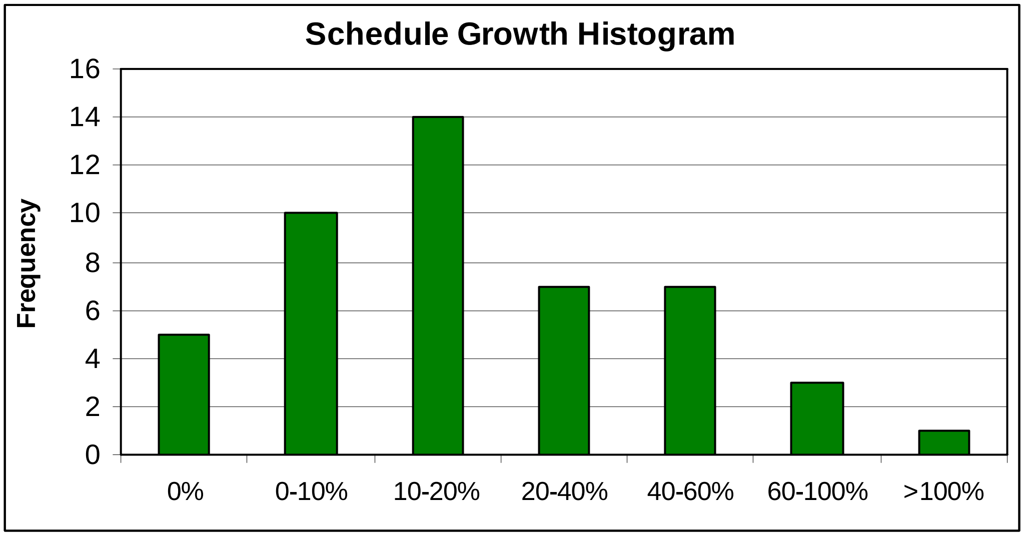

Figure 4. Graphical Summary of NASA Schedule Growth.

When it comes to fitting a distribution to schedule growth, there is no clear-cut winner as with cost growth. Different distributions provide better fits depending upon which test is used to rank the fits. See Table 6 for a comparison of the four distributions with the best test statistics.

Distribution	**Anderson-Darling**	**Chi-Square**	**Kolmogorov-Smirnov**
Weibull	**1.0943**	**3.33**	**0.1237**
Exponential	**17.8774**	**5.33**	**0.1437**
Beta	**0.3802**	**5.67**	**0.0768**
Gamma	**0.8502**	**11.333**	**0.1221**

Table 6. Comparison of Best-Fitting Distributions for Cost-Growth Data.

Note that for schedule growth, the lognormal distribution is not a candidate at all, since it cannot be fit to values at 0. At the 5% significant level, the Kolmogorov-Smirnov critical value is $\frac{1.36}{\sqrt{N}} = \frac{1.36}{\sqrt{48}} \approx 0.1963$, so we cannot reject any of the distributions in Table 6 based on that criterion. Also, the chi-square statistics are such that we do not reject any of these distributions at the 5% significance level. If we use the Anderson-Darling goodness-of-fit test, however, we can reject the exponential. However the goodness-of-fit statistics for the other tests are such that all four seem worthy of some consideration. One way to select a distribution would be to increase the significance level to the point where only one distribution is not rejected, although this would vary from test to test – the beta has the lowest Anderson-Darling and Kolmogorov-Smirnov test statistics, while the chi-square statistic is lowest for the Weibull. Another way to rank among distributions is to use the likelihood function, and use the Schwarz-Bayesian criterion (Ref. 22). The Schwarz-Bayesian criterion is an objective metric that compares the log likelihood values with a penalty for number of parameters, and is defined as

$$l(\theta) - \frac{r}{2} \ln(n),$$

where r is the number of distribution parameters, n is the number of data points, and $l(\theta)$ is the value of the log likelihood function for the fitted parameters θ (in the case of multiple parameters θ is a vector). The criterion chooses the distribution with the highest value. In the case of schedule growth $n=48$, and the number of parameters varies from a single parameter (rate) for the exponential, to four for the beta distribution. The Schwarz-Bayesian values are displayed in Table 7. As seen from the table, according to the

Distribution	SBC
Gamma	29.79
Exponential	15.42
Weibull	9.66
Beta	8.43

Table 7. Schwarz-Bayesian Values for the Beta, Exponential, Gamma, and Weibull Distributions.

