ANALYSIS AND QUANTIFICATION OF RISK OF COST OVERRUNS IN CONSTRUCTION PROJECTS

by

José Bermúdez Marcos
Ingeniero de Caminos, Canales y Puertos
Universidad de Santander, Spain
(1980)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS OF THE
DEGREE OF

MASTER OF SCIENCE
IN CIVIL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 1985

Signature of Author ______________________________
Department of Civil Engineering
November 15, 1984

Certified by ______________________________
Erik Vanmarcke
Thesis Supervisor

Accepted by ______________________________
François Morel
Chairman, Departmental Committee on Graduate Students

ANALYSIS AND QUANTIFICATION OF RISK

OF COST OVERRUNS IN CONSTRUCTION PROJECTS

by

José Bermúdez Marcos

Submitted to the Department of Civil Engineering
on November 15, 1984 in partial fulfillment of
the requirements for the Degree of Master of Science in
Civil Engineering

ABSTRACT

Studies on public and private construction projects have shown very frequent and sometimes substantial cost overruns.

We analyze all the factors involved in the uncertainty of the project's final cost. The main goal is to provide a methodology which quantifies that uncertainty through the incorporation of all the information available to the analyst.

Existing methods to propagate uncertainty from underlying elementary variables to the final cost figure are systematized and described. Once the project-specific distribution of the final cost is obtained, historical experience with similar projects is combined through Bayesian updating. Moreover, the same procedure is applied at different points in time over the project life so as to know the risk evolution.

The focus here is on the owner (public agency or private entity) and all the usual contractual arrangements are covered. However, the same procedures, with little modification, are applicable to the constructor instead of the owner.

Thesis Supervisor: Dr. Erik Vanmarcke

Title: Professor of Civil Engineering

ACKNOWLEDGMENTS

I wish to express my gratitude to Prof. Erik Vanmarcke for his guidance and support. Recognition is also given to the Fulbright Program and the Banco de Bilbao, which made possible my stay at M.I.T.

This work is dedicated to my parents.

TABLE OF CONTENTS

Page

CHAPTER 5 INFLATION

CHAPTER 6 UPDATING AND EVOLUTION OF THE COST UNCERTAINTY

CHAPTER 7 CASE STUDY AND CONCLUSIONS

LIST OF FIGURES

Figure | Page

LIST OF TABLES

CHAPTER 1

PROBLEM STATEMENT

1.1 Cost Variation in Construction Industry Projects

In the construction industry, the project that is completed to its initial estimate of cost is the exception rather than the norm. Since construction projects usually last several years, and many factors (subscil conditions, weather, etc.) are not completely known, the resulting variation in prices, quantities and the like, prevents accurate cost estimation. It is likely, therefore, that the actual final cost will differ from the original (definitive) estimate. Furthermore, as the project progresses and new information becomes available, the expected actual final cost will change.

The history of a large variety of projects shows such a cost variation, which is usually (and not surprisingly) a cost overrun.

If true costs were known beforehand many projects would not have been undertaken, since they would not have passed the cost-benefit analysis. With high cost overrun many owners can be trapped by the "sunk costs" principle. That is to say, during the project execution, cost overrun making the whole project uneconomic may become obvious, and yet the logical decision is to proceed to completion. As F.L. Harrison (1976) says: "The money already spent must be considered as sunk, or written off, and the decision to proceed must be based on the project's rate of return on the remaining forecast costs to completion". Great economic undertakings, such as developments in transportation or water regulation, have produced important cost overruns. However, these projects have turned out to be successful enterprises. What happens is that benefits (mainly indirect and long term ones) may be more grossly underestimated than costs.

Another problem derived from cost overrun is that of financing the unexpected costs. The owner will have to seek new arrangements with financial institutions, sometimes with the embarrassment of a project running over the budget and perhaps not meeting schedule, and therefore of doubtful profitability.

Most of the source of cost overruns is due to misinformation at the initial stages. Information can be gained through additional studies and surveys at a certain cost, which may be offset by the subsequent cost overrun reduction. This is why the owner's team must estimate the optimum level of detail and completeness in the design.

Sometimes the initial overoptimism of the developers or program managers and their will to carry out the project may produce a "buy-in", which is an intentional underestimation of costs in order to obtain a contract or to obtain approval to proceed on an effort with the hope that follow-on contracts, changes, or additional funding will compensate for the original low estimate (Murphy et al. 1974). Similar effect may be produced when the overoptimism is unintentional. It follows that often a more reliable cost figure can be obtained by means of a cost estimate conducted by a truly independent organization. Sometimes the "independent" government cost estimates can hardly be considered independent because of the vested interests of the agencies involved.

As we shall see in a subsequent section, there are many causes of cost overrun. Some are controllable, such as the completeness of the engineering surveys or the project administration; others are uncontrollable, such as inflation and weather. Therefore, the uncertainty in the final cost can be reduced but not eliminated. According to the contractual arrangement chosen for the project, that uncertainty or risk can also be shifted from owner to constructor up to a certain degree, as we will explain.

1.2 Cost Overrun Experience

In small projects a cost overrun of 10-20% is relatively common. In larger, longer projects, mainly those with a high development content or considerable uncertainty at the earlier stages, the cost overrun may be striking. For example, cost overruns have been reported of 50% in petrochemical plants, 140% in North Sea oil projects, and 210% in nuclear power stations (Harrison 1981). The Suez Canal costed twenty times its 1838 estimate and three times its 1887 estimate

(Merewitz 1972). Similar experience can be found in other large undertakings.

