The Portfolio Effect Reconsidered

Christian Smart, Ph.D., CCE/A
MCR, LLC

2007 Joint ISPA-SCEA Conference

"Creating Customer-Focused Success"

Cost Estimation

- Cost is a critical consideration in military and space systems analysis models and decision criteria
- The cost of systems is influenced by numerous factors that are not known with certainty
- Historically many NASA and military programs have been subject to cost overruns
  - "NASA Program Costs: Space Missions Require Significantly More Funding Than Initially Estimated," General Accounting Office, 1992
- Thus inclusion of uncertainty in cost estimates is critical for project planning

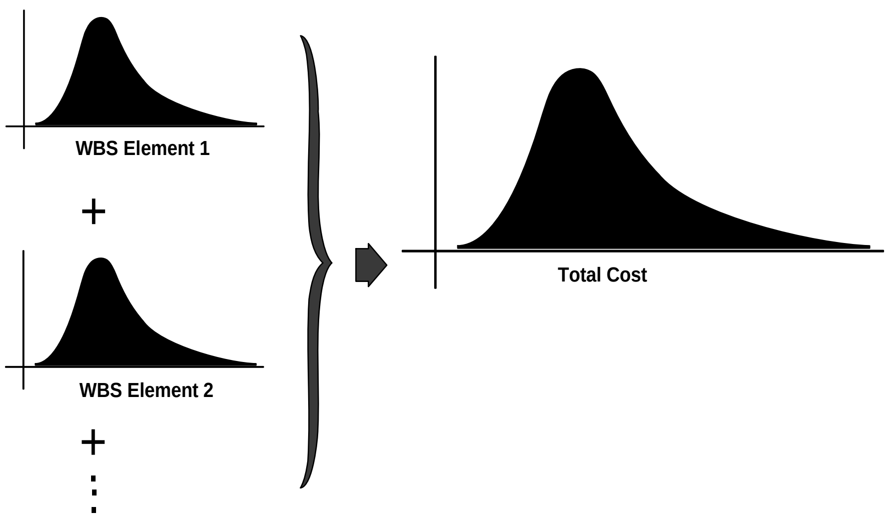
"Creating Customer-Focused Success"