Schwarz-Bayesian criterion, the gamma distribution is the best choice. The exponential is ranked second because it has a smaller parameter penalty than the Weibull or beta. See Figure 7 for a graphical comparison. The exponential and Weibull have the same color in the chart because they cannot be distinguished from one another because their graphs are almost coincident. But the Schwarz-Bayesian gives the exponential more credit than the Weibull since the Weibull adds a parameter but provides almost the same fit as the exponential. This means that the addition of the parameter in the Weibull has little or no value.

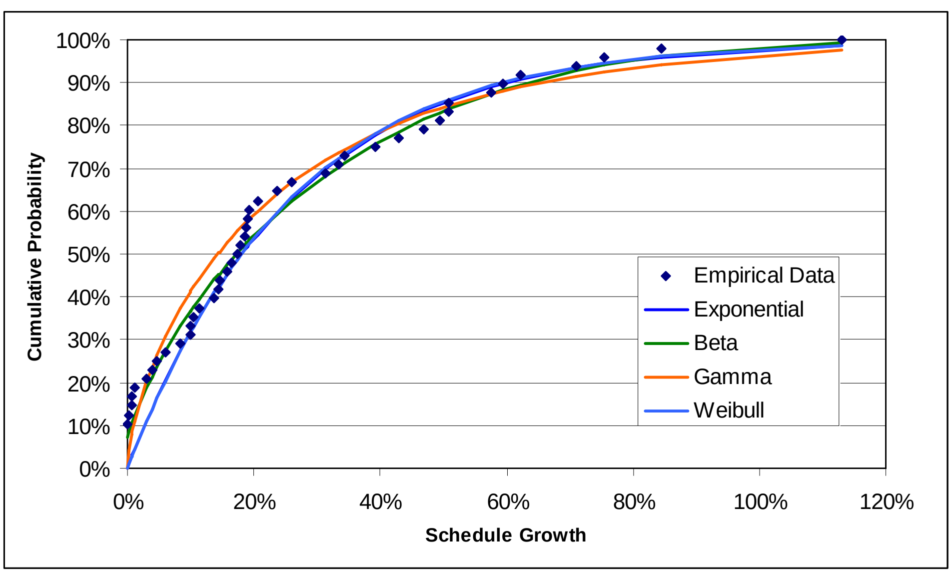

Figure 5. Graphical Comparison of Empirical Schedule Growth Data and Fitted Distributions.

The gamma distribution is a two-parameter distribution with density function defined as

$$f(x) = \frac{x^{k-1}\exp(-x/\theta)}{\Gamma(k)\cdot\theta^{k}},$$

where $\Gamma(k)$ denotes the gamma function evaluated at k. The fitted gamma parameters are $k = 0.655$ and $\theta = 0.39$.

Just like with cost growth and risk, schedule growth and risk are strongly linked. Schedule growth is smaller as a percentage than cost growth which implies that the schedule risk mean should be closer to the schedule point estimate than the cost risk mean to the cost point estimate. Schedule risk can also be modeled as a gamma distribution. Based on the empirical data the point estimate is at the 10th percentile, and the mean is 25% higher. Given this information, a gamma distribution can be derived that matches the mean and the desired percentile. The mean of a gamma distribution is $k \cdot \theta$ and the 10th percentile of a gamma distribution can be found by solving

