The Impact of Correlation on Risks in Programs and Projects [1]

By Bob Prieto

Chairman & CEO
Strategic Program Management LLC

One of the most under considered elements of cost and schedule risk is the correlation that exists within various WBS elements of a project or across projects comprising a program. Failure to adequately consider correlation between various activities and projects compounds the impact of other factors present in large complex projects.

These include:

- MAIMS – "Money Allocated Is Money Spent"
- Parkinson's Law – work expands to fill the time allotted
- Overconfidence in assessing uncertainties
- Complexity with hidden coupling – risk events are likely to affect multiple cost elements with the potential for cascading impacts
- State of technology
- Common management, staff and work processes
- Optimism bias
- Overly simplistic probabilistic cost analysis (PCA)

This paper looks at correlation in project and program risk assessments and some of the impacts of a failure to adequately consider such correlation in project risk assessments related to both cost and schedule.

Key points:

- Recognize that correlations exist not only among cost elements in a project but among the cost elements in different projects in a program
- As correlation grows, the probabilistic cost distribution curve broadens (higher standard distribution) requiring higher budgets at a given confidence level
- As the number (n) of correlated activities in a project or projects in a program grows so too does the variance in total project costs (proportional to n^2 at higher n). Correlation effects increase with the number of cost elements in a project or projects in a program

---

[1] How to cite this paper: Prieto, R. (2020). The Impact of Correlation on Risks in Programs and Projects; *PM World Journal*, Vol. IX, Issue XII, December.

- In the absence of any correlation, the probabilistic cost distribution narrows as the number of activities or projects increases (proportional to 1/√n)
- Correlation does not change the expected costs of individual cost elements but instead only changes the "portfolio"[2] standard deviation
- The behaviors in schedule risk analysis depends on whether tasks/projects are executed serially, rolled up or executed in parallel

Correlation and its impacts in projects and programs

Before delving into correlation more fully, it is important to highlight the effects of correlation and some of the erroneous behaviors failing to consider it may drive. In simple terms, as correlation grows, the probabilistic cost distribution curve broadens (higher standard distribution) requiring higher budgets at a given confidence level (P65; P80 etc.).

This can be seen in Figure 1[3].