Cost Uncertainty - Terminology

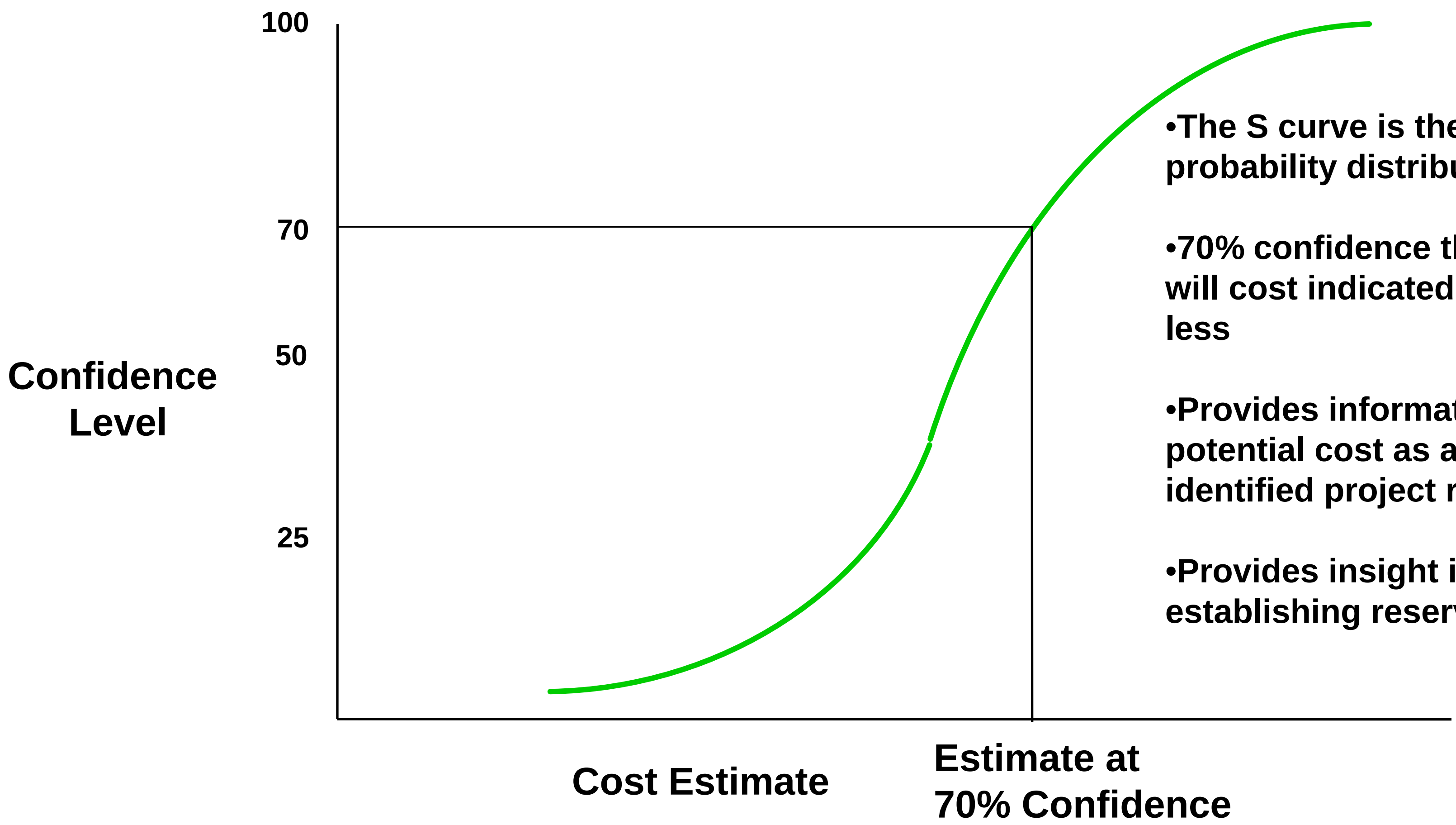
- *Risk* is the chance of uncertainty or loss
  - In a situation that includes potentially favorable and unfavorable events, risk is the probability that an unfavorable event occurs
- *Uncertainty* is the indefiniteness about the outcome of a situation
  - Uncertainty includes both favorable and unfavorable events
- *Cost Risk* is a measure of the chance that, due to unfavorable events, the planned or budgeted cost of a project will be exceeded
- *Cost Uncertainty Analysis* is a process of quantifying the cost estimating uncertainty due to variance in the cost estimating models as well as variance in the technical, performance and programmatic input variables
- *Cost Risk Analysis* is a process of quantifying the cost impacts of the unfavorable events

Risk-Adjusted Cost Estimates

- Cost for a project is built up by adding the cost of the WBS elements
- WBS elements have, historically, been viewed as deterministic values
- Each of these WBS cost elements is actually a probability distribution
- Adding the most likely costs of *n* WBS elements that are right skewed, yields a result that is less than 50 probable
  - Often only 30% probable
- These individual WBS element distributions are aggregated using analytic approximation or Monte Carlo simulation



S-Curves and Confidence Levels



- The S curve is the cumulative probability distribution
- 70% confidence that project will cost indicated amount or less
- Provides information on potential cost as a result of identified project risks
- Provides insight into establishing reserve levels

"Creating Customer-Focused Success"