$$0.10 = \frac{1}{\Gamma(k)}\int_{0}^{x/\theta} t^{k-1}e^{-t}dt .$$

While this is not a nice, easily solved equation, the parameters for a gamma distribution with given mean and 10th percentile can be derived using Excel's Solver capability and by using the built-in gamma distribution and gamma function. For example given a 72-month baseline schedule, given no other information about risk and by using history as a guide, one can set this value equal to the 10th percentile, and the mean equal to 1.25 times the point estimate, or 90 months. Using Excel Solver, setting the objective function to equal 90 months by varying the cells corresponding to the initial gamma parameters, with a constraint that the 10th percentile equal 72 months, one finds that the parameters of the gamma distribution are $k = 38.57$ and $\theta = 2.33$. Note that in this case, the variance is equal to $38.57 \cdot 2.33^{2} \approx 210$, and thus the standard deviation is approximately equal to 14.5, or 16.1% of the mean. That is the coefficient of variation for schedule risk based on history is approximately 16%. Comparing this with the 100% coefficient of variation implied by cost growth history, we see that on a percentage basis, schedules have much less risk than cost. See Figure 6 for a graphical comparison of relative cost and schedule risk, with both cost and schedule means set equal to 1. This does not imply that schedules are not risky, however. Schedule risk remains an important consideration for project managers, and projects should expect to experience schedule growth on a regular basis.

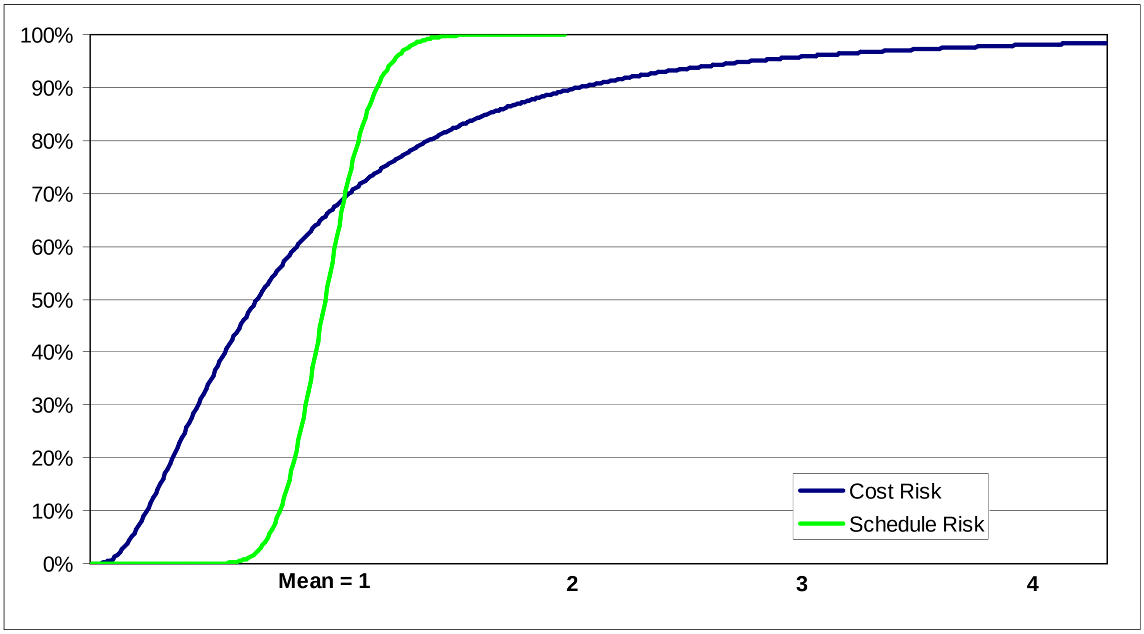

Figure 6. Schedule Risk Vs. Cost Risk Based on Empirical Data.

Cost Risk and Schedule Risk Correlation

Cost risk and schedule risk do not occur in isolation. They occur jointly and should be analyzed together. Project success involves meeting both a cost target and achieving a schedule, so confidence assessments should involve both as well. Recent NASA policy guidance has specified that confidence levels be set for joint cost and schedule, which is more stringent than considering cost or schedule alone. One simple way to assess joint confidence, proposed by Paul Garvey (Ref. 23), is to measure schedule risk and cost risk separately, then combine the two into a joint probability distribution by assigning correlation between cost and schedule. For the expanded 48-mission dataset, the correlation between cost and schedule growth is 71.5%. This provides a proxy for estimating cost and schedule risk correlation in the absence of any other correlation information.