In this section we will show statistics of cost overruns for different projects and also we will make reference to several regression models that attempt to relate the cost overrun to some characteristics of the project.

Extensive and systematic empirical analyses have been undertaken in the area of major weapon systems acquired by the U.S. Department of Defense. We will report the findings briefly for the sake of illustration in our analysis. They have been inventoried by Merrow et al. (1979) in a Rand Corporation study (R-2481-DOE).

The development of new weapons is characterized by a large amount of internal and external uncertainties. Internal (or technological) uncertainties relate to the possible incidence of unforeseen technical difficulties in any of the development stages of a specific weapon system. External uncertainties exist because of factors external to the individual project, such as changes in weaponry technology, strategic requirements or government policies, which may cause reorientation of the program or redesign of the system.

The following attributes differenciate the weapons developments from other types of projects:

-Noncompetitive features. Usually most weapons systems are unique enough so that the government selects the seller (constructor), who thus has some of the bargaining power of a monopolist.

- Presence of significant amounts of both technical and external uncertainty. There exists a large potential for unforeseen problems, for changes of scope and so on.
- More emphasis on maintaining development schedules and performance than on maintaining original costs, once contracts are signed. This is related to the contractor goal of survival and growth.

Marshall and Meckling (1959) studied a sample of 22 weapons systems and found, after adjusting for changes of scope and inflation, cost increases of 200 to 300%. The increase tended to be larger for systems incorporating many new ideas and major improvements in performance.

Summers (1967) developed a multiple regression model to explain the ratio of actual to estimated costs (cost factor) as a function of the following variables: a) timing of the estimate within the development program (fraction of project length); b) degree of technological advance required (based on expert judgement); c) actual project length and d) calendar year. The relation is clear: When there are major technological advances and the development time is long, the ratio is significantly greater than unity. Summers includes the timing of the estimate, because the later that a cost estimate is made, the more likely it will reflect more information about the project. The calendar year is included because he believes that estimating methodology improved since a certain year.

Harman (1970) prepared a regression model similar to that done by Summers. He found also that longer programs and larger technological advances (evaluated subjectively by experts) lead to higher cost factors. An underlying idea of his model is that there is an optimal project length. Deviations from it will increase cost.

Alexander and Nelson (1972) suggested a more objective index to measure technological advance in terms of expected time of arrival (TOA) of a demonstrated level of performance.

Several studies found that contract type is related to cost growth. Cost-plus-fixed-fee (CPFF) contracts, commonly used by the Department of Defense to acquire major weapons systems, provided no positive incentives for firms to be efficient, so that in 1962 there was a shift to incentive contracts of various types.

Important works about cost overrun experience have been done in the area of construction industry projects, with which we are more concerned. They are summarized by Merewitz (1972) and more extensively by Merrow et al. (1979) in a Rand Corporation study (R-2481-DOE).

Of the literature available, the best and most comprehensive work has been done on the cost experience of federal water-resource construction agencies.

The Corps of Engineers, in a 1951 study, revealed for 182 rivers, harbors, and flood control projects an overall ratio of actual cost to estimated cost (cost factor) of 2.24, prior to adjustment for inflation, and a ratio of 1.3, subsequent to adjustment for price level change. A later study showed a significantly improved record for the newly completed projects.

A Bureau of Reclamation report, analyzed by Altouney (1955), indicated for 103 projects completed or in construction up to 1955 an unadjusted ratio of 2.77 and an adjusted ratio of 1.96. Subsequent reports showed substantial improvement.

The Tennessee Valley Authority (TVA) has an unadjusted ratio of 1.22 for 25 dams completed between 1933 and 1966.

Hufschmidt and Gerin (1970) showed that agency performance, particularly that of the Corps of Engineers and the Bureau of Reclamation, has improved significantly in the post World War II period.

Table 1.1 shows the overrun experience of the three water resource agencies.

Merewitz (1972) has compiled and analyzed data on general public works cost experience. His figures, not adjusted for inflation, are presented in table 1.2. He concludes that cost overruns have been smaller in water and highway projects than the average public works

Table 1.1 Comparison of cost overrun experience in water-resource agencies

	Period of Record	No. of Projects	Actual/Estimated: Original Estimate	Actual/Estimated: Inflation Adjusted	Overrun Frequency (% of projects): Original Estimate	Overrun Frequency (% of projects): Inflation Adjusted
TVA	1933-66	25	1.217	-	45	-
Dams		9	.904	-	0	-
Steam plants		34	.947	-	34.4	-
Total, TVA						
Corps of Engineers						
1951 Report	pre-1951	182	2.241	1.306		-
1964 Report	1954-65	68	.998	.770	51	29
Bureau of Reclamation						
1955 Report	pre-1955	103	2.770	1.960	89	86
1960 Report	1935-60	79	1.366	1.136	67	52
1960 Report	1946-60	54	1.094	.959	52	35

Source: Merrow et al. (1979).

Table 1.2 Summary of public works cost overrun experience.