Interpretation of Confidence Levels

- Funding at 30% confidence means that the probability of the actual cost exceeding the budget is 70%
  - Funding at higher confidence levels will decrease the chance of a cost overrun, for example:
    - Funding at 50% confidence means that the probability of the actual cost exceeding the budget is 50% (coin toss)
    - Funding at 70% confidence means that the probability of the actual cost exceeding the budget is 30%
- Funding at a high level of confidence will help prevent cost overruns, but at a price
  - If funds are not needed, other projects that could use those funds may be delayed or cancelled
- Decision makers must therefore balance the need to fund programs sufficiently to prevent cost overruns with the ability to sufficiently fund a wide array of programs

"Creating Customer-Focused Success"

Confidence Levels and Multiple Systems

- Programs often consist of multiple projects
- Program planners often have to budget at both the individual project level and the total program level
- One example is the NASA's Crew Launch Vehicle (CLV)
  - The CLV is a two stage launch vehicle that consists of multiple separate projects, including the first stage, the upper stage, and the liquid rocket engine for the upper stage
  - Cost risk must be managed both for CLV by managing risk at the CLV level and at the lower levels
    - What is the relationship between the confidence level between the lower levels and the CLV level?

Modern Portfolio Theory

- Modern Portfolio Theory, as expounded by Nobel Laureate Harry Markowitz, states that investors can reduce risk and increase risk-adjusted returns by diversifying their investments
- Diversification is often called the “only free lunch on Wall Street”

"Creating Customer-Focused Success"

Application of Portfolio Theory to Systems Risk

- First researched by Tim Anderson of the Aerospace Corporation, the portfolio effect can be applied to understanding the relationship between confidence at the systems level and confidence at the individual project level
- The basic idea is that by diversifying among projects, decision makers can fund individual projects at lower levels of confidence while achieving higher confidence levels when all the missions are considered together (i.e., a portfolio of missions)
  - This phenomenon is the "portfolio effect"

Anderson, Timothy P. "The Trouble With Budgeting to the 80th Percentile" ; The Aerospace Corporation; 72nd Military Operations Research Society Symposium; June 22 – 24, 2004.

"Creating Customer-Focused Success"

Example

- If we want to ensure, say, an 80% probability that our program budget will not be exceeded, then we need to determine the individual percentiles that, when summed, correspond to the 80th percentile of the program cost

Project	μ	σ	61st %ile
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

Source: Anderson, Timothy P. "The Trouble With Budgeting to the 80th Percentile" ; The Aerospace Corporation; 72nd Military Operations Research Society Symposium; June 22 – 24, 2004.

Potential Pitfalls

- “Behold, the fool saith, ‘Put not all thine eggs in the one basket’ – which is but a manner of saying, ‘Scatter your money and your attention;’ but the wise man saith, ‘Put all your eggs in one basket and watch that basket!”
  - Mark Twain, Pudd’n’head Wilson, 1894
- The diversification effect only works if
  - There is a sufficiently large number of projects.
  - None of the projects is very large compared to the rest.
  - Funding at the overall level is at a high percentile (~ 80th percentile).
  - Projects have low correlation with one another.
  - Project risks follow a Normal, Lognormal or similar distribution with a small tail

"Creating Customer-Focused Success"

A Comparison with NASA History

- In 2004 Matt Schaffer of NASA HQ collected and analyzed budget data on cost growth for NASA missions
  - Comprised 50 missions from the 1990s – present
  - Cost growth ranged from -25% to +193%
  - Average cost growth was 35%
  - 76% of the missions had budget overruns
    - Similar to studies by Goddard and GAO
  - Data are conservative
    - Does not completely account for changes in requirements and scope before ATP (accounting for this would reduce the reported cost growth for some missions, e.g., Rossi XTE)
  - 12% of the missions experienced cost growth of more than 100%

"Creating Customer-Focused Success"

