

WORKING P A P E R

Quantitative Risk Analysis for Project Management

A Critical Review

LIONEL GALWAY

WR-112-RC

February 2004

This product is part of the RAND Corporation working paper series. RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.



PREFACE

This paper is the final report of the RAND Internal Research and Development (IR&D) project "Risk Management and Risk Analysis for Complex Projects: Developing a Research Agenda." The aim of the project was to survey how quantitative risk management and risk analysis methods were applied to the planning and execution of complex projects, particularly those which planned to utilize new and untried technologies. One recent RAND study indicated that such methods, while widely advocated, were not used to plan and manage a critical government satellite development project. This paper recommends several research areas in which RAND could contribute to evaluating the utility of these methods and improving their applicability.

This paper results from RAND's continuing program of self-sponsored independent research. Support for such research is provided, in part, by donors and by the independent research and development provisions of RAND's contracts for the operation of its U.S. Department of Defense federally funded research and development centers.

CONTENTS

Preface.....iii
Figures.....vii
Summary.....ix
Acknowledgments.....xi
1. INTRODUCTION.....1
 Risk, Risk Analysis, and Risk Management.....1
 Origin of Project.....3
 Goals and Methodology.....3
2. PROJECT RISK ANALYSIS.....5
 Three Elements of Project Risk Analysis.....5
 Historical Overview.....8
 Schedule Risk8
 Cost Risk11
 Performance Risk11
 U.S. Government Mandate for Risk Analysis12
 Empirical Retrospective Studies of Schedule and Cost Risk12
 Pedagogical Literature.....14
 Search for a Critical Literature.....15
 Conversations with Experts.....19
 State of the Art.....20
3. CRITICAL EXAMINATION OF PROJECT RISK ANALYSIS.....21
 Known Problems.....21
 Level of Aggregation of Tasks or Costs21
 Elicitation of Probabilities22
 Correlations23
 Feedback Effects24
 Is There Enough Data?25
 Summary26
 What is a Good Project Risk Analysis?.....26
 Answer #1: Accuracy26
 Answer #2: Aid to Structure Thinking27
 What Would a Critical Evaluation Look Like?.....28
 Barriers to a Critical Analysis.....29
4. CONCLUSIONS AND RECOMMENDATIONS.....31
 Is Quantitative Project Risk Assessment Useful or Even
 Feasible?31
 Need for a Critical Literature.....32
 Opportunities for RAND.....34
Bibliography.....37

FIGURES

2.1.	Nominal Schedule Risk for a Project	6
2.2.	Nominal Cost Risk for a Project	7

SUMMARY

INTRODUCTION

One of the major intellectual triumphs of the modern world is the transformation of risk, the possibility of untoward events, from a matter of fate to an area of study. *Risk analysis* is the process of assessing risks, while *risk management* uses risk analysis to devise management strategies to reduce or ameliorate risk. In project management, these techniques are used to address the questions "how long will this project eventually take?" (schedule risk), "how much will it finally cost?" (cost risk), and "will its product perform according to specifications?" (performance risk).

PROJECT RISK ANALYSIS

After extensive development beginning at the start of the 20th century, the methods of risk analysis recommended by the pedagogical literature are the stochastic Critical Path Method (CPM) for schedule risk, and a stochastic simulation of costs from the Work Breakdown Structure (WBS). Both techniques require a specification of uncertainty for time and cost for tasks to complete, followed by a Monte Carlo simulation for the total time and cost.

However, there is a striking lack of literature on the *use* of the techniques. This study conducted unstructured interviews with a number of researchers and practitioners. The universal statement about the general utility of quantitative project risk analysis was that it is clearly useful, because it is so widely used and so widely recommended. However, this was always followed by comments that project risk analysis is not well understood by project management. There was also agreement, confirmed by a literature search that virtually all of the evidence for its utility was anecdotal.

CRITICAL ANALYSIS OF PROJECT RISK MANAGEMENT

There is a set of issues, which need to be addressed in a critical evaluation of these techniques: what level of aggregation should be

used for the components of the schedule or cost? How should probability distributions be elicited? How to deal with correlations? How to take account of adaptive strategies? How to deal with limited information?

How do we judge a good risk analysis? If we are using the estimates to plan reserves or compare competing proposals, we require accuracy of the estimates. Alternatively, we could use the quantitative risk analysis framework (which requires measures of probabilities and impacts) primarily to force us to think hard about the project, whatever the final estimates say. If accuracy is the goal, a critical evaluation would be straightforward: collect information from projects, document cost and schedule estimates, and see how close they came to the final numbers. Evaluating the second criterion would require an ethnographic approach, entailing how insights from the analysis process affected management decisions.

CONCLUSIONS AND RECOMMENDATIONS

A program of critical evaluation in the open literature would help resolve these issues. How could RAND help? RAND has a reputation for doing work with organizations that might not trust each other with proprietary information but who do want an honest evaluation. The DoD and NASA should be interested in this research because it requires its contractors to do risk analysis, and bases decisions on the results. For example, NASA management could mandate project risk analysis for a selected group of projects and compare their results with a group that does not use the methods. There are also research issues in the areas of probability assessment and risk communication.

ACKNOWLEDGMENTS

This report has tried to synthesize current literature and practice in the area of project risk management in order to identify areas where future RAND research could be conducted. Much of the early research on cost and schedule risk was done at RAND, and many RAND colleagues are currently involved in a wide range of research that touches on these issues. Many of these colleagues gave me the benefit of their long and varied experience in these areas in conversations, recommendations for reading, other contacts to consult, in bringing new literature to my attention as they ran across it. In addition, several attended an internal seminar where I presented interim results and made valuable comments on that briefing. In alphabetical order they are: Jim Bigelow, Cynthia Cook, Paul Davis, Mel Eisman, John Friel, Dana Johnson, Mike Kennedy, Rob Leonard, Rosalind Lewis, Rich Mesic, Kip Miller, Isaac Porche, Jim Quinlivan, Dan Relles, Greg Ridgeway, Giles Smith, Fred Timson, David Vaughan, and Henry Willis. Former RAND colleague Ian Coulter corresponded via email.

In the course of this research I also interviewed by phone and email a number of external experts on project risk management. Thanks to Steven Book (MCR), Terry Williams (University of Strathclyde), Liam Sarsfield (NASA), Charles Bosler (Risk Services & Technologies), David Hilson (Project Management Professional Solutions Ltd.), Joe Hamaker (NASA), Michael Stamatelatos (NASA), Jay Kadane (Carnegie Mellon University), Lawrence Klementowski (Sekai Electronics), Henry Stefanou (Project Management Institute), Edmund Conrow (Management and Technology Associates), and Dr. Stephen Grey (Broadleaf Capital International).

Thanks to RAND reference librarian Amy Atchison for handling the technical aspects of the literature search, and to Richard Bancroft and Leroy Reyes of the classified library.

Finally, thanks to my colleague Tim Bonds for providing the original idea for this project, encouraging the author to submit it for an IR&D grant, and providing continual encouragement and help for the duration of the project.

As always, any inaccuracies or errors of fact or interpretation are the sole responsibility of the author.

1. INTRODUCTION

RISK, RISK ANALYSIS, AND RISK MANAGEMENT

A strong case has been made that one of the major intellectual triumphs of the modern world is the transformation of risk, the possibility of untoward events, from a matter of fate (essentially preordained and impossible to anticipate or mitigate) to an area of study, which can be anticipated, quantified, and dealt with, or at least ameliorated by good management.¹ This is exemplified in the development of insurance practices beginning in the 1700s and accelerating towards the end of the 20th century with the advent of sophisticated mathematical techniques such as probability and Bayesian statistics, the collection of large amounts of data, and the vast, almost unimaginable increases in computing power. Together these developments made it possible to consider quantifying risk, and then assessing the cost-effectiveness of mitigation efforts. Further, these techniques were increasingly applied not just in traditional areas of insurance such as life or fire protection, but to more complex risks in the environment, engineering design, and general management problems.

Before proceeding further, we define some of the key terms that will be used in the rest of this report.

A *risk* is an event, which is

- uncertain
- has a negative impact on some endeavor

For example, to a life insurance company the timing of deaths of its policyholders are risks. The company never knows precisely who among their insurees is going to die in a given period of time (uncertainty) and each death costs them a payout equal to the face value of the policy (negative impact on profitability).

Risk analysis is the process of quantitatively or qualitatively assessing risks. This involves an estimation of both the uncertainty of

¹ Bernstein, 1998.

the risk and of its impact. Again, an insurance company can estimate the number of deaths in a given period based on demographic information about their insurees; this estimate, coupled with information about their policies, in turn allows them to estimate the amount of money they will have to pay off in the time period in question. In general, these estimates will not match the exact amount of money paid out, but a key part of the uncertainty analysis will allow the insurance company have an idea of how likely different payoffs are in a range around their estimate.

Risk management is the practice of using risk analysis to devise management strategies to reduce or ameliorate risk. In order to deal with an estimated payoff, the insurance company may revise its investment strategy, change eligibility for insurance, target different populations for sales of policies, or even cancel policies if possible to control the amount of money they expect to pay out and insure that they make a profit.