Type of Project	No. of Projects	Actual/Estimated Cost
Water Resources	49	1.39
Highway	49	1.26
Building	59	1.63
Rapid Transit	8	1.51
Ad Hoc (*)	15	2.11
Grand Mean	180	1.50

(*) Ad hoc projects are essentially first- or one-of-a-kind projects, such as the Albany Mall (Albany, New York), the World Trade Center (New York City), the Long Beach Queen Mary, the Trans Alaska Oil Pipeline, and the New Orleans Superdome.

Source: Merewitz (1972) and Merrow et al. (1979).

experience. The cost overrun experience for ad hoc, building, and perhaps rapid transit projects, has been higher than the average.

Tucker (1970) examined 107 civil works projects and obtained a regression equation for the cost growth as a function of the project length, the estimated project cost, the calendar year and the timing of the estimate within the project length. The equation indicates that longer, more costly projects have higher cost growth; over time, cost estimate growth has tended to become less of a problem; and estimates made later in the life of a project tended to be closer to actual costs. He found also that the cost overruns in buildings are the greatest and most predictible, while water and highway projects experienced lesser, yet more eccentric, cost growth.

Mead (1977) studied cost overruns in more recent, very large projects (table 1.3). He makes adjustments for unanticipated inflation and scope change, thus providing an indication of the potential importance of these factors. Mead's unadjusted weighted average ratio is greater than any generated by Merewitz in table 1.2, perhaps due to a bias arising from the selection of the projects for analysis (large and complex).

The occurrence of cost overruns (or cost underestimation errors) is not exclusive of civil works projects; indeed, it appears to take place in every public and private capital venture. Table 1.4 -from the Rand Corporation study R-2481-DOE - summarizes cost overrun experience, allowing comparison of different areas.

1.3 Causes of Cost Overrun

The following causes are mentioned most often in the literature as those that drive up costs in construction projects (Merrow et al. 1979):

—Deviations from estimate scope. Inadequate anticipation at the time of an early estimate of all the physical installations and essential aspects necessary to fulfill the purpose of a project and shifts in legal, administrative and

Table 1.3

Cost overrun in major construction projects completed between 1956 and 1977

(adjusted for unanticipated inflation and changes in project scope)

Project	Initial Estimate		Actual Result		Unadjusted Ratio of Final to Initial Cost	Ratio After Adjustment		Compound Annual Rate of Cost Overruns, After Adjustments (in percent)[a]
	Amount (millions)	Date	Amount (millions)	Date Completed		For Unanticipated Inflation	For Change in Scope of Project	
Bay Area Rapid Transit Authority	$996.0	1962	$1640.0	5/76	1.647	1.297	1.037	0.31
New Orleans Superdome	46.0	1967	178.0	7/75	3.870	3.219	3.219	15.73
Toledo Edison's Davis-Besse nuclear power plant, Ohio	305.7	1971	466.0	5/75	1.524	1.401	1.401	11.89
Trans-Alaska Oil Pipeline (Alyeska)	900.0[b]	1970	7700.0[c]	7/77	8.556[c]	6.926	4.250	22.96
Cooper Nuclear Station, Nebr. Pub. Power Dist.	184.0	1966	395.3	74	2.148	1.748	1.748	7.23
Rancho Seco Nuclear Unit No. 1, Sacramento	142.5	1967	347.0	74	2.435	2.026	1.239	3.11
Dulles Airport, Washington, D.C.	66.0[c]	1959	108.3[c]	62	1.641[c]	1.641[d]	1.486	14.10
Second Chesapeake Bay Bridge	96.6[c]	1968	120.1[c]	6/73	1.243[c]	1.104	1.104	2.00
Frying Pan Arkansas Project								
Ruedi Dam	12.8[c]	1962	22.9	72	1.789[c]	1.636	1.145	1.36
Sugar Loaf	6.1	1952	10.2	73	1.672[c]	1.500	1.500	3.75
Boustead Tunnel	9.2[c]	1962	21.2[c]	73	2.304[c]	2.078	1.233	1.92
Rayburn Office Building, Washington, D.C.	64.0[c]	1956	98.0[c]	6/66	1.531[c]	1.531[d]	1.342	2.99
Weighted average					3.93	3.21	2.21	10.07

SOURCE: Walter J. Mead, with George W. Rogers and Rugus Z. Smith, *Transporting Natural Gas from the Arctic*, American Enterprise Institute for Public Policy Research, Washington, D.C., 1977, pp. 88-89.

[a]The compound annual rate expression is used only as a convenient method of comparing initial cost estimates with the sum of all actual costs at the termination of the project. This device permits a comparison of overruns on several projects having different construction periods.

[b]In May 1974, the Alyeska Pipeline Service Co. re-estimated capital cost at $4 billion; then in October 1974, costs were again estimated at $6 billion for the completed pipeline. By June 1975, the estimate was raised to $6.375 billion. In 1969, the $900 million cost estimate for Alyeska assumed a capacity of 500 mb/d. The scope was changed to permit a capacity of 1.2 million b/d. The cost of this change in scope was $700 million, raising the initial capital cost estimate to $1.6 billion.

[c]Does not include interest.

[d]Observed inflation was less than anticipated.