A Comparison with NASA History

- Consider the example presented above

Project	Mean	σ	30th Percentile	100% Growth	Probablity of 100%+ Growth
Project 1	$1,696	$539	$1,413	$2,827	1.80%
Project 2	$1,481	$404	$1,269	$2,538	0.44%
Project 3	$1,395	$435	$1,167	$2,334	1.55%
Project 4	$874	$288	$723	$1,446	2.35%
Project 5	$840	$219	$725	$1,450	0.27%
Project 6	$1,449	$371	$1,254	$2,509	0.21%
Project 7	$1,638	$537	$1,356	$2,713	2.27%
Project 8	$1,031	$259	$895	$1,790	0.17%
Project 9	$1,271	$323	$1,102	$2,203	0.19%
Project 10	$1,937	$602	$1,621	$3,243	1.50%
Total	$13,612	$1,317			

- Note that I have added two columns one that represents the cost that represents 100% cost growth and another column that represents the probability of the occurrence of that amount of cost growth

"Creating Customer-Focused Success"

A Comparison with NASA History

- For the 10 missions in the table, the probability of 100%+ growth for one mission varies from 0.17% to 2.27%
- For Project 1, the probability that it will experience at least 100% cost growth is 1.8%
  - Denote this as event *A* and the probability of this event's occurrence as *P(A)*
  - The complement of this event, denoted by A', is the probability that Project 1 experiences less than 100% cost growth
  - Note that from basic probability

$$P(A) + P(A') = 1 \text{ or } P(A') = 1 - P(A)$$

  - The probability that Project 1 will experience less than 100% cost growth is therefore

$$P(A') = 1 - P(A) = 1 - .018 = 98.2\%$$

"Creating Customer-Focused Success"

A Comparison with NASA History

- For Project 2, the probability that it will experience at least 100% cost growth is 0.44% (denote this as event *B* and the probability of this event's occurrence as *P(B)*)
  - The probability that it will experience less than 100% cost growth is

$$P(B') = 1 - P(B) = 1 - .0044 = 99.56\%$$

- For the two projects, Project 1 and Project 2, the probability that at least one experiences 100% cost growth is given by the formula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Note that we assumed that A and B are independent, which means that

$$P(A \cap B) = P(A)P(B)$$

A Comparison with NASA History

- Therefore, based on the independence assumption, the probability that at least one of the two experiences 100%+ cost growth is given by the formula

$$P(A \cup B) = P(A) + P(B) - P(A)P(B)$$

- The probability that at least one project experiences 100%+ cost growth is thus

$$P(A \cup B) = 0.018 + 0.0044 - 0.018 \cdot 0.0044 \approx 2.23\%$$

"Creating Customer-Focused Success"

A Comparison with NASA History

- Note that this can be more easily calculated by noting that

$$P(A \cup B) = 1 - P((A \cup B)')$$
$$= 1 - P(A' \cap B') \quad \text{(De Morgan's Laws)}$$

- Since the independence of A and B implies the independence of A' and B'

$$P(A \cup B) = 1 - P(A' \cap B')$$
$$= 1 - P(A')P(B')$$
$$= 1 - (1 - P(A))(1 - P(B))$$

"Creating Customer-Focused Success"

A Comparison with NASA History

- This latter formula can be extended to find the probability that at least one of these 10 missions will experience at least 100% cost growth is 10%
- Denote the probability that the *i^th^* mission ($i = 1,\ldots,10$) experiences at least 100% cost growth by A_i
- Probability that at least one of the 10 missions experiences more than 100% cost growth is given by the formula

$$1 - \prod_{i=1}^{10} \left(1 - P(A_i)\right)$$

- Using this formula, the probability that at least one of the 10 missions given in the example experiences more than 100% cost growth can be calculated as 10.1%

"Creating Customer-Focused Success"

A Comparison with NASA History

- But based on a historical 100%+ cost overrun rate equal to 12%, the probability that at least one mission in 10 will experience 100%+ cost growth is

$$1-\prod_{i=1}^{10}\left(1-P(A_i)\right)=1-\left(1-0.12\right)^{10}$$

$$=1-0.88^{10}$$

$$\approx 72.1\%$$

- Note the striking difference between the overrun probability given in the example (~10%) vice the probability based on history

A Comparison with NASA History

- Part of the discrepancy for the results in the table may be due to the fact that the 10 example missions are less risky on average than history would suggest
- The average coefficient of variance for the 10 missions in the table have an average coefficient of variance of approximately 29%
- How does this compare to history?