As noted above, these ideas and methods of risk analysis and risk management have moved into many other areas...For example, in engineering design reliability estimates of different parts are combined with an assessment of the impact on system performance of the failure of the parts. This analysis has in turn been used to direct resources for modification and redesign to those areas of aircraft, nuclear reactors, and other complex man-made systems where improvements have the most effect on reducing potential failures. Success in this area has led to expanding the practice of assessing and managing risks to economies and eco-systems. In this report we will be concerned with the use of these techniques in managing complex projects, where some of the important questions are "how long will this project eventually take?", "how much will it finally cost?", and "will its product perform according to specifications?"

Before a project begins and while it is in progress none of these questions can be answered with certainty and project managers and customers are concerned with both how uncertain the answers are and what the potential impact of deviations may be. Risk analysis and risk management techniques are designed to answer just these questions.

ORIGIN OF PROJECT

This project had its origin in events surrounding a RAND review of a major U.S. government space program. The program had had a number of technical and managerial difficulties and RAND was asked to help evaluate the project by sitting in on the review briefings and independently evaluating the written materials provided. The author of this report was asked to comment on the risk analysis documented by the contractors to demonstrate that the revised project would meet new schedule and cost targets. After conducting an examination of several of the summary and detail briefings, it was puzzling that the prime contractor did not appear to have used any of the techniques for evaluating schedule and cost risk from the extensive pedagogic literature on risk analysis (although at least one of the subcontractors did use these methods).²

GOALS AND METHODOLOGY

The contrast between practice in the project under review and the pedagogic literature led to a more detailed review of the pedagogic literature on risk analysis and project management. This second review raised questions as to whether these methods, although widely advocated, would in fact be useful for complex high-technology projects such the one under review, which are very complex and require the contractors to push the technological envelope in several diverse areas simultaneously.

With the encouragement of the RAND project leader for the review, this project was proposed as a RAND IR&D (Internal Research and Development) project. As the project proposal put the central question:

It is striking that even elementary textbooks on risk analysis devote chapters to the use of such techniques in project management; are there reasons why these were not used in the space project? Are they systematically underutilized in such contexts in both defense and non-defense work? If so, why?

² See Bedford and Cooke, 2001, or Vose, 2000, for current versions of this literature. We use the term pedagogic literature to mean textbooks on risk analysis designed for students in business, operations research, and similar fields.

The methodology proposed to answer this question was to do a literature survey of methods used for risk analysis and risk management of complex projects to see which methods were found to be useful in practice.

The study was also planned to include a case study (if feasible) of the use of quantitative risk assessment techniques in the management of an actual project, whether successful or not. In the course of this project, we contacted several companies and government agencies and laboratories to explore the possibility of interviewing managers on a successful or unsuccessful use of quantitative management techniques. Unfortunately, while managers and practitioners talk in generalities about the benefits of quantitative risk analysis in project management, no one was comfortable discussing actual projects in detail despite our attempts to get top management support for the study.

2. PROJECT RISK ANALYSIS

THREE ELEMENTS OF PROJECT RISK ANALYSIS

There are three basic concerns in project management:

1. Schedule. Will the project go over schedule?
2. Cost. Will the project overrun its budget?
3. Performance. Will the output satisfy the goal(s) of the project?

At the start and up until the end of a project, the answer to each of these questions is unknown, and a yes answer to any or all of the questions is taken to be an undesirable consequence.³ So by the definition of risk in the previous chapter each of these elements should be subjected to a risk analysis (preferably quantitative) that will help project managers decide whether the project is in jeopardy of not meeting its commitments and whether or not to take action to mitigate the risk.⁴

What is schedule risk? It is the probability that a project will overrun its schedule. Conceptually we would like to see an analysis such as the following nominal graph of schedule risk for an imaginary project:

³ There is an obvious scale issue here: one day or one dollar over is not a problem. We'll generally ignore this issue; the methods discussed below give an indication of how big these overruns will be with the implicit assumption that a decision maker then transforms that based on his/her utility function.

⁴ In Chapter 1 it was stated that the two pieces of quantitative risk assessments were (1) to estimate the probability of an untoward event and (2) to estimate its consequences. In formal decision analysis, the consequences are measured in terms of utility to the decisionmaker, not in actual physical units (dollars, months). This is because the consequences are rarely linear in those units (a five million dollar overrun may be much more than five times as distressing as a one million dollar overrun because of reporting requirements, oversight, etc. However, in discussing project risk the physical scales are almost universally used, presumably because the scales are considered to be easily interpretable. The decisionmaker can therefore overlay his or her utility mentally on the products of the risk analysis. See DeGroot, 1970 for a clear exposition of utility in decision analysis.

Schedule Risk for Project X

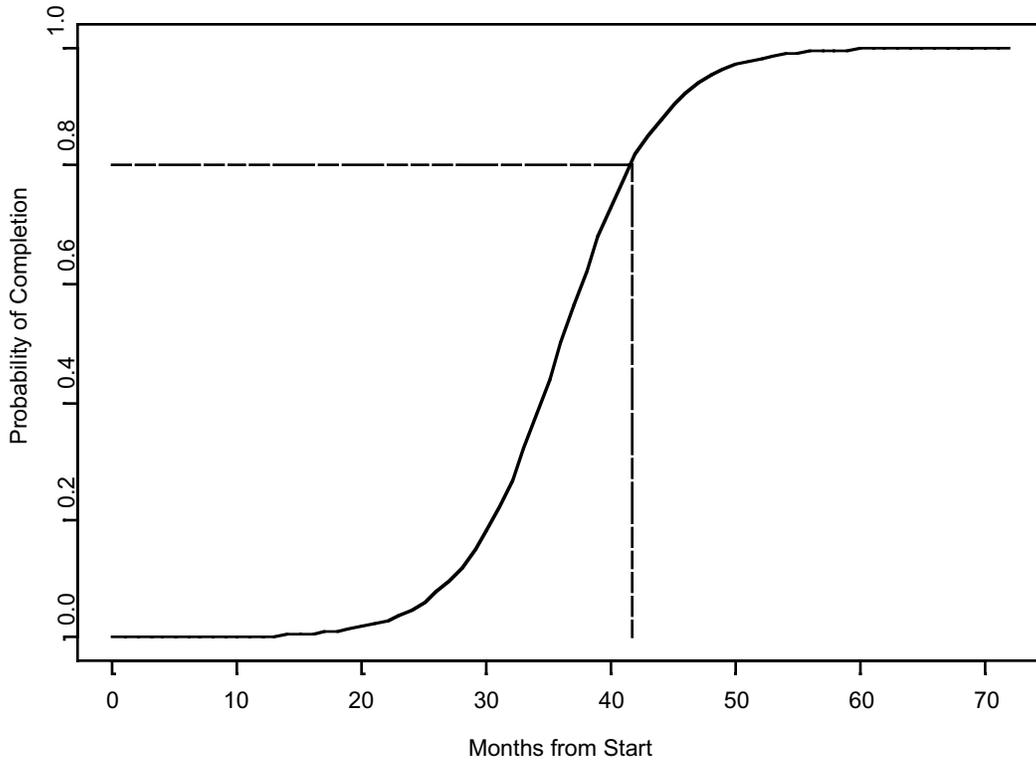


Figure 2.1--Nominal Schedule Risk for a Project

If the deadline for the project is 42 months this analysis indicates that the probability of completing the project in that time is 80%. There is only a 50% chance of completion in 35 months, and we are virtually certain that the project will be completed in 70 months.⁵

For cost we would like to see a similar curve, with the difference being that on the x-axis we would have cost at completion rather than months to completion:

⁵ This is a cumulative density function (cdf) for the time to completion. The same result could be shown with a probability density curve, which gives the probability that the length of the project lies within a specified segment, but given our interest in statements of the form "completion in no more than x months", the cdf is more convenient.