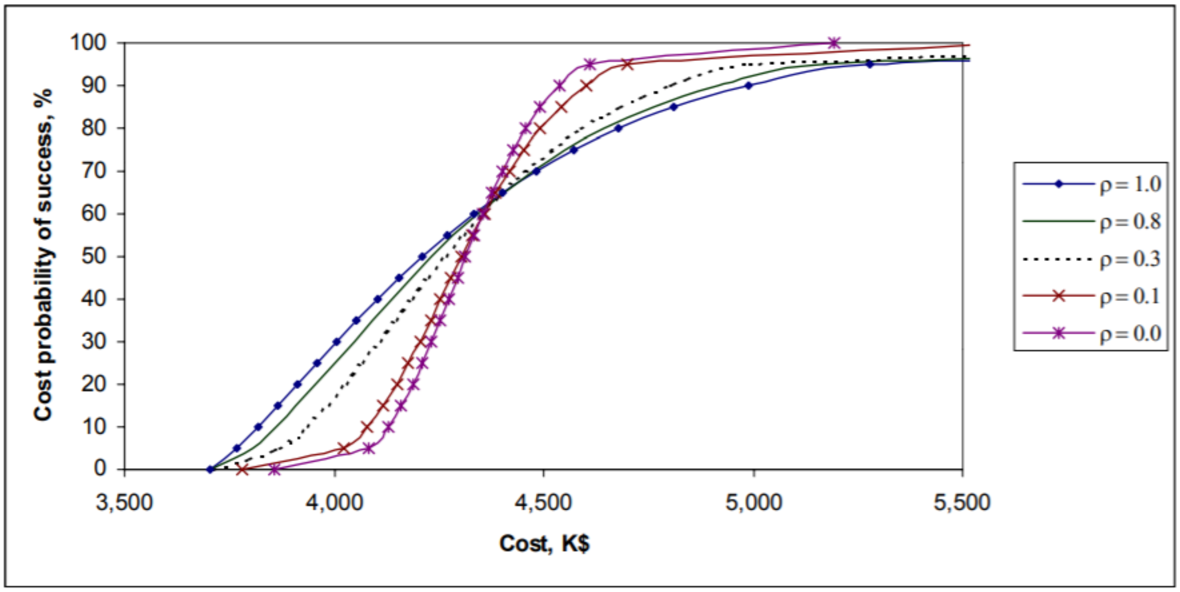


Figure 1

As the number of correlated activities in a project or projects in a program grows so too does the variance in total project costs (proportional to n^2 at higher n). Correlation effects increase with the number of cost elements in a project or projects in a program[4].

---

[2] All WBS cost elements in a project or projects in a program.
[3] Kujawski, Edouard & Alvaro, Mariana & Edwards, William. (2004); Figure 1b
[4] Book, 1999 and 2000/2001

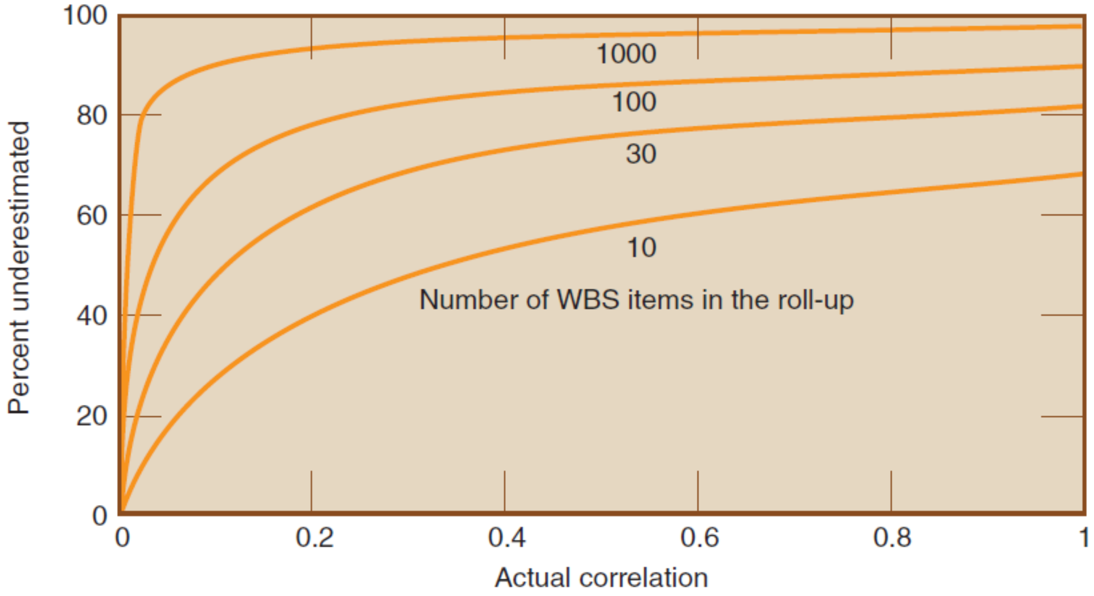


Figure 2[5]

Conversely, in the absence of any correlation (independent random variables), the probabilistic cost distribution narrows as the number of activities or projects increases (proportional to 1/√n). **This often leads to a management decision to unacceptably decompose cost elements leading to a normally distributed cost distribution. This conflicts with the reality that large complex projects are characterized by distributions skewed to the right and much broader than what decomposition might suggest.**

Positively correlated elements have the effect of spreading the distribution of total cost. We must also recognize that correlations exist not only among cost elements in a project but among the cost elements in different projects in a program. The former are driven by characteristics such as complexity, common staff and processes, while the later arise from organizational and programmatic factors common across projects.