"Creating Customer-Focused Success"

A Comparison with NASA History

- Cost risk growth data can be used as a means to check the results of risk analyses against reality
- These checks can be used to determine if the amount of risk in the cost risk analyses is too high or too low
- Use 2004 NASA HQ cost growth study

"Creating Customer-Focused Success"

A Comparison with NASA History

- Cost risk is the probability of exceeding the initial estimate
- Cost growth is the actual amount that the initial estimate is exceeded
- Assumption - the initial budgets in the cost growth database are point estimates (no risk is included)
- By assuming that the initial estimates are point estimates, we can relate cost risk to cost growth
  - For example, if the point estimate represents the 30^{th} percentile of a cost risk distribution, then the ratio of the 70^{th} percentile to the 30^{th} percentile represents potential cost growth
- For A, B two points of a cost risk distribution (A > B), with B as an initial reference point, the following formula relates cost growth to cost risk
  - Cost growth = A/B
- A cost growth distribution is simply the ratio of various percentiles on a cost risk distribution relative to an initial reference point, such as the 30^{th} percentile

"Creating Customer-Focused Success"

A Comparison with NASA History

- Assume that the 30th percentile on a cost risk S-curve represents the point estimate (initial estimate)
  - From experience, the point estimate is typically at or below the 30th percentile
- Assume that the risk distribution is Lognormal
- For NAFCOM estimates, the ratio of the standard deviation to the mean is typically between 1/3 and 1/2 of the mean
- The ratio of the 70th percentile to the 30th percentile of a lognormal is

$$\frac{e^{P+0.5244Q}}{e^{P-0.5244Q}} = e^{1.0488Q}$$

- When $\sigma = a\mu$, it follows that

$$Q = \sqrt{\ln(1+\frac{(a\mu)^2}{\mu^2})} = \sqrt{\ln(1+a^2)}$$

- When a = 1/3, the ratio of the 70th percentile to the 30th percentile is 1.4
- Thus, a reasonable rule of thumb for the ratio of the 70th percentile to the point estimate is 1.4