Cost Risk for Project X

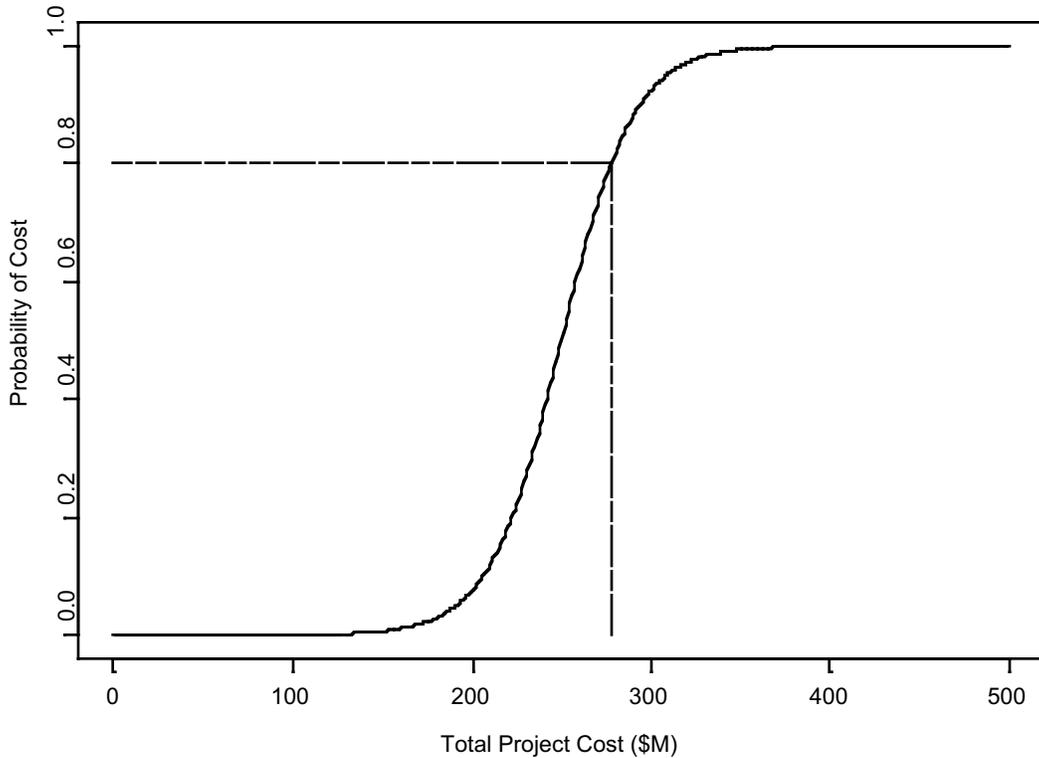


Figure 2.2 Nominal Cost Risk for a Project

We expect these curves to be different at different stages of the project; for example during the planning stages both curves might have a less steep slope, indicating that there is considerable uncertainty about the length of the project or its total cost, i.e., there is larger probability of going past 70 months in duration. As the project progresses the curves may shift in either direction as events happen that increase or decrease schedule or costs. And, as the project nears completion, the curve should become steeper as we become more certain about the final completion date and cost. These curves encapsulate what we want in a risk assessment of cost and schedule.⁶ Generating these

⁶ For examples in the literature where these curves are advocated, see e.g., Bedford and Cooke, 2001, and Glennan et al., 1993, and Raymond, 1999.

curves in a rigorous and credible way is not trivial for large and complex multiyear projects.

HISTORICAL OVERVIEW⁷

Large and complex projects have always needed a substantial management structure to insure that workers and materials came together in an organized fashion to achieve the tasks at hand. However, up until the 19th century such projects, while technically sophisticated by the standards of the time, either had long time horizons or did not require the solution of numerous engineering challenges in different substantive areas. However the many engineering projects of the late 19th century, such as high rise buildings, large canals and railways (often through challenging terrain) required more sophisticated techniques to keep track of the many different tasks which were required to be done in parallel.

Schedule Risk

The first such quantitative technique of modern project management in the area of schedule risk analysis was the Gantt chart, developed by Henry Gantt in 1917. It provided a graphical summary of the progress of a number of project segments by listing each segment vertically on a sheet of paper, representing the start and duration of each task by a horizontal line along a time scale, and then representing the current time by a vertical line moving from left to right. It is then easy to see where each task should be, and to show its current status.

The Gantt chart does have a serious drawback in managing complex projects: it does not easily show the interrelationship of tasks. In complex projects, many tasks have precedence requirements, i.e., they require other tasks to be substantially or fully completed before they themselves can be begun. Showing these relationships, especially as the number of tasks becomes larger, is no longer feasible on a Gantt chart.

⁷ Much of this is taken from Morris, 1994. The brief overview given here necessarily ignores many of the different approaches to quantitative project management that were explored in the post-World War II era. The aim here is to trace the evolution of the approaches to project risk analysis that are the accepted techniques in use today.

Instead, computers must be used to set up and maintain the network of tasks; this advance awaited the post-World War II development and widespread deployment of computing power.

The first project to avail itself of these resources was the U.S. Navy's Polaris program, which began in the mid-1950s to develop nuclear submarines, which could launch nuclear-tipped ICBMs. The Polaris Special Projects Office (SPO) was under the command of Vice Admiral William F. Raborn, who directed his staff to survey the project management techniques available in American industry to manage technologically complex programs. They found little. Raborn directed a small group of SPO staff and outside contractors to develop a useful control system for the Polaris project, and within a few weeks they developed the Program Evaluation Review Technique or PERT.

The basis of PERT was a detailed diagram of all anticipated tasks in a project, organized into a network, which represented the dependence of each task on the ones that needed to precede it. In addition, planners would estimate or *elicit* a probability distribution for the time each task would take from expert engineers. In early versions of PERT the experts were asked to give three estimates: pessimistic, optimistic, and most likely.⁸ With a number of other mathematical assumptions, it was then possible to derive and compute a probability distribution for the time to completion of the entire project.

PERT was a great success from a public relations point of view, although only a relatively small portion of the Polaris program was ever managed using the technique. And this success led to adaptations of PERT such as PERT/cost that attempted to address cost issues as well. While PERT was widely acclaimed by the business and defense communities in the 1960s, later studies raised doubts about whether PERT contributed much to the management success of the Polaris project. Many contended that its primary contribution was to deflect management interference by

⁸ Optimistic was often taken to be the minimum time and pessimistic the maximum time to do a task. Most likely was sometimes interpreted literally or as an average time for task completion.

the Navy and DoD by providing a "cover" of disciplined, quantitative, management carried out by modern methodologies.⁹

At about the same time that PERT was being invented, a similar planning and management technique was developed by DuPont. The Critical Path Method (CPM) also used a network representation, but initially did not try to estimate probability distributions for task durations.¹⁰ The non-stochastic nature of the network allowed for easier computation; it also facilitated the computation of the *critical path*, the set of tasks that drove the final project length. Various enhancements to CPM allowed the systematic exploration of alternative resource allocations to reduce this time, subject to cost constraints (whose assignments also were a matter of judgment).

The initial deterministic nature of CPM seems not to have been considered a drawback to its users. However, the increasing amount of computing power available led naturally to the inclusion of probability distributions for task durations in CPM. While the analytic simplicity of PERT was lost, rapidly increasing computer power allowed straightforward Monte Carlo simulation to be substituted for the PERT assumptions. The addition of stochastic task durations implies that tasks in turn are on the critical path with some probability, also estimated using the Monte Carlo results. With this development, the integral resource allocation enhancements apparently have been largely lost, at least in mainstream practice. Stochastic CPM is now the preferred methodology for assessing schedule risk in project management.¹¹

⁹ The classic study of the Polaris project is Sapolsky, 1972. His book treats the use (or lack thereof) of PERT within the project in detail.

¹⁰ Morris suggests that this was due to its use by DuPont in construction planning in which task durations were known with some accuracy.

¹¹ A further extension of PERT/CPM known as GERT, the Graphical Evaluation and Review Technique was developed by Pritsker at RAND in the mid-1960s as an outgrowth of NASA work (Morris, 1994). See also Pritsker, 1966 and Pritsker, 1979 for further extensions. Many of these generalizations have been subsumed into CPM.

Cost Risk

With the exception of the cost estimation and resource allocation optimization techniques noted above that once were embedded in CPM, most quantitative cost risk analysis has been done with techniques largely separate from those for schedule risk analysis.

The technique used for cost analysis of complex projects is based on the Work Breakdown Structure (WBS).¹² The WBS breaks a complex project down into components, services, facilities etc., with each succeeding level going to a finer level of detail.¹³ WBS cost estimation builds on the WBS by simply attaching a cost to each element and summing to a total. For a quantitative risk analysis in project planning, experts in relevant areas are asked to specify a probability distribution for each part of the WBS and then Monte Carlo simulation is used to estimate a probability distribution for the total project cost. As with CPM, the method is conceptually straightforward, although it does raise questions about the process of elicitation and possible correlations in costs for related components.

Performance Risk

Unlike schedule and cost risk analysis, where the methodologies are largely generic across all project types, methods of performance risk analysis tend to be much more tightly tied to subject area. Further, quantifying the relationships between different aspects of performance may be much more difficult. For example, ultimate performance of the space project reviewed by RAND will depend on software, power supply reliability, ground control facilities, and decision support systems.

There have been some efforts at constructing quantitative estimates of performance risk for aerospace systems, using physical relationships between performance parameters.¹⁴ However, current practice seems to be to use a mix of quantitative methods and models, where available, for

¹² When a project is in the stages of initial planning, cost estimating relationships are often used to get a rough estimate of costs based on hypothesized characteristics of the new project. See below for more information on this approach.

¹³ Morris, 1994, p. 44-45.

¹⁴ E.g., Timson, 1968, and Timson, 1970.

subsystems and then use a subjective judgment approach to estimate an aggregate risk for system performance risk.¹⁵ Performance risk will not be discussed further in this report, but to the extent that performance issues drive cost and schedule changes, the continued treatment of these risks in isolation for planning purposes may strongly affect the possibility of doing a good job of risk analysis in each area separately.

U.S. Government Mandate for Risk Analysis

There is a perception that the U.S. Government, particularly the Department of Defense, requires specific types of risk analysis for projects. It is true that after the success of PERT in the Polaris program DoD required the use of PERT for the management of projects; however, mandating the use of PERT specifically was fairly short-lived.¹⁶ Instead, later DoD acquisition regulations required only that risk analysis and risk management be used to help DoD manage risk. Specific *techniques* do not appear to be required at the Department level.¹⁷

Empirical Retrospective Studies of Schedule and Cost Risk

When used as planning tools, CPM/PERT and WBS-based estimates can be characterized as prospective, bottom-up techniques for estimating schedule and cost risks. An alternative strategy is to take a retrospective, top-down approach: review the history of past projects to find out how much they cost and how long they took, and compare these figures to budget and schedule estimates made at various earlier stages of the project, particularly the planning stages. The empirical

¹⁵ For a more modern approach to quantifying performance uncertainty, see Porche et al., forthcoming. For some examples of an approach that mixes quantitative and qualitative methods, see Bodilly, 1993a, and Bodilly, 1993b.