---

[5] See Book 2000/2001. This graph illustrates the importance of working with the numeric correlations between WBS items. Assuming these correlations to be zero causes a detrimental effect on the estimation of total-cost uncertainty. Shown is the percentage by which the sigma value (standard deviation) of the total-cost distribution is underestimated, assuming WBS inter-element correlations to be zero instead of the actual value (usually represented by ρ, the Greek letter rho). The horizontal axis tracks ρ, and the vertical axis tracks the percentage by which the total-cost sigma value is, for each nonzero correlation value, underestimated if the correlations are instead assumed to be zero. Each curve is keyed to a unique value of n, the number of elements in a roll-up. As the four curves show, the percent by which sigma is underestimated also depends on the number of WBS items for which the pairwise correlations are incorrectly assumed to be zero. For example, if n = 30 WBS items, and all correlations between WBS items (ρ) are 0.2, but the estimator assumes they are all zero, the total-cost sigma values would be underestimated by about 60%. (This is meant to be a generic illustration and therefore is only approximately true in any specific case. It has been assumed that the sigma values for the WBS items are the same throughout the entire structure.)

So, what is correlation?

Correlation measures linear dependence between two random variables and as such only provides us with a partial picture of their dependence. It does not indicate causality and even a correlation coefficient of 1.0 does not indicate causality, only perfect dependence. When we speak of correlation we typically are referring to "Pearson's" product moment coefficient[6]. When data is nonlinear, non-parametric correlation may be more robust.

Correlation does not tell the whole story as was demonstrated by Anscombe's quartet, where four data sets with nearly identical correlation and other significant statistical properties look very different when graphed[7].