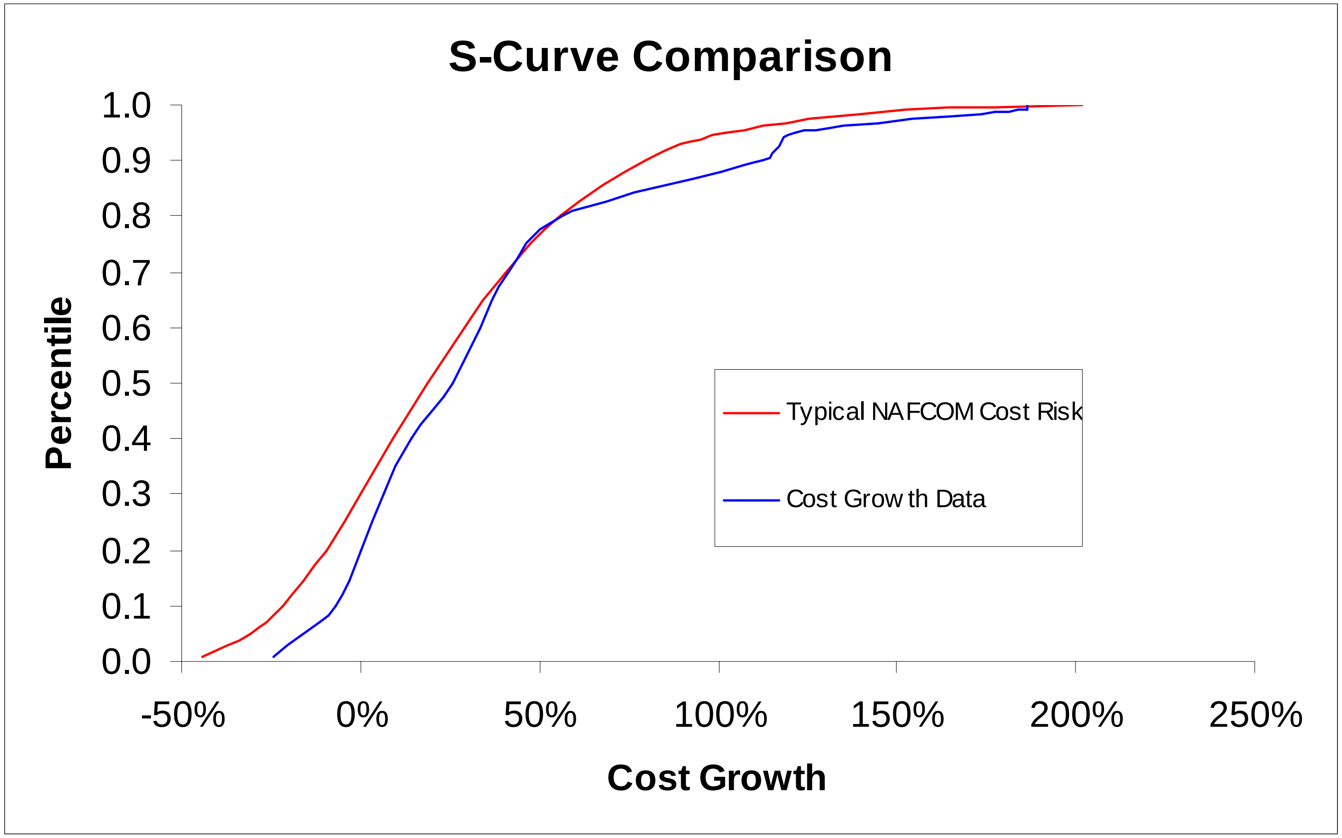
"Creating Customer-Focused Success"

A Comparison with NASA History

- The derivation of the rule of thumb on the previous chart provides a method to convert cost risk into cost growth
  - Select a reference point, and divide each percentile by the initial reference point
  - In this analysis, it is assumed that the 30th percentile is the reference point
  - Once the cost risk has been normalized to a cost growth curve, it can be directly compared to the cost growth data
  - Assume that a = 1/3 for "typical" risk

“Creating Customer-Focused Success”

A Comparison with NASA History



- Notice how closely a “typical” NAFCOM risk distribution fits the bulk of the actual cost growth data
  - Provides confidence that cost risk estimates produced by NAFCOM are realistic

"Creating Customer-Focused Success"

A Comparison with NASA History

- Based on history, a coefficient of variance that reflects NASA history is equal to 1/3 or approximately 33.3%
- Recall that the average coefficient of variance for the 10 missions in the example have an average coefficient of variance of approximately 28.9%
- The table below illustrates the adjusted standard deviation so that the standard deviations are more in line with recent NASA experience

Project	Mean	σ	CoV	Adjusted CoV	Adjusted σ
Project 1	$1,696	$539	0.318	0.365	$620
Project 2	$1,481	$404	0.273	0.314	$465
Project 3	$1,395	$435	0.312	0.359	$500
Project 4	$874	$288	0.330	0.379	$331
Project 5	$840	$219	0.261	0.300	$252
Project 6	$1,449	$371	0.256	0.294	$427
Project 7	$1,638	$537	0.328	0.377	$618
Project 8	$1,031	$259	0.251	0.289	$298
Project 9	$1,271	$323	0.254	0.292	$371
Project 10	$1,937	$602	0.311	0.357	$692
Average	$1,361	$398	0.289	0.333	$457