¹⁶ Klemenstowksi, 1978.

¹⁷ Driessnak et al., 2003 judge that the new interim version acquisition policy (which superseded DoD 5000.1 and 5000.2 in 2002) emphasize risk analysis and risk management because there are more references to the topics. But both they and Shepherd, 2003, writing in the same issue of *Acquisition Review Quarterly*, lament the lack of use of such methods in today's program offices.

relationship between estimates and actual time and cost can then be used to adjust the planning estimates to get figures that are hopefully closer to the final ones.

There are two variants on this approach. The first is to take an essentially descriptive approach: a number of projects are compared directly in terms of cost and schedule over- or under-run. Most commonly this is done essentially as a univariate analysis, with many different types of projects considered together, possibly broadly stratified by a characteristic such as platform type, total planned cost, etc. Often the aim is to test for a time trend. Much of this type of work was done at RAND in the 1950s-1970s,¹⁸ starting with defense projects and then branching out into other major infrastructure projects. There have been some recent additions to the literature,¹⁹ but at RAND this work effectively ended in the early 1990s.

The primary result of these studies is to find that most projects do in fact overrun on time and schedule, and may have unanticipated performance shortfalls, although typically the major cost of the schedule and budget slippage is to achieve performance goals that were not achievable with initial resource allocations. The amounts of slippage observed have not improved with the passage of time, and they suggest that the limits set for triggering a Nunn-McCurdy breach²⁰ may be too low. When covariates have been available and used in regression analyses they have not been too informative.

The alternative method is for analysts to assemble a set of projects along with their characteristics (size, technological maturity, management style, etc.) and then fit a regression-type equation to explain the final cost and time of the project using the program

¹⁸ See e.g., Marshall and Meckling, 1959, Perry et al., 1959, Perry et al., 1971, Merrow et al., 1979.

¹⁹ Conrow, 1995, Drezner et al., 1993, Glennan et al., 1993, the Spring 2003 issue of *Acquisition Review Quarterly*, published by the Defense Acquisition University.

²⁰ The Nunn-McCurdy amendment was originally part of the 1982 Defense Authorization Act, and called for the termination of developmental defense systems whose cost grew by more than 25%. There can be wavers and the actual details are somewhat complex, but Nunn-McCurdy is not toothless: it was used to terminate the Navy Area-Wide Missile Defense program in December 2001.

characteristics as covariates. For example, in a study of weather satellite projects, each satellite program would be characterized by its cost duration, technology, satellite weight, an indicator of initial technology, and so forth. This latter technique has a long history in cost studies, where the technique is called CER (for Cost Estimating Relationship) and is still in use today. However, it is not clear from the literature reviewed that these techniques are widely used in projecting costs for a specific project; instead their application seems to be in looking at technology trends or in very preliminary thinking about projects.²¹ Similar approaches could be used in analyzing time to completion, but this does not appear to be common practice; only the descriptive approach is used, although there are a few studies that mix the techniques.

All of these studies face serious challenges. The first is that retrospective information on project planning has been surprisingly hard to come by. Historical materials from project offices were not systematically preserved, especially in the early post-war years. When some historical information was available it was often hard to draw useful comparisons between programs due to a lack of comparable milestones. Even in the relatively structured DoD programs, cost and schedule estimates were missing for key milestone reviews, or there was uncertainty about the actual stage of the program at the review. Finally, characteristics of programs often changed substantially as a project proceeded, which could largely invalidate earlier estimates and the extent of these changes was undocumented in records that were preserved.

PEDAGOGICAL LITERATURE

The pedagogical literature on project risk management (largely in textbooks, but also including tutorials and training seminars sponsored by professional societies), as with most such work, focuses its attention on presenting recommended techniques for use today. There are a few textbooks on project risk analysis per se, but most of this material is available in books on general risk analysis, which typically

²¹ For a recent example, see Younossi et al., 2002.

contain a chapter on project risk, with the material sometimes divided into two separate chapters on schedule and cost risk.²²

The recommended approach to schedule risk analysis is stochastic CPM. Most texts contain a brief explanation of how to elicit probabilities, followed by an explanation of the logic of arranging tasks into networks, doing the Monte Carlo simulation and understanding the two major outputs, the distribution of time to project completion and the probabilistic nature of the presence of a task on the project's critical path. Most texts then present an example, sometimes drawn from practice, but usually simplified considerably and illustrate the computations and the outputs with the example.

Similarly, the discussion of risk analysis in the cost area depicts the WBS-based approach, again with discussion of elicitation, doing a Monte Carlo simulation, and interpreting the output. Examples are straightforward and often simplified from real problems.

The striking lack in the textbook literature is that there is little literature cited on the use of the techniques. That is, there are no pointers to a critical literature about the techniques such as when they are useful, or if there are any projects or project characteristics that would make it difficult to apply these methods. There are also few or no sets of case studies that would illustrate when the methods worked or failed.

SEARCH FOR A CRITICAL LITERATURE

The evaluation of a critical literature is, of course, one way to address the primary goal of this study: a critical literature on project risk analysis should contain case studies of a wide variety of project management situations which would allow us to address the question of whether these quantitative technique work well in the types of projects of interest: large, technologically-complex system developments. It should also provide some theoretical or methodological grounding as to what problems are likely to crop up. For example,

²² See, e.g., Cooper and Chapman, 1987, Vose, 2000, Bedford and Cooke, 2001.

probability elicitation for task duration or cost may be difficult for these projects, which could limit the applicability of the techniques.

Lacking pointers from the pedagogical literature, we conducted several bibliography searches of the business, management, and operations research literature.²³ These searches provided abstracts, which were scanned for the types of articles in which we were interested. We confined our searches to the 1970s and later because the earlier literature was dominated by articles on PERT that did not utilize the more flexible methods made possible with modern computing.

The articles and documents found could be divided into three groups:

- Theoretical articles addressing computing algorithms, analytic results for special cases, and similar issues.
- Tutorials describing the use of methods for particular industries usually using "toy" examples.
- A (very) few describing the use of the techniques for some particular problem. None of these were the type of projects of interest here.

This lack of a critical literature has not gone unnoticed in the field, nor is it new. At the Air Force Institute of Technology (AFIT) Klementowski took as his research project for his master's degree²⁴ an analysis of the usage of PERT/CPM and related techniques in the aerospace industry. He found only half of the organizations he studied used these techniques at all, and that intensity of usage varied.²⁵ As

²³ RAND reference librarians searched for articles about CPM and PERT using various combinations of phrases and keywords in a variety of databases that index business, operations research, and related methodological journals, books and working papers. These included EconLit, Business & Management Practices, Applied Science and Technology Abstracts, Periodicals Abstracts, Science and Technical Information Network, and the National Technical Information Service. Other web searches were carried out, as well as a search of the website of professional societies such as the Project Management Institute (PMI).

²⁴ Klementowski, 1978.

²⁵ Unfortunately, as acknowledged by the author, funding and time constraints did not allow him to do a formal sample. Instead, he depended on a mixture of personal contacts and suggestions from the people he interviewed to select organizations and people to interview.

part of his research he also reviewed the available literature and characterized it with the observations that "theory tends to dominate published material" and there is "little documentation of practical applications of PERT/CPM."

In one of the most recent papers Gill and Hilson²⁶ reported the results of a survey at a workshop of risk analysts on the role of quantitative risk analysis techniques in project risk management. The main finding was that the participants were not agreed as to when the various techniques should be used and whether they would provide useful information to project managers. As the authors put it

...there is no unambiguous set of criteria for determining when to use quantitative modeling techniques.

Hilson published a second paper²⁷ in which he noted one reason why there was no consensus:

Problems currently arise from the fact that existing evidence is ... *anecdotal* (instead of providing hard measurable data"

He goes on to encourage the profession to develop such a body of case studies and critical work, noting that without a body of evidence, the "undoubted" benefits of risk analysis in project management must "be taken on trust." Why should project risk analysis lack a critical literature? It is clearly mature enough, since it dates back to the 1960s. There has been a vast growth in complex projects, especially those that push the limits of different technologies.²⁸ And the exponential increase in computing power has removed the technical limits on the use of these techniques on very large and complex projects. All of these developments should lead to widespread use and detailed knowledge of applicability, *if* the techniques are of use.

However, there are several reasons that would tend to inhibit the development of a critical literature.

The results of the survey are therefore not generalizable to a larger population.

²⁶ Gill and Hilson, 1998.

²⁷ Hilson, 1998.

²⁸ Morris, 1994.

First, the community of practice is by nature fragmented. It includes academics in business and operations research, professional managers, risk analysis consultants and software developers. This fragmentation means that the project management literature is also split up, spread among operations research journals, business and trade publications, the risk analysis literature, and the burgeoning project management literature. The publication venues range from established, long-running academic journals, through interdisciplinary, cross-field "startups", to conference proceedings.