Source: Merrow et al. (1979)

Table 1.4 Summary of cost overrun experience

Items Estimated	Mean of Actual to Estimated Cost	N	Standard Deviation
Weapons, 1950s	1.89	55	1.36
Weapons, 1960s	1.40	25	.39
Public works			
Highway	1.26	49	.63
Water projects	1.39	49	.70
Building	1.63	59	.83
Ad hoc	2.14	15	1.36
Major construction	2.18	12	1.59
Energy process plants (*)	2.53	10	.51

(*) Because so few energy process plants have been constructed, the mean and standard deviation were calculated on the basis of the inflation-adjusted ratio of the last available estimte to the first available estimate. This probably results in a significant understating of the mean for these plants.

Source: Merrow et al. (1979).

political conditions over time, are underlying causes of these changes. We can distinguish:

*<u>Scope changes</u>. These are changes in conceived project size and characteristics between the original estimate and final design and construction. Scope changes are more common for long projects because they allow more time for such changes and greater probability that the owner's needs will change. Project scope may be changed for a variety of both exogenous and endogenous reasons: a) the owner may "change his mind" due to economic or political conditions; b) additional elements may be needed due to changes in environmental and safety standards; c) deliberate omissions ("buy-ins") to obtain approval to go ahead with the project may have occurred.

*<u>Design changes</u>. They are sometimes the consequence of scope changes and, as in the case of scope changes, correlate with project length. Design modifications may have significant impact on costs, since such changes may engender "ripple effects" throughout the entire project. The sources of design changes are, in addition to exogenous conditions, inadequacies and incompleteness of preliminary surveys, which can substantially affect estimation accuracy.

The later that scope and design changes are executed in the construction stage, the larger the impacts on costs will be.

*Other changes. We can include here those stemming from the lack of completeness and accuracy in the design and specifications and the "normal" variation in quantities (due to loss, spoilage, rework) and crew productivities.

—Management and organizational performance. Project management is responsible for the organization, leadership, control, coordination, and integration of the efforts of the personnel and organizations involved in a construction effort. Project management and organizational performance can have a considerable impact on the cost of a project. Many factors may produce overrun problems:

*Organizational structure. Appropiate project organizational structure and administrative procedures appear essential for controlling costs. The Trans Alaska Pipeline System (TAPS) illustrates this case. According to an investigatory report, duplication of supervision and decisionmaking, cumbersome decision chains, unclear lines of authority, and fragmentation of responsibility in the organizational structure, together with communication, coordination and liaison problems between project groups, led to the cost overrun.

According to Murphy et al. (1974), the projectized form of organizational structure, unlike the matrix form, is most often associated with project success. In general, it is important for the project manager of a large, long duration project to have key functions of the project team under him. He should be delegated sufficient authority by the owner to make important decisions and direct the project team. Murphy also says that public projects involving the cooperation, funding and participation of several government agencies often result in the creation of elaborate bureaucratic structures, decision delays, "red tape", and relatively diminished success.

*Organization manning. The competence, expertise, historical experience, and efficient utilization of personnel are factors affecting cost growth. These personnel may either be the direct staff of the agency ordering construction (owner), or contracted employees and organizations which may specialize in particular types or phases of projects.

A very few large organizations operate on a "force account", i.e., in-house construction force, basis. The Tennessee Valley Authority (TVA) is a case in point, in contrast with the Corps of Engineers and the Bureau of Reclamation. TVA's use of the force account has led it to important economies due to relative certainty of labor costs and productivity levels, and the administrative and technical expertise accumulated over years.

Contracting, however, usually constitutes the most effective option for most construction projects, since it allows the availability and utilization, for short periods, of organizations and specialists with administrative and technical expertise and historical experience not otherwise available.

It is important to involve the key personnel in the project as early as possible. Key personnel are usually the project manager, the chief design engineer, the construction supervisor, and the chief operating engineer who supervises startup.

*Project planning and control. The lack of effective utilization of planning systems and management control techniques may lead to cost overruns.

Murphy et al. (1974) affirm that networking systems, such as PERT and CPM, do contribute to project success, specially when initial over-optimism and/or "buy-in" strategy has prevailed in the securing of the contract; but the importance of those networking systems is far outweighted by other factors including the use of project tools known as "systems management concepts". These include work breakdown structures, life cycle planning, systems engineering, configuration management, and status reports. They found that the overuse of PERT-CPM systems may hamper success.

According to Merrow et al. (1979), every project presumably has an optimal schedule. Deviations from that "appropriate" schedule through either overly compressed or overly relaxed schedules may produce cost increases.

—<u>Exogenous causes</u>. A number of factors outside the control of project management can increase cost over the original estimate:

*<u>Inflation</u>. Where inadequate allowance for construction industry price level changes are contained in project estimates, estimation error will result. Economic changes are difficult to forecast. Usually, the costs of the different components of the project will not follow exactly the overall price change or general inflation, nor will the construction industry price change as typified by several construction cost indexes. The costs will depend on the particular project undertaken and the area market-conditions. In areas of high construction and economic activity, extra costs may be incurred due to tight market

conditions.

In general, where the period of time between the estimate and the completion of project construction is long, unanticipated construction price level changes are likely to be an important factor behind cost overrun.

We will devote Chapter 5 to analyzing inflation in more depth.

*<u>Regulatory, legal and political conditions</u>. Modification of these conditions can cause scope and design changes to meet new needs and standards, as we have already commented.