A Comparison with NASA History

- The table below shows the probability of 100% cost growth for the adjusted example

Project	Mean	σ	30th Percentile	100% Growth	Probability of 100%+ Growth
Project 1	$1,696	$620	$1,371	$2,742	4.58%
Project 2	$1,481	$465	$1,237	$2,475	1.62%
Project 3	$1,395	$500	$1,133	$2,265	4.09%
Project 4	$874	$331	$700	$1,401	5.59%
Project 5	$840	$252	$708	$1,416	1.11%
Project 6	$1,449	$427	$1,225	$2,451	0.95%
Project 7	$1,638	$618	$1,314	$2,628	5.44%
Project 8	$1,031	$298	$875	$1,750	0.79%
Project 9	$1,271	$371	$1,076	$2,152	0.88%
Project 10	$1,937	$692	$1,574	$3,148	4.01%
Total	$13,612	$1,514			

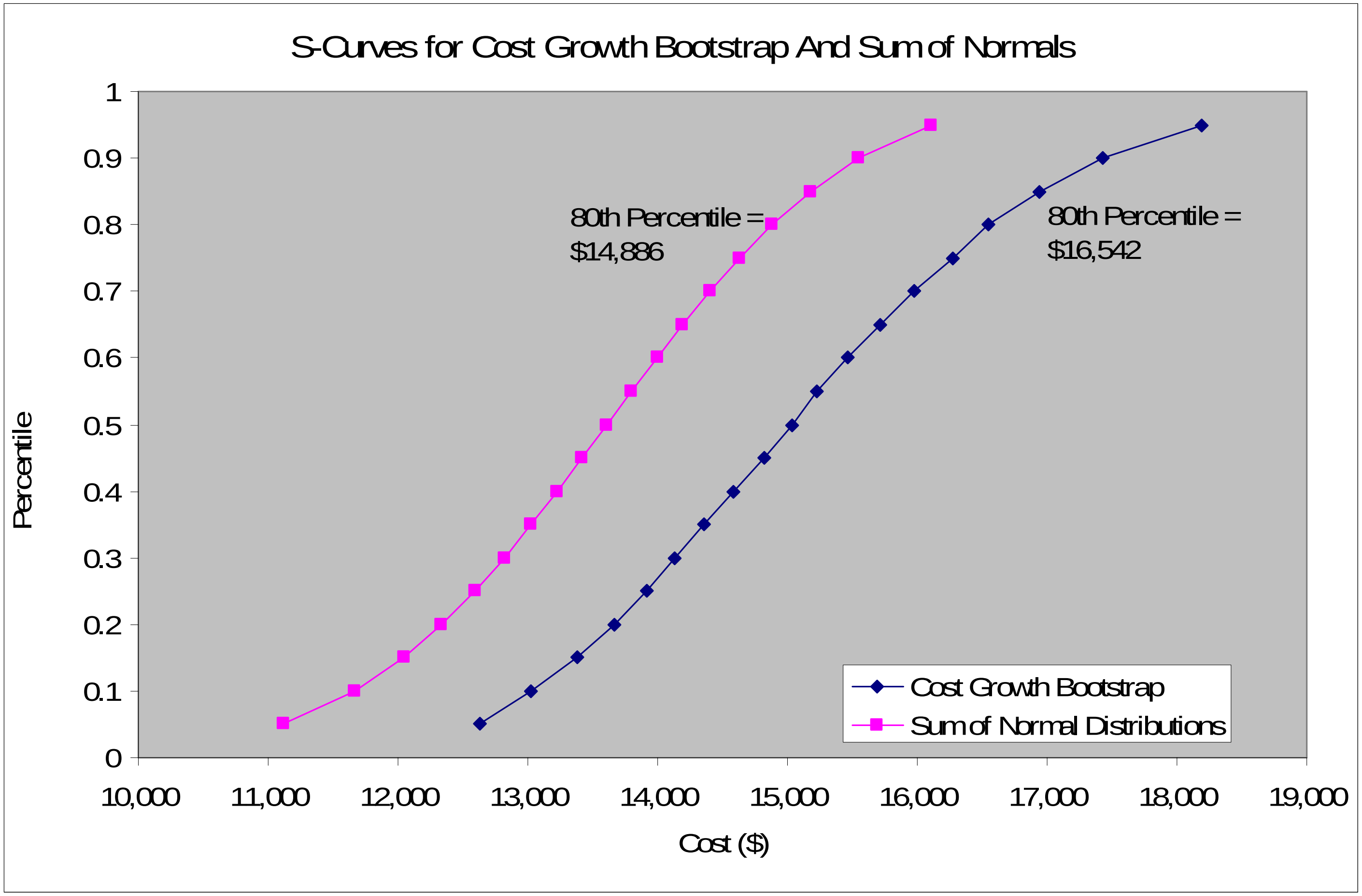
- Using the formulas discussed earlier, the probability of 100%+ cost growth for at least one of the 10 missions in the adjusted table is 26%
- While higher than 10%, it is much lower than the 71% suggested by history