Second, the evolution of projects as they face challenges and are modified means that the evaluation of a risk analysis can be hard to define precisely. An early schedule projection may be perfectly competent, but a change in requirements or a mid-course modification in technology would make the original estimate "inaccurate".²⁹ This leads to the more fundamental question of just how one would evaluate a risk analysis. What are the criteria for "good"? Which analyses should be evaluated?

Third, details of business projects and management methodologies and data are often considered proprietary.³⁰ The experience of what works and what doesn't in project management is often considered to be a key competitive advantage, especially in high-technology areas. In addition, national security projects may impose levels of classification on project details that effectively prohibit meaningful comparisons.

Fourth, a critical analysis of project management experience would show *who* has succeeded and who has not, in quantitative terms. As long as there is general agreement that a risk analysis needs to be done whether or not there is evidence about what it does or how well it works, there is little incentive to assemble such evidence.

Finally, documenting and preparing a report on the risk analysis and its results requires resources above and beyond those required for the project proper, resources which will be used for the "greater good" of the project risk analysis community. There will be little or no

²⁹ This was one of the difficulties noted by the retrospective studies on cost and schedule overruns cited above.

³⁰ Hillson, 1998, noted this as well.

direct payback to the authors or their parent organization. In many cases, funding essentially is gone when the project reaches the stage where it can do this kind of report.

CONVERSATIONS WITH EXPERTS

In the absence of a critical literature, I conducted some unstructured interviews with a number of researchers and practitioners in the field of project management risk analysis. These included academics, consultants, government officials, and members of the risk analysis special interest group (SIG) of the Project Management Institute, the U.S. professional organization for project management techniques.³¹ The interviews inquired specifically about the utility of project risk analysis in the types of projects considered here.³²

The universal answer to the initial question about the general utility of quantitative project risk analysis was that it is clearly useful, because it is so widely used and so widely recommended. However, this was always followed by comments (echoed in the professional and tutorial literature) that project risk analysis is not well-understood, not well-integrated into project management (except in exceptional cases), not easily explainable to senior decision-makers, etc. When pressed further, most agreed that *customer* demand that a risk analysis be performed was one of most common reasons for doing it, especially when working for the government. There was also general agreement that most if not all of the evidence for its utility was anecdotal, i.e., that there was little empirical evidence of how useful quantitative techniques were.³³

There was also some disagreement as to whether qualitative or quantitative techniques were the most important. Hillson acknowledged that the profession was in fact divided on the issue.³⁴ One respondent said that in his experience, at the early stages of a project the

³¹ www.pmi.org.

³² Interviewees (about 15 total) included academics in business schools, private risk analysis consultants, risk analysis software developers, and senior U.S. government engineers.

³³ The Hillson, 1998 reference is one of the few explicit statements of this in print. See also Driessnack et al., 2003.

³⁴ Hillson, 1998.

usefulness of a project risk analysis was inversely proportional to the number of decimal places. This person was in a firm that wrote risk assessment software to help organize overall risk analyses, and they had removed a Monte Carlo capability because they thought the added complexity added little or nothing to the usefulness of the product. On the other hand, another respondent, a private risk analysis consultant, insisted that he had never found a qualitative analysis that gave the insight to a schedule of a simple quantitative analysis.

STATE OF THE ART

This chapter has drawn together a brief history of quantitative project risk analysis, together with a description of the current techniques and review of the empirical justification for their use. For schedules the current method of choice is stochastic (Monte Carlo) CPM as we have defined it here. Cost risk analysis is based on a Monte Carlo simulation from distributions of cost for elements of a WBS. While intuitively appealing and logical, there appears to be little published empirical evidence on the actual utility of these techniques, especially in the management of high-technology projects.³⁵ While professional practitioners testify to its usefulness, and the use of the techniques are required by customers (especially the U.S. DoD), it would be desirable to have more objective evidence as to the techniques' utility. In the absence of such evidence, in the next chapter we turn to a discussion of just what an empirical justification for these techniques would consist of.

³⁵ As noted, the method *is* reportedly widely used in construction.

3. CRITICAL EXAMINATION OF PROJECT RISK ANALYSIS

Given that there is little or no critical discussion on project risk analysis in the professional literature, in this chapter we discuss some known issues which should be treated in a critical evaluation of the subject, followed by a discussion about how such an evaluation would be carried out.

KNOWN PROBLEMS

Even in the absence of critical work on the subject, there are a set of problematic issues some of which have been enumerated by authors of the pedagogical literature and some that are suggested by work in other fields. Difficulties in any of these areas would affect the ability of practitioners to carry out a project cost or schedule risk analysis and so a critical study of the methods would need to address these issues.

Level of Aggregation of Tasks or Costs

A first obvious question is what level of aggregation should be used for the components of the schedule or cost. One could imagine using the highest level available, that of the whole project. This is obviously not credible: when asked how long it would take to complete a complex project or how much it would cost, risk analysts would be under severe pressure to produce a number that fits the client's schedule or budget. If the estimate did *not* fit, the analyst would be hard pressed to justify the estimate or to answer the next question, which would ask what adjustments could be made to modify the project to fit the constraints. Alternatively, too fine-grained a disaggregation of tasks or costs would be enormously burdensome to deal with as experts attempted to determine probability distributions for very small tasks or cost distributions for tiny pieces of hardware.

Cost estimators recommend going down to various levels of the WBS, depending on the stage of the project,³⁶ but in schedule risk analysis there is no similar structure that defines level of detail: deciding what tasks go into the schedule network is a matter of experience. Added to these issues of aggregation is the simple observation that in a complex project, which is technologically immature, it may be impossible to specify individual tasks at the beginning, let alone their interrelationships or all of the cost elements.

Elicitation of Probabilities

Assuming that the task network and the WBS have been assembled at the appropriate level of aggregation, the next task is to determine the probability distributions of the times to completion of each task and the probability distribution of each cost element. One source of such information is from previous experience. However, such data is not available in the open literature, and, if proprietary, may be limited by the type of projects done by a particular company or consulting firm. And for complex projects that are at the cutting edge of technology there simply may not be much data that is directly relevant. In this case the pedagogical literature in general recommends that expert opinion be consulted to elicit the required probability distributions for task durations and task and component costs.³⁷

In the pedagogical literature the actual process of elicitation is not usually treated in detail, other than stating the basic properties of probability distributions, recommending some particular distributions such as the triangle or the Weibull for practical use,³⁸ and, in some cases, drawing insights from some of the psychological literature insights about the biases most people exhibit when trying to specify probability distributions.³⁹ For specifics, most books suggest using

³⁶ One recent tutorial presentation suggested that no more than 30 total independent WBS items be used (Black, 2000).

³⁷ An exception is Conrow, who is skeptical of the ability of experts to contribute quantitative risk information of sufficient accuracy to be of use. Conrow, 2000.

³⁸ However, some authors have pointed out problems with using accepted distributions such as the triangle, e.g., Moran, 1999.

³⁹ See, e.g., Kahneman et al., 1982, and Bell et al., 1988.

the triangle distribution,⁴⁰ or they use a small number of references in the engineering literature on probability elicitation.⁴¹ Most of these techniques involve either eliciting from the expert the most likely, maximum and minimum costs or time durations, or the estimation of a mean and variance.

Somewhat oddly, none of the pedagogical literature references the statistical literature. One of the major foundational paradigms of statistics is the Bayesian paradigm which takes probability to be a measure of subjective information; as a corollary, the elicitation of the probability distributions on unknown quantities such as task duration or component cost is a key element of this version of statistical inference. In consequence, probability elicitation has received substantial attention from the Bayesian statistical community.⁴² For example, while the risk analysis literature focuses on matching a few values to a standard distribution, the statistical literature includes methods for cross-checking elicited distributions by feeding back the implications of the current distribution to the subject, and allowing the subject to then modify the distribution based on this information to better reflect their beliefs. However, the project management literature has few or no references to this work.

Correlations

As several authors in the literature have noted,⁴³ durations for some related tasks and costs for some related hardware might be correlated because they are affected by common external events or by some common characteristics. Virtually all of the advanced texts on project risk analysis recommend that correlations be taken into account when specifying probability distributions for quantities that may be

⁴⁰ Often referring to Morgan and Henrion, 1990.

⁴¹ See, e.g., Williams, 1992.

⁴² For a review by a prominent Bayesian statistician who has been instrumental (with his students) in developing and using practical elicitation methods, see Kadane and Wolfson, 1998. A book-length treatment of these techniques is Meyer and Booker, 2001. Note that this literature extends back to the early 1980s (see Kadane, 1996, and Kadane et al., 1980).

⁴³ E.g., Vose, 2000, and Book, 2000.

correlated. This amounts to eliciting not a univariate probability distribution, but a multivariate one.

However, this latter task is known to be much more difficult for human subject to do without using various conditioning arrangements (asking the expert to develop a probability distribution for one random variable while fixing the value of others),⁴⁴ and most of the project risk analysis literature does not recommend much more than taking correlation into account.⁴⁵ At most, it may recommend eliciting correlations as well, although with no references given.