Cost growth may also be produced by court and regulatory delays. Koehn et al. (1978) conducted a study on cost increases due to delays initiated by government agencies. These agencies include the Environmental Protection Agency (EPA), the Nuclear Regulatory Comission (NRC), the courts, and other local, state and Federal government units. The results were based on a questionnaire sent to the U.S. main consulting firms. It was found that the weighted percentage of construction cost that may be allocated to delays due to government regulation is as much as 47.2% (unweighted, 30.3%)(*).

(*) The authors say that the percentages may be a bit inflated since a greater number of consultants who have experienced long delays may have returned the questionnaire than those who have experienced short or no delays.

Another factor that can influence the project cost is the bid system used by the owner for constructor selection and the administrative procedures (mainly, change procedures).

*Other exogenous factors. These include bad weather, labor strikes, and failure of vendors to deliver materials or equipment item as planned.

—Limitations on estimating methods. Merewitz (1972) and Merrow (1979) demonstrate a correlation between project type and the accuracy of cost estimates for various types of public works projects. Ad hoc, first-of-a-kind construction projects, which have large design and construction uncertainties surrounding them, were found to have the greatest deviations from estimated costs.

For water projects, according to Merrow, several authors have demonstrated substantial differences in the cost estimation error experienced across types of projects. Frequency of overruns was least for dredging, locks and dams, and greatest for local protection works and reservoirs - flood control and multipurpose projects.

The limitations on estimation techniques clearly stem from the uncertainty embodied in the type and complexity of projects.

1.4 Objective of the Study

We have seen that the cost of construction projects is quite uncertain. Studies on public and private projects have shown very frequent cost overruns.

Our goal is to quantify that uncertainty, incorporating all the information available. Particular information on the specific project under study is combined with that gathered through past experience with similar projects.

The viewpoint with which we deal is that of the owner, who is the party ultimately interested in the risk that a project is "carrying". The sought output is therefore the distribution of the owner's actual final cost. From that distribution, it is straightforward to obtain such interesting results as the expected final cost, its standard deviation or "spread", the risk of not meeting the initial estimate - or the authorized budget -, and the adequate contingency amount to be allocated. Furthermore, the distribution can be obtained at different point in time during construction and, thus, the evolution over time of the risk can be known.

We are not concerned with cost-benefit analysis. The projects to which these methodology applies are those that have passed that phase successfully, and their definitive design is completed or well under way.

As far as the contractor is concerned, the distribution of his actual cost can be determined in an analogous way to the "in-house" construction case.

We analyze the different contractual arrangements to carry out a construction project and how they affect the cost uncertainty, i. e., how the total cost figure in each arrangement may be disaggregated into components to better explain the total uncertainty. Procedures available to propagate the uncertainty from elementary components to the total are systematized and described.

As it can be inferred from Chapter 2, the literature is varied. Some studies are too simplistic, many forget essential aspects, and others are limited to only a particular type of contractual arrangement. Construction reality is different: there are many possibilities of owner-contractor agreement that can shift the risk from each other, internal relations - correlations - between components often exist, and the historical experience provides information that should not be disregarded.

The methodology described may be useful to either private owners or public agencies. It can be used as an internal tool to know the project risk so as to take corrective steps on time and to avoid late "surprises". In addition, it may help support the petition of funds from financial institutions.

The case study in Chapter 7 illustrates the practical application of the material covered on a typical construction project.

CHAPTER 2

BACKGROUND

2.1 M. Gates (1971) (*)

As a part of a much broader study about the different contingencies in construction contracting, Gates addresses the constructor's problem of calculating the percentage markup (profit) to be added to the cost in order to be assured (up to a certain degree) that the true cost does not exceed the bid price.

His aim is therefore to obtain the distribution of the total cost from the distribution of more elementary components.

He considers the project decomposed into operations or items. The cost C_i of an operation i is equal to the product of the number of units of work Q_i (quantity), and the unit cost U_i.

Each input (C_i and U_i) is random with a triangular distribution which is defined (Fig 2.1) by its average value $\overline{X}$ and the maximum error, e%, of X. The variance can be easily calculated.

The cost of a certain item i has a mean:

$$\overline{C}_i = \overline{Q}_i \times \overline{U}_i$$

(*) Gates, M., "Bidding Contingencies and Probabilities", Journal of the Construction Division, ASCE, Vol. 97, No. C02, Nov., 1971.

and a variance:

$$\sigma_{C_i}^2 \approx \sigma_{Q_i}^2 \; \bar{U}_i^2 + \sigma_{U_i}^2 \; \bar{Q}_i^2$$

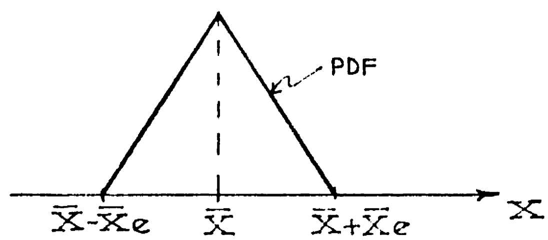


Figure 2.1

Triangular distribution
(symmetrical)

The total cost of a project including n different items of work (with independent cost contribution) has a mean:

$$\bar{T} = \sum_{i=1}^{n} \bar{C}_i$$

and a variance:

$$\sigma_T^2 = \sum_{i=1}^{n} \sigma_{C_i}^2$$

The author assumes that the total cost is also triangularly distributed and suggests the use of a standard triangular distribution to calculate the 'markup' over the estimated total cost $\bar{T}$ in order to have a particular probability of breaking even.