Conclusion

NASA and other agencies have made significant strides in adopting modern financial risk management techniques. However, the portfolio effect is a pitfall that government agencies should avoid relying upon, as it is an imaginary free lunch that does not exist. However, cost overrun insurance offers the promise of a true portfolio effect if it can be implemented well. Despite these strides, however, the focus to date has been entirely upon setting budgets based on confidence levels. As we have demonstrated, this is not sufficient. Projects budgeted at the 70^{th} percentile will on average require a great deal more money to complete. Setting reserves at level near the bulk of the distribution, such as at the 70^{th} percentile, is a recipe for disaster without additional set asides for missions that exceed these limits.

References

1. Anderson, T. P. "The Trouble With Budgeting to the 80th Percentile", presented at the 72nd Military Operations Research Society Symposium, 2004.
2. Smart, C., "The Portfolio Effect Reconsidered," Presented at the 2007 Joint ISPA-SCEA Conference, New Orleans, LA, June 2007.
3. Smart, C., "The Fractal Geometry of Cost Risk," Presented at the 2008 Joint ISPA-SCEA Conference, Noordwijk, the Netherlands, May, 2008.
4. Smart, C., "Cost Growth Correlation for Shuttle and Apollo," unpublished white paper, 2007.
5. Shaffer, M., "NASA Cost Growth: A Look at Recent Performance," 2003, unpublished Powerpoint presentation, NASA HQ.
6. United States Government Accountability Office, "Space Missions Require Substantially More Funding Than Initially Estimated", 1992, GAO-NSIAS-93-97, U.S. GAO, Washington, D.C.
7. Smart, C., "Predicting Differences Between Estimated and Actual Cost and Schedule," 2002, unpublished white paper.
8. Smart, C., "Cost and Schedule Interrelationships," Presented at the 2007 NASA Cost Symposium, Denver, Colorado, July 17-19, 2007.
9. Mog, R.A., and D.L. Thomas, "An Historical Assessment of the Impact of Technology Readiness Estimation on NASA Programmatic Cost Growth," MSFC Presentation, August, 2003.
10. Freaner, C., and Bitten, B., "SMD Cost Database," Excel Spreadsheet, 2008.
11. Rutkowski, B., "Cost Growth Management Reserves," unpublished white paper, 2007.
12. MSFC Engineering Cost Office, "TSS, OTA, OMV, ET, SSME, SRM Historical Cost Growth, Cost Drivers," unpublished white paper, 1992.
13. Pine, D., "NASA Cost Growth Experience & the Role of Cost Estimating," Presentation to NASA HQ, 1992.
14. Mandlebrot, B.B., *The Fractal Geometry of Nature*, 1983, W.H. Freeman and Company, New York.
15. Mandlebrot, B.B., *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, 1997, Springer Verlag, New York.
16. Cramer, J. "The Only Free Lunch on Wall Street," Mad Money (television), February 6, 2007, transcript available at http://www.madmoneystocks.com/jim-cramer.aspx?id=2/6/2007&seg=4 .
17. Heinlein, R.A., *The Moon is a Harsh Mistress*, 1966, G. P. Putnam's Sons, New York.
18. Friedman, M., *There's No Such Thing as a Free Lunch*, 1975, Open Court Pub. Co., Chicago.
19. Gray, J., *Janos Bolyai, Non-Euclidean Geometry, and the Nature of Space*, 2004, Burndy Library Publications, Cambridge, MA.
20. Kiss, E., *Mathematical Gems from the Bolyai Chests: Janos Bolyai's Discoveries in Number Theory and Algebra*, 1999, TypoTEXLtd. Electronic Publishing Co., Budapest.
21. Bowers, N.L., *Actuarial Mathematics*, Society of Actuaries, Schaumburg, Ill., 1997.
22. Schwarz, G., "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461-464, 1978.

23. Garvey, P.R., *Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective*, New York, Marcel Dekker, 2000.