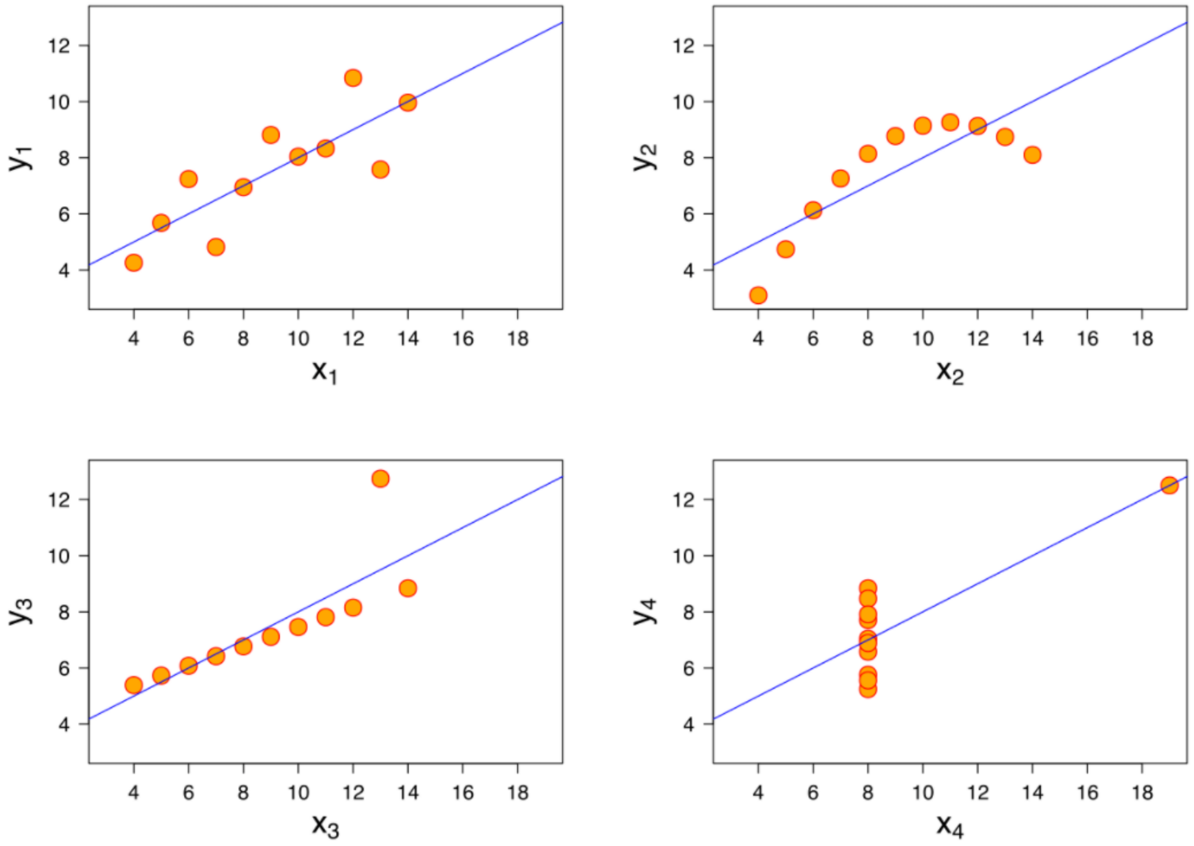


Figure 3

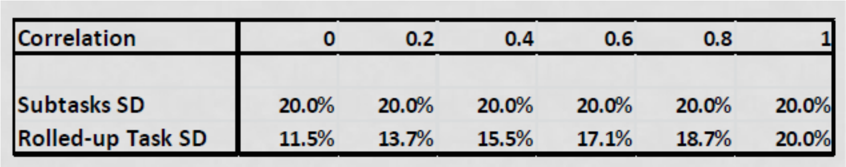
---

[6] Other definitions for correlation include rank correlation and Kendall's Tau. Both are non-parametric measures.
[7] https://en.wikipedia.org/wiki/Anscombe%27s_quartet

The correlation coefficient is dependent on the variance[8] of the data and degrades with volatility of the data. Correlation does not change the expected costs of individual cost elements but instead only changes the "portfolio"[9] standard deviation. This in turn changes the shape of the traditional S-curve, increasing budget requirements for confidence levels greater than the expected value (P50) (where we are less confident in point estimates). Conversely, below expected values we are actually more confident in point estimates than without correlation. These relationships can all be seen in Figure 1.

The behaviors in schedule risk analysis depends on whether tasks/projects are executed serially, rolled up or executed in parallel. Serial execution has a high correlation coefficient which tends to tilt the S-curve while the variance of rolled up tasks is dependent on the variances of the individual subtasks and the degree of correlation.

This can be seen in the following table[10] for a specific example.



Correlation	0	0.2	0.4	0.6	0.8	1
Subtasks SD	20.0%	20.0%	20.0%	20.0%	20.0%	20.0%
Rolled-up Task SD	11.5%	13.7%	15.5%	17.1%	18.7%	20.0%

In the situation where we are able to execute the various tasks or projects in parallel, we see increasing correlation in parallel tasks. This reduces mean duration (more synchronized to dominant task/project) but increases the variance as seen in Figure 3.

---

[8] Square of the standard deviation

[9] All WBS cost elements in a project or projects in a program.

[10] Kuo; NASA

Column1	Deterministic Value	Deterministic Probability	Mean	Standard Deviation
corr=0	300	9%	364	52.4
corr=50	300	20%	350	58.2
corr=100	300	54%	301	60.3

Figure 3[11]

Sources of correlation

Having discussed the implication of correlation on cost and schedule probabilistic assessments it is worthwhile to identify some sources of correlation that we may find in large complex programs. These are summarized in Table 1. Other sources of correlation exist and the table is meant to suggestive.

Table 1
Sources of Correlation

Cost Correlation	Event Correlation
Project competence	Changed requirements
Project organization	Stakeholder influence/action
Project management and processes	Market conditions
Decision and approval processes	Economic trends
Estimation and risk processes	F/X rates

---

[11] Kuo; NASA

Table 1
Sources of Correlation

Wages, benefits, payroll taxes	Trade actions
Productivity	Regulatory changes/actions
Raw material costs	Low frequency high impact events of scale
Design development	Archaeology finds
Means & methods	
State of technology - common new technology/materials	
Uncertainty factors/known unknowns	
Budgeting and contingency management strategy and approach	
Packaging and contracting strategy	
Schedule precedencies	
Shared/common assumptions	
Failures/delays at interfaces	
MAIMS – “Money Allocated Is Money Spent”	
Parkinson’s Law – work expands to fill the time allotted	
Optimism and other biases consistently applied	
Location factors	

Correlation between pairs of projects in a program can be calculated as the sum of the products of the standard deviation for each common risk divided by the product of the standard deviation of all risks in each project. Selection of appropriate distributions for each risk is important. As the number of projects in the program grows calculation of overall program correlation becomes more challenging and more important. Ignoring dependency among cost components results in underestimation of total cost variance.