The viewpoint considered is that of the constructor; therefore, this case is analogous to the 'in-house' construction case (from owner's viewpoint) that we will study later.

The approach is very simplistic. The unit cost of every item could have been clearly disaggregated further into its components (quantity, productivity, unit costs), since the constructor knows - or is supposed to know - the participation of each one in the cost of every item. Also, the overhead costs, so difficult to allot rightly among items, should have been considered separately.

The author does not take into account possible dependencies or correlations that may exist and usually exist inside and between items.

Finally, the distribution of the total cost is obviously not triangular, but probably closer to normal or lognormal, according to the well known Central Limit Theorem (for sums or products, respectively). The standard normal distribution, widely tabulated, then can be used to calculate the markup.

2.2 J.E. Spooner (1974) (*)

The items or accounts into which the direct cost is broken down are now disaggregated further. The author

---

(*) Spooner, J.E., "Probabilistic Estimating," Journal of the Construction Division, ASCE, Vol. 100, No. CO1, March, 1974.

expresses the cost of an item as:

$$\text{Cost of item } i = \underset{L_i}{\text{Cost of labor}} + \underset{M_i}{\text{Cost of materials}}$$

where:

L_i = Quantity x Productivity x Labor rate

M_i = Quantity x Unit Cost

Each component is a random variable with given mean and variance.

He further considers correlation between L_i and M_i, within each item i, since both are linear functions of the quantity. For some items, he also assumes correlation between the cost of labor (L_i) because the same crew may carry out different tasks involving different items.

Then, the variance of the total cost is:

$$\sigma_T^2 = \sum_i \sigma_{L_i}^2 + \sum_i \sigma_{M_i}^2 + 2 \sum_i \text{Cov}(L_i, M_i) + \sum\sum_{i \neq j} \text{Cov}(L_i, L_j)$$

The author assumes that the correlation coefficient of L_i and L_j is equal to one and that the coefficients of variation are constant for all the items in order to arrive at a simpler expression.

Based on the central limit theorem and some simulations he adopts a normal distribution for the output. The level of risk for a project with a certain allotted contingency may then be calculated.

For the input components, either beta, traingular or uniform distributions are considered. They are defined by the optimistic, pessimistic and most likely value of the variable; the optimistic and pessimistic values being the extremes (0 and 100 percentiles).

This study only refers to 'in house' construction. On the other hand, equipment and overhead cost items are neglected. Inflation is an important factor that could have been separated as input.

Correlation certainly exists, but it is not limited only to the components mentioned. Given the number of items and the correlations between them, a rigorous analytical handling of the problem usually becomes cumbersome.

The adoption of a normal distribution for the output can be done if the dependency is weak, but not in every case.

Usually, the optimistic and pessimistic values judged by an expert do not correspond to the extremes of a distribution, but rather to the 5 and 95 (or 10 and 90) percentiles.

2.3 K.C. Carrier (1977) (*)

In a computerized project management information system, he uses a technique similar to the anterior and to that developed for the PERT in order to arrive at the distribution of the final cost.

(*) Carrier, K.C., "A System for Managing Escalation and Contingencies," Transactions American Association of Cost Engineers, 1977.

The project is broken down into contract packages. Personnel intimately concerned with the project are required to provide an expert estimate of the optimistic (C_o), pessimistic (C_p) and most likely value (C_m) of final cost of the individual contract packages.

The three cost estimates are fitted to a beta distribution. The mean and variance of each contract package is calculated using the formulae:

$$\bar{C}_i = \frac{C_{o_i} + 4\,C_{m_i} + C_{p_i}}{6}; \quad \sigma^2_{C_i} = \left[\frac{C_{p_i} - C_{o_i}}{6}\right]^2$$

Assuming independency of the contract packages, he applies the central limit theorem. The final cost of the project then follows a normal distribution defined by a mean:

$$\bar{T} = \sum_i \bar{C}_i$$

and a variance:

$$\sigma_T^2 = \sum_i \sigma_{C_i}^2$$

Although the central limit theorem is applied correctly, the independency assumed is unusual in most projects. If the dependency is strong, the distribution of the final cost does not follow a normal distribution.

Another remark is about the level of disaggregation. The work packages clearly could have been decomposed into their components, which are better known.

2.4 H.C. Bjornsson (1977) (*)

In a study about risk analysis of construction project cost estimates, he uses MonteCarlo simulation and describes a method for coping with the problem of correlation.

The project is divided into items, the costs of which are considered random variables. As in the previous cases, optimistic, pessimistic and most likely values of the variables are assessed by experts. From these three values a triangular distribution is defined for each variable.

A simulation is carried out and it is found that the total cost follows a normal distributions as the central limit theorem suggests for the case of independent random variables.

Then he addresses the correlation issue. He uses the "range sampling", a method similar to the "discriminate sampling" (Eilon, 1973), which we will describe in Chapter 4.

Ranges of values are assigned between the dependent variables. The specification of a positive dependency is done as follows: a sample from the dependent variable will be generated by a random number that is similar (within a

(*) Bjornsson, H.C., "Risk Analysis of Construction Cost Estimates," Transactions American Association of Cost Engineers, 1977.

certain range depending on the specified magnitude of co-variation) to the random number that was used to determine the value of the independent variable. In the case where +1 dependency is assumed, exactly the same random number is used in the sampling procedure.