This problem is compounded by the fact that correlations between random variables which can only take positive values (as duration and cost are) are not unconstrained. Unlike normal distributions, where the correlations between an unlimited set of normal random variables can be specified arbitrarily, subject to an overall condition of positive-definite correlation matrix, the constraints on correlated positive random variables are stronger and far less intuitive.⁴⁶

Given that correlations between task durations and cost elements should be assessed (and Book, 2000 make a very strong argument for this) the difficulty of doing so is an obstacle to applying quantitative risk methods.

Feedback Effects

Although much of the empirical research on cost and schedule changes have documented overruns in both categories (largely attributed to specification changes), it has also been noted in the risk analysis literature that managers have adaptive strategies available to them in the course of a project. For example, they can reallocate resources such as manpower to shorten timelines for some tasks if the project begins to fall behind. This kind of adjustment is yet another reason

⁴⁴ Kadane, 1996.

⁴⁵ An exception is Book, 2000, who sets forth empirical recommendations based on experience at Aerospace Corp on cost analysis.

⁴⁶ Arnold and Press, 1989, and Book, undated. Book makes the additional point that @RISK, an Excel add-on from Microsoft, claims to allow the user to specify correlations between events, but (at his writing) did not disclose the method used. Book noted that one likely method did not satisfy the assumptions cost risk analysts used.

for distrusting initial risk analyses as having predictive validity: if schedules begin to slip or costs rise there are feedback mechanisms that can act to bring the schedule and/or cost back to compliance. At least one author⁴⁷ has suggested that the process of modeling a complex project should include this kind of feedback to account for possible adaptations in the course of a project's life. Williams suggests modeling project managers' behavior with intelligent agents that will formulate and implement adaptive strategies during simulated project histories to get a better (i.e., more accurate) cost and schedule projection.

While Williams is pursuing research on this topic, there are reasons to doubt that this approach will work. If the job of specifying task networks and constructing a sufficiently detailed WBS are daunting, the job of trying to anticipate adaptations seems exponentially harder: trying to build in to such an agent subject matter expertise (especially in advanced technology areas where uncertainties abound) as well as project management skill seems intractable now. (However, to the extent that a stochastic CPM analysis can focus attention on tasks that have a high probability of being on the critical path, some ideas for adapting could be preformulated and then played in a simulation.)

Is There Enough Data?

As noted above, one expert in this field had the opinion that the value of an early risk analysis was inversely proportional to the number of decimal places. One argument against quantitative methods in advanced-technology projects is that there simply is not enough information with which to make judgments about time and cost. There may not even be enough information to specify the tasks and components. In these cases, even elicitations of expert opinions may not be credible or precise enough to carry out an informative analysis.⁴⁸

⁴⁷ Williams, 2002.

⁴⁸ Williams (Williams, 2002) has commented that often distributions built on individual elicitations are so wide that they are useless for planning purposes or to help award contracts. Conrow (Conrow, 2000) is also skeptical about replacing actual data with judgment.

Summary

Although the currently advocated methods of schedule and cost risk analysis are clear and logical in concept, there are many practical issues that raise questions about how they can be applied in practice, particularly in the areas of interest here: complex, technologically challenging projects such as those in aerospace. The lack of an empirical literature is especially felt here in trying to sort out which of these obstacles are really problematic.

WHAT IS A GOOD PROJECT RISK ANALYSIS?

Given the lack of a critical literature on project risk analysis, as well as the somewhat ambivalent comments by the experts interviewed, it seems worthwhile to step back a bit and ask a more fundamental question namely, what are the characteristics of a good project risk analysis? Another way to ask the question is to ask what the users want out of the analysis. There are two major answers to the latter question, although they are not explicitly asked in any of the professional literature reviewed for this report. However, in order to think about what a critical literature would look like we need to at least have a tentative answer.

Answer #1: Accuracy

The most obvious answer is that we require accuracy of the estimates. That is, given a specified final cost or project duration, we can calculate from the Monte Carlo simulation the probability of achieving that cost or duration (provided the project specifications do not change). If we use a criterion probability of schedule or cost for planning, say 80%, we can interpret the duration or cost at the 80% point in at least two different ways:

1. (Frequentist) Over many different projects, four out of five will cost less or be completed in less time than the specified cost or duration.
2. (Bayesian) We would be willing to bet at odds of 4 to 1 that the project will be under the 80% point in cost or duration.

We obviously require accuracy if we are using the estimates to plan reserves of time or money or if we are using them to compare competing proposals.⁴⁹ If the methods do not give us accurate answers in this sense they cannot be used.

Answer #2: Aid to Structure Thinking

There is an alternative view that has gained some currency because of the types of difficulties described above. Given these difficulties, we realize that our estimates will be off, and that our measures of uncertainty (quantiles, standard deviations) from the simulations understate the uncertainty in our estimates simply because there are unknowns that are not taken into account. Therefore, our quest to rate methods, analysts and contractors using estimates of schedule and cost, especially early in a project, is misguided. All estimates will be off, and trying to quantify these uncertainties will leave us with bounds, which are so wide as to be meaningless. And we have reason to believe that feedback effects will also act to mitigate schedule and cost slippage to some extent.⁵⁰

Therefore, according to this line of thought, we should use the quantitative risk analysis framework not for its explicit results, but because it forces us to

- Think hard about different aspects of the project
- Try to put numbers on probabilities and impacts
- Discuss (argue) with colleagues who have different ideas and perceptions

This view has some implicit support in the literature: writers talk about the side advantage of having to think things through carefully in the process of producing the estimates on which a risk analysis is based. However, there seems to be no one who then argues that the process is what is valuable, not the numbers. This view was

⁴⁹ Book, 2000 suggests using estimates in just this way, as does Roberts et al., 2003.

⁵⁰ Williams, 2002. One government cost analyst who was a trained engineer said that cost estimates were "almost always" greater than the 80% point of the simulated distribution.

argued explicitly in the operations research literature by Hodges,⁵¹ who argued that one valuable effect of the process of writing a computer model (planning the model, coding it, and assembling input data), was that it forced analysts to look carefully at the process they were trying to model. This could be a useful exercise even if the model was never used or was invalid or incorrect.

As noted in the previous chapter, there is virtually no literature that addresses the choice of either answer, although given the empirical literature on cost or schedule overruns we expect that the methods may not be accurate in the sense used above.

WHAT WOULD A CRITICAL EVALUATION LOOK LIKE?

Given these two different criteria for a good project risk analysis, what do they imply for the structure of a critical evaluation?

If accuracy were the goal, a critical evaluation would be straightforward: it would follow the structure of the empirical studies cited above in the descriptive literature of the 1950s-1980s. The procedure would be to collect information from a variety of projects, document cost and schedule estimates at defined stages of each project, and then see how close they came to the final cost and schedule. If the project scope changed substantially, new analyses would be required (and older ones possibly discarded). If this approach were taken, it would also be of interest to document and evaluate the accuracy of the subordinate steps, i.e., the cost and duration of individual tasks. The problems with this approach have been documented in the studies cited: lack of complete documentation of estimates and methods, ambiguity about the stage of project development corresponding to specific estimates, and changes in specification, among others.

If, however, the real benefit of project risk analysis is as an aid to structure critical thinking about the project and its management, evaluation would be much more complex. This type of evaluation would require an ethnographic approach, entailing the collection and examination of the risk analyses as well as the recording and examination of project management meetings, including the information

⁵¹ Hodges, 1991.

discussed and how decisions were arrived at. These examinations would try to tie specific risks discussed and acted on with the insights gained from performing the risk analyses. Evaluation would consist of metrics such as counting the number of risks that were identified and acted on a priori and which and how many risks were not foreseen. Accuracy could also be evaluated as a by-product.

BARRIERS TO A CRITICAL ANALYSIS

Taking either approach to critically analyzing the performance of quantitative risk analysis methods in project management would be expensive, although focusing on the value of risk analysis as an aid to critical thinking would be much more costly.⁵² However, the basic methodologies are well established.

There are several other obstacles that would need to be confronted in doing this kind of critical analysis. One is security concerns. National security projects would have classification issues, especially in looking at the details of projects if not their overall cost or schedule performance (although for some black programs even these numbers are classified and even the existence of the program may be classified).

Far more projects are likely to run into questions of proprietary knowledge with details of individual projects, including the perceived and actual capabilities of companies and their people to complete specific tasks.

This leads to even more fundamental considerations of politics and personal and professional reputation. Submitting to any retrospective or prospective evaluation of project risk management techniques and practices may reveal poor performance in either analysis or performance or both. This in turn could affect selection in future competitions for private companies and political problems for government agencies. The

⁵² This type of analysis would require not just a comparison of estimates and final values, but also the review and analysis of the discussions in key decision meetings, and repeated interviewing of decisionmakers to assess how they are interpreting and using risk analysis products. This would be rather labor intensive, although the insights obtained on risk perception by senior decisionmakers would be enormously valuable.

result is a natural reluctance to reveal details, which would only lead back to the current practice of generic lists of projects to which risk analysts made "substantial" contributions.

4. CONCLUSIONS AND RECOMMENDATIONS

IS QUANTITATIVE PROJECT RISK ASSESSMENT USEFUL OR EVEN FEASIBLE?