"Creating Customer-Focused Success"

A Comparison with NASA History

- In order to further study the effects of reality on portfolio risk, we performed a bootstrap simulation of the cost growth data
  - 5,000 iteration Latin hypercube simulation
- For the cost growth simulation, for each sample
  - Used the 30th percentiles from the pervious table as the point estimates
  - For each of the 10 missions, simulated cost growth using the actual NASA cost growth data (the bootstrap)
    - For each of the 10 missions, we obtained a percentage P_i that ranges from -25% to +193%
  - Multiplied each 30th percentile by $1+P_i$ $(i= 1,\ldots,10)$
  - Summed the 10 values to obtain a total cost
- The end result is a cost risk distribution that is based on actual recent history

A Comparison with NASA History



"Creating Customer-Focused Success"

A Comparison with NASA History

- Notice that the two S-curves in the figure above have a similar shape, but the cost growth bootstrap S-curve is shifted substantially to the right of the sum of the Normal distributions
- Normal - the 80th percentile is approximately $14,885
  - At this level, the individual projects could be budgeted at the 61st percentile
- Cost growth data - 80th percentile is much higher at approximately $16,542. At this level, each individual project must be budgeted at the 74th percentile

Project	Mean	σ	74th Percentile
Project 1	$1,696	$620	$2,094.78
Project 2	$1,481	$465	$1,779.90
Project 3	$1,395	$500	$1,716.83
Project 4	$874	$331	$1,087.08
Project 5	$840	$252	$1,002.03
Project 6	$1,449	$427	$1,723.48
Project 7	$1,638	$618	$2,035.30
Project 8	$1,031	$298	$1,222.62
Project 9	$1,271	$371	$1,509.97
Project 10	$1,937	$692	$2,382.39
Total	$13,612	$1,514	$16,554.37

"Creating Customer-Focused Success"

A Comparison with NASA History

- The differences are caused by
  - Skewness in the cost growth data
    - Normal is symmetric
    - Lognormal may correct for this
  - Fatter tail than the Normal distribution
    - Lognormal also has a thin right tail
      - May have to use a fatter-tailed distribution to represent cost risk such as Pareto, Cauchy, or a discrete risk distribution
      - Another possibility is to use a lognormal to represent risk for nominal cost risk and perform scenario analysis to estimate the likelihood and consequences of events that may cause high cost growth

"Creating Customer-Focused Success"

Conclusions

- Under the right conditions, the portfolio effect can allow decision markers to fund individual projects at a lower level while still achieving a high level of confidence for the overall program
- However, in many situations the portfolio effect does not apply or is relatively weak
- Also, the portfolio effect relies to some extent on a low probability of a large overrun, which is more common than theory predicts
  - This conclusion is tentative because it is based on several assumptions and further analysis is required