Other factors impacting project and program risk assessment

Several factors beyond correlation can impact the output and more importantly validity and conclusions from a probabilistic risk assessment. Some of these are outlined in the introduction and succinctly discussed here.

MAIMS – "Money Allocated Is Money Spent"

This is the financial analog of Parkinson's Law and is a major contributor to cost overruns or higher than necessary expenditures in the delivery of a program. One telltale sign that this effect is in full play is in multi-project programs where the final cost performance index is at 1.00 for a large number of the individual projects. This is not the result of "perfect" management but rather the willful consumption of any underrun that may have existed. The MAIMS principle effectively makes any potential savings from underruns unavailable to cover overruns elsewhere in the program.

Typical project cost analysis assumes an "ideal" project or program where savings on one element are made available to other elements. The presence of MAIMS in program or project contexts drive to an alternative strategy on establishing budgets and dynamically managing contingency and risk pools.

MAIMS acts to increase a probability distribution function's mean and reduce its standard deviation. (Effectively, values less than the allocated amount are assumed to be equal to the allocated budgets in the statistical assessment of total project or program costs.)

Optimism bias

People have a systematic bias towards overconfidence. Thus, many cost distribution approaches that rely on expert judgement to set several values (minimum, most likely, maximum for example) lead to distributions that are too tight and even weaker assessments on extreme values. Methods exist to reduce bias in assessing uncertain quantities but are not embraced in the engineering and construction industry[12].

Overly simplistic probabilistic cost analysis (PCA)

Numerous over-simplifications are often present in engineering and construction risk analysis. These include:

- Selection of distributions for various cost elements. In particular the use of triangular distributions which have been blamed for unrealistically low and high estimates. The triangular distribution leads to underestimates, and potentially significant underestimates because it has an upper bound.
- Omission of interrelationships among cost elements (correlation, discussed above)
- Process by which budgets are established and allocated (is MAIMS likely to be present)

---

[12] Direct Fractile Assessment Method; Alpert and Raiffa, 1982

- Approach to management of contingencies (levels and process for release)
- Confidence levels in underlying assumptions (available equals good; tough to ascertain is ignored)
- How low frequency, high impact events are considered
- Absence of adequate sensitivity analysis

Summary

Ignoring correlation results in an underestimation of total cost variance that grows more significant when we are dealing with a multi-project program or portfolio. Similarly, excluding correlation between variables in schedule estimation is significant. While there may be inadequate historical data to calculate correlation coefficients, ignoring correlation's presence is not an acceptable strategy. Sensitivity analyses for different levels of correlation may be performed and guidance is provided in the recommendations that follow.

Correlation is especially important when projects are concurrent but selection of the concurrency period should consider overall program or contract durations (especially in task order type contracts) or sequential contracts with added coupling hidden in complexity.

Recommendations:

- Treat correlation among cost elements and projects in a program realistically. Seldom are projects or programs "ideal" with no correlation between cost elements.
- Correlations in large complex engineering and construction programs of 0.3 to 0.6 are not atypical. Consider sensitivity to variation in assumed correlation levels. Rule of thumb - 0.4 is now almost the de facto correlation coefficient for cost estimates, 0.5 for schedule.
- Consider use of either the open-ended three-parameter Lognormal or Weibull distribution for cost analysis[13]. There are other acceptable distributions but emphasis should be placed on the values of input data (to minimize bias) and an open-ended distribution (triangular is not open-ended)
- Recognize that all statistical analysis represents a macroscopic view and that a complete assessment of risks and risk response also requires a microscopic view using other risk assessment methods such as decision trees. This microscopic view for low probability high consequence events helps address overconfidence and optimism biases. Ensure good quality information at the tails.