There has been extensive development of methods for quantitative project risk analysis over the past two decades. Starting from the need to coordinate and control large, technically complex projects, theoreticians and practitioners have developed several related methodologies to estimate cost and schedule for these projects and to attach uncertainty estimates to these numbers. The methods have had varying degrees of use in different areas: for example, CPM is in widespread use in the construction industry, where technological innovation is not the dominant feature of a project.

The picture in the area of high technology projects is more mixed. Despite the substantial history of the methods, it seems fair to say that practitioners and users are ambivalent about the usefulness and applicability of these techniques to these types of projects. While the Polaris program touted PERT as a breakthrough in project management, as noted above not even a majority of the tasks in the project were controlled with PERT.⁵³ Klementowski's thesis in the late 1970s, although limited for generalization by the sampling technique, showed less than a majority of organizations using CPM/PERT techniques. And the interviews conducted for this report revealed a similar ambivalence: respondents affirmed the usefulness of the techniques in general, but did not provide much in the way of particular examples.⁵⁴

In particular, there does not seem to be a consensus as to whether the techniques should provide accurate estimates of cost and duration or whether they are primarily frameworks for exploring uncertainty, with the quantitative aspects serving to force completeness and rigor of thought. Both arguments are used to justify the techniques' use. Accuracy is required if the estimates and their uncertainty measures are

⁵³ Morris, 1994, p. 31, Sapolsky, 1972.

⁵⁴ Much of this was attributed to the proprietary nature of the details and/or classification issues.

to be used for selection purposes and for planning. However, what retrospective empirical evidence there is from the 1950s to today indicate that most projects overrun in both dimensions. And there is no systematic investigation (in the open literature) that documents applicability of the second justification. (As noted, this type of evaluation would require significant resources to carry out.)

NEED FOR A CRITICAL LITERATURE

Given that there is a large community of risk analysis practitioners and academics that are convinced of the efficacy of the methods, why is a critical literature needed at all? There are two basic reasons.

First, the methods are not uniformly used even in the aerospace and defense industries or even within a single government organization such as NASA. This indicates a practical lack of consensus: for whatever reason, not all managers agree on their utility. A critical literature would help form this consensus: where and when do the methods work? Are they accurate in the traditional sense? Do they get better as projects proceed? A critical literature would provide a public basis of a "standard of practice" which is not available now.⁵⁵

Second, application of these methods is expensive and time-consuming. If they are not useful for whatever type of usefulness is required, then they should not be used or they should be modified so that they are useful.

The difficulty of developing a critical literature in quantitative project risk management brings to mind two other fields of practice: classical psychoanalysis and chiropractic treatment. Classical psychoanalysis (Freudian psychoanalysis and its derivatives) aims to treat mental and emotional problems by having a patient talk to a therapist, verbalizing ideas and thoughts as they occur, with the therapist responding as dictated by the branch of psychoanalysis they profess. Chiropractic treatment focuses on manipulation of a patient's spine to treat a variety of illnesses, from chronic back and limb pain to organic diseases.

⁵⁵ Hillson, 1998, Conrow, 2000 (preface), Driessnak et al., 2003.

In these two health areas as in project risk analysis, the methods advocated are not a priori implausible. Verbalizing thoughts might give insights into mental states, spinal manipulation might certainly help at least some back problems, and assessing risk, eliciting probability distributions, and doing an overall simulation is a straightforward way of addressing questions about uncertainty of cost and schedule. However, in each case it is essential to go beyond a plausibility argument to see if the proposed methods actually work. In the cases of psychoanalysis and chiropractic, proof of empirical efficacy is largely a matter of pointing out that since people seek the treatments, the treatments must be helpful. Janet Malcolm, in her book on psychoanalysts in New York City,⁵⁶ profiled one psychoanalyst who did research on analysis, and she discussed the unease and even distaste expressed by practicing analysts for his work (and noted that he was essentially unique in the New York psychoanalytic community). RAND colleague Ian Coulter, who studied chiropractic and other alternative medical systems, once observed to the author that wealthy chiropractors tended to found schools of chiropractic, whereas they would contribute more to the acceptance of the field by funding basic research.

None of the three areas lend themselves to easy experimentation. The classical clinical trial paradigm tests a drug against a placebo. Current clinical trial design is much more complex, but the three areas share special problems with issues like sampling, blinding (hiding treatment status), and placebos. However, evaluations have been done, most notably of chiropractic,⁵⁷ and this could be done for project risk management as well, particularly with the backing of the DoD and/or NASA.

A vigorous program of evaluation and a critical literature are no guarantee that best practices in any area will be accepted immediately or universally, but it is a basic necessity to have an informed discussion.

⁵⁶ Malcolm, 1982.

⁵⁷ E.g., Shekelle, 1998.

OPPORTUNITIES FOR RAND

This is clearly an area where research is needed. The lack of a body of empirical evidence, the often-cited reluctance of managers to use risk analysis techniques in project management, and the ambivalence of risk practitioners themselves over key issues such as applicability all call for a program of evaluation of these techniques and their application, especially in the area of complex, technologically advanced projects. The risk analysis community is rich in experience but that experience is primarily anecdotal; if project managers should be using risk analysis as a key management tool, as stipulated by customers and government organizations such as DoD and NASA, they should be convinced by empirical studies of the usefulness of the techniques.

One important part of the plan for this project was to formulate a agenda for research in project risk analysis based on the literature review and interviews, especially work that benefited from RAND's strengths and key competencies. There are several possibilities.

First, the lack of a critical literature, especially the evaluation of a set of case studies, has been noted here and elsewhere as a lack in this area. However, as noted at the end of chapter 3, there are many barriers to producing such a set of case studies in the open literature because of issues of classification, proprietary practices, and general sensitivity to retrospective critical review of project management decisions. However, practitioners all agree that such evaluation would be enormously useful. RAND's reputation for dealing with these concerns in both its defense and non-defense work makes RAND ideally suited to conducting such a study. Besides RAND's long experience in the defense area, centers such as the Institute for Civil Justice and RAND Health often deal with confidential and politically sensitive analyses of decisionmaking and its consequences. RAND has an important niche doing work with confidential information from organizations that might not trust each other with proprietary information but who do want an honest evaluation. Different commercial companies might be willing to provide detailed information on their project risk analysis successes and failures in return for protection of proprietary information and a

comprehensive, rigorous, objective evaluation of the effectiveness of these techniques.

Even if commercial companies are not interested enough to support this work, the DoD should be interested in this research because it requires its contractors to do this kind of analysis, which is costly, and because it makes decisions, at least putatively, based on the results of these analyses. It could facilitate the agreement of its contractors to provide the data from a variety of projects of differing sizes and complexity. While this would be a natural project to cut across DoD, the individual services may also have an interest in sponsoring this type of work.

One other government agency might be interested in this type of work: NASA. Several interviews in this project were with senior NASA engineers who were charged with overseeing various kinds of risk analyses. While NASA has pioneered various methods in risk analysis, the interviewees all indicated that the application of these methods to project risk was not uniform, at a minimum.⁵⁸ Under the current administrator NASA moved its emphasis from a limited number of very large, very expensive projects, to encouraging smaller, cheaper, high-risk projects that emphasized engineering innovation to solve traditional problems.⁵⁹ One aspect of those projects was to try to evaluate project risk, rather than simply applying resources to buy out any encountered or anticipated risk. While there has recently been a move away from this "smaller, cheaper, faster" approach (and the recent release of the report of the Columbia accident may spur further far-reaching changes at NASA), there is still a push to insert more project risk analysis into project management. If NASA project managers really do not think that these methods provide value to their management decisions, this sets up the conditions for a "clinical trial."

NASA management could mandate project risk analysis for a selected group of projects and compare their results with a group that does not use the methods. This could provide a clear picture of how much the

⁵⁸ Friel and colleagues noted this as well in unpublished RAND research (Friel, et al., 2002).

⁵⁹ Sarsfield, 1998, Muir and Simon, 1999, Sarsfield, 2000.

methods can really contribute, as well as an evaluation of accuracy, methods of risk communication, and other issues of interest to the risk community (see below). Further, the methods used could be controlled to be uniform or a limited set of variants, which would eliminate some of the variability inherent in an observational study of DoD contractors or private corporations that might use very different methods.

Setting up, conducting and evaluating such an experiment would be complex, and would probably require a commitment of years. But it could help NASA decide whether or not to embed these methods in its management practices with empirical evidence, while also providing a more general service to the field of project management as a whole in contributing to the scanty critical literature.

Finally, there are still research issues in the areas of probability assessment and risk communication. As noted in chapter 3, the crucial step of assessing probability distributions for task durations and task and component costs gets short shrift in the pedagogical literature of project risk analysis. Even in the Bayesian statistical literature, where elicitation of prior distributions is nominally a key step in an applied analysis, not much attention has been paid to the process. This may be because of the importance of human psychology and its difficult relationship with probability reasoning; much of the literature that has addressed this area deals with situations where humans do not do well with probabilities.

Even given the ability to do elicitation, and to carry through the computations demanded by the methodologies discussed here, there remains the problem of presenting risks to project management in meaningful ways. Again, the professional literature cites this as a problem but gives few if any solutions. Empirical studies seem to be non-existent. Research on methods for presenting risks and/or using them to make decisions that have been extensively empirically tested in the project management area would be very valuable.