When negative dependency is considered, the dependent variable is sampled by the antithetic random number. If the random number x is used to select a value of the in-dependent variable, then the random number (1 - x) is used to sample from the distribution of the dependent variable. If the magnitude of the negative relationship is assumed to be z, then a random number from a uniform distribution (1 - x - z, 1 - x +z) is used to generate a value of the dependent variable.

Another simulation is carried out on an example, assum-ing correlation between some cost items. A certain skewed distribution (nonnormal) is then obtained for the total cost.

Here also we can make comments about the level of dis-aggregation. If the cost items were further divided into their components (quantity, productivity, unit costs), the ranges of variation and the correlations between them could be more accurately determined. Inflation, an important factor in the total cost variation, should have been also explicitly separated.

The same remark as before may be made about the consideration of the optimistic and pessimistic values as the extremes (0 and 100 percentiles) of the input distributions.

CHAPTER 3

CONTRACTUAL ARRANGEMENTS AND THE OWNER'S RISK IN CONSTRUCTION INDUSTRY PROJECTS

3.1 Types of Contractual Arrangements

The owner's risk of cost overrun in a project depends largely on the type of contractual agreement between him and the contractor. Certainly, each type involves different risk sharing by both parties.

There are several alternative contractual (and organizational) approaches to the construction of a project. Each type has its strengths and weaknesses, and the type chosen will take into account the particular circumstances of each case.

In this section, we will describe the various types of contract forms and we will analyze how the uncertainty in the final cost may be disaggregated or "explained" in each type.

3.1.1 Owner-Builder or "In-house" Construction

This approach is also referred to as "force account." Historically, many city, state and federal agencies, as well as large private companies, have performed the actual construction with their own forces. Sometimes, consultants are utilized for some or all of the detail design and contractors for the hiring and supervision of the labor force.

Many of the owner-builders have developed design-construct divisions, mainly when the volume of work is relatively large and constant over a long period of time.

Clearly, in this approach the owner bears all the risk of the venture costs. The actual total cost is random since all its components, such as quantities to be executed, labor or equipment cost, are uncertain.

We can express the total cost of the project as:

$$\underset{TC}{\text{Total Cost}} = \underset{DC}{\text{Direct Cost}} + \underset{OC}{\text{Overhead Cost}}$$

where DC includes all the labor, equipment and material costs that can be associated with some activity or item of work while OC includes the indirect and administrative general costs of the projects.

For estimating, cost control and accounting purposes, DC is broken down into accounts or items according to functions, predominant craft or other common characteristics. A similar division is made for OC although usually on a more empirical basis. Therefore, we may write:

$$DC = \sum_{i=1}^{n} DC_i$$

$$OC = \sum_{j=1}^{m} OC_j$$

where the indexes i and j refer to the item.

Each item of direct cost can be decomposed into labor cost (LC_i), equipment cost (EC_i) and material cost (MC_i):

$$DC_i = LC_i + EC_i + MC_i$$

Those components can be further disaggregated in the following fashion:

$$LC_i = Q_i \times P_i \times LU_i$$

$$EC_i = Q_i \times M_i \times EU_i$$

$$MC_i = Q_i \times MU_i$$

in which,

Q_i = quantity of item i (for example, m^2).

P_i = productivity of labor in item i (man-hours/m^2).

M_i = productivity of equipment in item i (machine-hours/m^2).

LU_i = unit cost of labor in item i (\$/man-hour).

EU_i = unit cost of equipment in item i (\$/machine-hour).

MU_i = unit cost of material in item i (\$/m^2).

If several trades or labor categories take part in the production of the same item, either they would have to be considered separately or we could use a weighted average unit cost. The same happens in case of several machines or materials. The procedure, however, remains as described.

The different accounts or items OC_j of the overhead cost, such as mobilization, office expenses, indirect personnel, topography and the like, could also be disaggregated into components. In each case, the optimum degree of disaggregation or level of detail clearly varies. Relative weights of components to the total would have to be considered. Nevertheless, for the statistical treatment that we are developing here, it seems reasonable not to divide the OC_j accounts further.

The total construction cost of the project for the owner is:

$$\tilde{TC} = \sum_{i=1}^{n} \left[\tilde{Q}_i \times (\tilde{P}_i \times \tilde{LU}_i + \tilde{M}_i \times \tilde{EU}_i + \tilde{MU}_i) \right] + \sum_{j=1}^{m} \tilde{OC}_j \qquad (3.1.1)$$

The random variable $\tilde{TC}$ (*) appears now as a multiplication and addition of random variables, i.e., the variation of $\tilde{TC}$ can be explained through the variation of its components. Decomposed in this way, the task of quantifying its uncertainty becomes easier. The owner's estimating team can now guess a most likely value and a range of variation, or even a distribution of values for each variable. The

(*) The sign "~" over a variable means that such variable is random.

uncertainty can then be "propagated" to the total cost $\tilde{TC}$ using one of the methods that will be described later.

An important issue is that of <u>correlation</u> between the component random variables. It may happen that the productivities $\tilde{P}_i$ (or $\tilde{M}_i$) of labor (or machinery) in different items are correlated, since the same crew (or the same equipment) may be used to execute the work associated with both items. In most practical cases, it happens that $M_i = P_i$.