---

[13] Kujawski, Edouard & Alvaro, Mariana & Edwards, William. (2004)

- Recognize that the standard probable cost assessment assumes an “ideal” project and provides management with a false sense of confidence. Standard probable cost assessments are a source of major cost overruns even when high contingencies are deemed to have been provided.
- Use an appropriate number of elements in the work breakdown structure in a project and the number of projects in a program. Subdividing costs into too many smaller pieces leads to a false sense of security and an erroneous outcome.
- In large programs specifically incorporate the MAIMS principle in cost modeling and budget management practices.
- Recognize other human and organizational influences on project cost analysis and consider together rather than separately
- Recognize that systems thinking is essential when addressing correlation in large complex projects and programs.

References

Kujawski, Edouard & Alvaro, Mariana & Edwards, William. (2004). Incorporating psychological influences in probabilistic cost analysis. Systems Engineering. 7. 195 - 216. 10.1002/sys.20004.

S.A. Book, Why Correlations Matter in Cost Estimating, 32nd Annual DOD Cost Analysis Symposium, Williamsburg, 1999.

S.A. Book, Estimating Probable System Cost, Crosslink 2 (1) (Winter 2000/2001), 12-21

M. Alpert and H. Raiffa, "A progress report on the training of probability assessors," in Judgment under uncertainty: Heuristics and biases, D. Kahneman, P. Slovic, and A. Tversky (Editors), Cambridge University Press, 1982, pp. 294-305.

EPA/630/R-97/001, Guiding Principles for Monte Carlo Analysis; March 1997

F. Kuo, Everything you want to know about correlation but were afraid to ask; NASA

P. Bakhshi, A. Touran, A Method for Calculating Cost Correlation among Construction Projects in a Portfolio; International Journal of Architecture, Engineering and Construction Vol 1, No 3, September 2012

Ince, P. and Buongiono, J. (1991). “Multivariate stochastic simulation with subjective multivariate normal distribution.” Symposium on System Analysis in Forest Resources, Charleston, South Carolina, United States.

About the Author

Bob Prieto

Chairman & CEO
Strategic Program Management LLC
Jupiter, Florida, USA

Bob Prieto is a senior executive effective in shaping and executing business strategy and a recognized leader within the infrastructure, engineering and construction industries. Currently Bob heads his own management consulting practice, Strategic Program Management LLC. He previously served as a senior vice president of Fluor, one of the largest engineering and construction companies in the world. He focuses on the development and delivery of large, complex projects worldwide and consults with owners across all market sectors in the development of programmatic delivery strategies. He is author of nine books including “Strategic Program Management”, “The Giga Factor: Program Management in the Engineering and Construction Industry”, “Application of Life Cycle Analysis in the Capital Assets Industry”, “Capital Efficiency: Pull All the Levers” and, most recently, “Theory of Management of Large Complex Projects” published by the Construction Management Association of America (CMAA) as well as over 750 other papers and presentations.

Bob is an Independent Member of the Shareholder Committee of Mott MacDonald. He is a member of the ASCE Industry Leaders Council, National Academy of Construction, a Fellow of the Construction Management Association of America, Millennium Challenge Corporation Advisory Board and member of several university departmental and campus advisory boards. Bob served until 2006 as a U.S. presidential appointee to the Asia Pacific Economic Cooperation (APEC) Business Advisory Council (ABAC), working with U.S. and Asia-Pacific business leaders to shape the framework for trade and economic growth. He had previously served as both as Chairman of the Engineering and Construction Governors of the World Economic Forum and co-chair of the infrastructure task force formed after September 11th by the New York City Chamber of Commerce. Previously, he served as Chairman at Parsons Brinckerhoff (PB) and a non-executive director of Cardno (ASX)

Bob serves as an honorary global advisor for the *PM World Journal* and Library and can be contacted at rpstrategic@comcast.net.