BIBLIOGRAPHY

- Arnold, Barry C. and S. James Press, "Compatible Conditional Distributions," *Journal of the American Statistical Association*, 84, 152-156, 1989.
- Bedford, Tim, and Roger Cooke, *Probabilistic Risk Analysis: Foundations and Methods*, Cambridge University Press, Cambridge, UK, 2001.
- Bell, David E., Howard Raiffa, Amos Tversky, *Decision Making: Descriptive, Normative, and Prescriptive Interactions*, Cambridge University Press, Cambridge, UK, 1988.
- Bernstein, Peter L., *Against the Gods: The Remarkable Story of Risk*, John Wiley & Sons, New York, NY, USA, 1998.
- Black, Hollis, "Fundamentals of Assessing Risk: Quantifying Uncertainty in Cost Estimates with Modeling and Monte Carlo Simulation," Space Systems Cost Analysis Group (SSCAG) Meeting, Ogden, UT, February 17-18, 2000.
- Bodilly, Susan J., *Case Study of Risk Management in the USAF B-1B Bomber Program*, RAND, Santa Monica, CA, N-3616-AF, 1993a.
- Bodilly, Susan J., *Case Study of Risk Management in the USAF LANTIRN Program*, RAND, Santa Monica, CA, N-3617-AF, 1993b.
- Book, Stephen A., *Problems of Correlation in the Probabilistic Approach to Cost Analysis*, Technical Paper, The Aerospace Corporation, undated.
- Book, Stephen A., "Schedule Risk Analysis: Why It is Important and How to Do It," SCEA National Conference and Educational Workshop, Advanced Training Session, Phoenix, AZ, 11-14 June 2002.
- Book, Stephen A., "Estimating Probable System Cost," *Crosslink*, The Aerospace Corporation, Winter 2000/2001, p. 12.
- Book, Stephen A., "Schedule Risk Analysis: Why It is Important and How to Do It," SCEA National Conference and Educational Workshop, Advanced Training Session, Phoenix, AZ, 11-14 June 2002.
- Conrow, Edmund H., "Some Long-Term Issues and Impediments Affecting Military Systems Acquisition Reform," *Acquisition Review Quarterly*, Vol. 2, No. 3, pp 199-212, Summer 1995.
- Conrow, Edmund H., *Effective Risk Management: Some Keys to Success*, American Institute of Aeronautics and Astronautics, Reston, VA, 2000.
- Cooper, Dale, Chris Chapman, *Risk Analysis for Large Projects: Models, Methods, and Cases*, John Wiley, Chichester, 1987.

- DeGroot, Morris H., *Optimal Statistical Decisions*, McGraw-Hill, New York, NY, 1970.
- Driessnack, LTC John, Noel Dickover, Marie Smith, "Risk Community Building Inside the Program Management Community of Practice (PM COP)", *Acquisition Review Quarterly*, Defense Acquisition University, Washington, D.C., Spring 2003, pp. 107-114.
- Drezner, J., J. Jervaise, R. Hess, Hough, D. Norton, *An Analysis of Weapon System Cost Growth*, RAND, Santa Monica, CA, MR-291-AF, 1993.
- Friel, John, Noreen Clancy, James Dryden, Robert Hirsch, Liam Sarsfield, and Jerry Sollinger, RAND, Santa Monica, CA, 2002, unpublished.
- Gill, Harpreet, and David Hilson, "RISK SIG Workshop: The Role of Quantitative Risk Analysis," *Project*, Association for Project Management, pp. 26-27, July 1998.
- Glennan, Thomas K., Susan J. Bodilly, Frank A. Camm, Kenneth R. Mayer, Timothy Webb, *Barriers to Managing Risk in Large Scale Weapons System Development Programs*, RAND, Santa Monica, CA, MR-248-AF, 1993.
- Hilson, David, "Project Risk Management: Future Developments," *International Journal of Project and Business Risk Management*, 1998.
- Hodges, James S., *Six (or so) Things You Can Do With a Bad Model*, RAND, Santa Monica, CA, N-3381-RC, 1991.
- Kadane, Joseph B., ed., *Bayesian Methods and Ethics in a Clinical Trial Design*, Wiley, New York, NY, 1996.
- Kadane, Joseph B., J. Dickey, R. L. Winkler, W. Smith, and S. Peters, "Interactive Elicitation of Opinion for a Normal Linear Model," *Journal of the American Statistical Association*, 75, 845-854, 1980.
- Kadane, Joseph B. and Lara J. Wolfson, "Experiences in Elicitation," *The Statistician*, 47, pp. 3-19, 1998.
- Kahneman, Daniel, Paul Slovic, Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, UK, 1982.
- Klementowski, Lawrence J., *PERT/CPM and Supplementary Analytical Techniques: An Analysis of Aerospace Usage*, Master's Thesis, Air Force Institute of Technology, Wright-Patterson AFB, Dayton, OH, AFIT/GSM/SM/78S-11, 1978.
- Malcolm, Janet, *Psychoanalysis: The Impossible Profession*, Random House, New York, NY, 1982.

- Marshall, Andrew W. and W. H. Meckling, *Predictability of the Costs, Time, and Success of Development*, RAND, Santa Monica, CA, P-1821, 1959.
- Morrow, E. W., S. W. Chapel, C. Worthing, *A Review of Cost Estimation in New Technologies*, RAND, Santa Monica, CA, R-2481-DOE, 1979.
- Meyer, Mary A., Jane M. Booker, *Eliciting and Analyzing Expert Judgment: A Practical Guide*, Society for Industrial and Applied Mathematics-American Statistical Society, Alexandria, VA, 2001.
- Moran, Steve, "Cost-Risk Analysis: Ready - Fire - Aim???", Space Systems Cost Analysis Group (SSCAG) Meeting, 21-22 January 1999.
- Morgan, M. Granger, Max Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, Cambridge, UK, 1990.
- Morris, Peter W. G., *The Management of Projects*, Thomas Telford, London, UK, 1994.
- Muirhead, Brian K., William L. Simon, *High Velocity Leadership: The Mars Pathfinder Approach to Faster, Better, Cheaper*, HarperBusiness, New York, NY, 1999.
- Perry, R. L., D. DiSalvo, G.R. Hall, A. J. Harman, G. S. Levenson, G. K. Smith, and J. P. Stucker, *System Acquisition Experience*, RAND, Santa Monica, CA, RM-2072-PR, 1969.
- Perry, R. L., G. K. Smith, A. J. Harman, S. Henrichsen, *System Acquisition Strategies*, RAND, Santa Monica, CA, R-733-PR/ARPA, 1971.
- Porche, Isaac, Henry Willis, Martin Ruszkowski, *Framework for Quantifying Uncertainty in Electric Ship Design*, RAND, Santa Monica, CA, DB-407-ONR, forthcoming.
- Pritsker, A. Alan B., *GERT: Graphical Evaluation and Review Technique*, RAND, Santa Monica, CA, RM-4973-NASA, 1966.
- Pritsker, A. Alan B., *Modeling and analysis using Q-GERT networks*, Wiley, New York, NY, 1979.
- Raymond, Fred, "Quantify Risk to Manage Cost and Schedule," *Acquisition Review Quarterly*, Spring 1999, pp. 147-155.
- Roberts, Barney, Clayton Smith, and David Frost, "Risk-Based Decision Support Techniques for Programs and Projects," *Acquisition Review Quarterly*, Defense Acquisition University, Spring 2003.
- Sapolsky, Harvey M., *The Polaris System Development: Bureaucratic and Programmatic Success in Government*, Harvard University Press, Cambridge, MA, 1972.

Sarsfield, Liam, *The Cosmos on a Shoestring: Small Spacecraft for Space and Earth Science*, RAND, Santa Monica, CA, MR-864-OSTP, 1998.

Sarsfield, Liam, *The Application of Best Practices to Unmanned Spacecraft Development: An Exploration of Success and Failure in Recent Missions*, RAND, Santa Monica, CA, DB-319-NRO, 2000 (U.S. Government distribution only).

Shekelle, Paul G., "What role for chiropractic in health care?", *New England Journal of Medicine*, v. 339, no. 15, Oct 8, 1998.

Shepherd, Bill, "Managing Risk in a Program Office Environment," *Acquisition Review Quarterly*, pp. 124-139, Spring 1993.

Timson, Fred S., *Measurement of Technical Performance in Weapon System Development Programs: A Subjective Probability Approach*, RAND, Santa Monica, CA, RM-5207-ARPA, 1968.

Timson, Fred S., *Technical Uncertainty, Expected Contract Payoff, and Engineering Decisionmaking in a System Development Project*, RAND, Santa Monica, CA, RM-6044-ARPA, 1970.

Vose, David, *Risk Analysis: A Quantitative Guide*, John Wiley & Sons, New York, NY, USA, 2000.

Williams, Terry W., "Practical Use of Distributions in Network Analysis", *Journal of the Operational Research Society*, Vol. 43, No. 3, pp. 265-270, 1992.

Williams, Terry W., *Modeling Complex Projects*, Wiley, New York, NY, 2002.

Younossi, Obaid, Mark V. Arena, Richard M. Moore, Mark A. Lorell, Joanna Mason, John C. Graser, *Military Jet Engine Acquisition: Technology Basics and Cost-estimating Methodology*, RAND, Santa Monica, CA, MR-1596-AF, 2002.