The quantities $\tilde{Q}_i$ may also be correlated. For example, in a project involving excavation of a "fixed" volume of material, the quantity (m^3) of rock is negatively correlated with the quantity (m^3) of soil found. Co-variation may exist also between particular unit costs in a region or period.

Therefore, correlation between the direct cost of some items, $\tilde{DC}_i$, will exist sometimes and can be difficult to quantify. Neglecting this aspect, however, can result in an erroneous outcome.

The impact of correlation is clear. When independent variables are added, the effect of variation of one variable can be offset by the variation in the opposite direction of another variable. In the case of dependencies, the effect of the variation of the variables can either be strengthened (positive correlation) or systematically compensated (negative correlation).

In Chapter 4, several possible approaches to cope with correlation will be suggested.

3.1.2 Negotiated Cost-Plus-Fee Contract

In this type of contract, the contractor agrees to perform the work for a fixed or variable fee covering profit and home-office costs, with all field costs being reimbursable at actual cost. These reimbursable costs are the direct cost DC and part of the overhead cost OC. This form is used for situations in which a firm bid cannot be made in advance of performance, because the project has not been fully specified.

These are several variations:

- Cost-Plus-Incentive-Fee Contract. It is devised to give greater profit motivation to the contractor than exists in other cost-type contracts.

In this type, the owner and contractor at the outset negotiate an estimated target cost, a target fee, and a fee adjustment formula. After execution of the contract, the fee payable to the contractor is determined in accordance with the formula. The formula may typically provide that the contractor would be penalized 25% of the actual cost overruns above the target estimate and rewarded by 25% of the underruns. In some cases, a minimum and a maximum fee is agreed.

The total construction cost TC for the owner is:

$$TC = \text{Reimbursable Cost} + \text{Fee} = RC + F \qquad (3.1.2)$$

The reimbursable cost may be disaggregated as we did before:

$$RC = \sum_{i=1}^{n} \left[Q_i \times (P_i \times LU_i + M_i \times EU_i + MU_i) \right] + \sum_{j=1}^{\ell} OC_j \qquad (3.1.3)$$

where ℓ is the number of reimbursable overhead cost items.

The fee is equal to the target fee plus the adjustment:

$$F = F_o - 0.25 \times (RC - RC_o) \qquad (3.1.4)$$

where F_o = target fee

RC_o = target (reimbursable) cost.

Substituting equations (3.1.3) and (3.1.4) in equation (3.1.2), we have:

$$\tilde{TC} = 0.75 \times \left\{ \sum_{i=1}^{n} \left[\tilde{Q}_i \times (\tilde{P}_i \times \tilde{LU}_i + \tilde{M}_i \times \tilde{EU}_i + \tilde{MU}_i) \right] + \sum_{j=1}^{\ell} \tilde{OC}_j \right\} + \underbrace{F_o + 0.25 \times RC_o}_{\text{constant}}$$

In the case of a minimum and a maximum fee, the procedure would remain the same, although with more equations.

We have again expressed the r.v. $\tilde{TC}$ as a function of its component r.v.'s. The expression is quite similar to that of "in-house" construction.

Most of the risk is assumed by the owner. The small part borne by the contractor is reflected in the premium that he will include in the fee agreed in order to offset such risk.

On the other hand, the owner does not pay (any premium) for the contingencies that do not take place.

- Cost-Plus-Award-Fee Contract. This is a rather new type which has been used mainly by government agencies.

The contractor is reimbursed by the owner for his actual allowable costs in performing the contract. He receives also a base fee (which does not vary regardless of the performance level he achieves) and, in addition, he is given the opportunity through superior performance to earn an additional award fee which may be two or three times the amount of the base fee.

For our purpose, which is to disaggregate the total cost up to a certain level so that its variation can be better explained through that of its components, this case is analogous to "cost-plus-incentive-fee" contract. Therefore, we will not give the expression for $\tilde{TC}$. The same holds for the rest of "negotiated cost-plus-fee" contract types, which will be described for the sake of completeness.

- Cost-Plus-Fixed-Fee Contract. It provides that all field costs will be reimbursed to the contractor at actual cost. He is paid a fixed fee of an amount agreed to at the beginning of the contract. The fixed fee varies only if there are changes made in the contract, but does not vary because of cost overruns or underruns in the orginal estimate. Thus, there is no incentive to the contractor, other than his pride, to save money. This type of contract, once widely used, has been largely superseded by cost-plus-incentive-fee or cost-plus-award-fee contracts.

- Cost-Plus-Percentage-of-Cost Contract. The contractor is paid his actual costs plus a percentage of those costs as his profit. The higher the costs are, the greater his profit is. The incentive is obviously not to save money. Despite this weakness, it has been used with contractors of presumed high integrity.

3.1.3 Unit Price Contract

In this type of contract the prices of specified units of work (items) are fixed, so that the total cost to the owner will vary with the actual quantities of units put in place. It applies best where the details and general character of the work are known, but the quantities are subject to variation within reasonable limits.

Most of the risk passes now to the contractor, who will include in his bid unit prices, in addition to direct costs, overhead costs and profit, an allowance for contingencies.

The total cost to the owner is:

$$\tilde{TC} = \sum_{i=1}^{n} \left[\tilde{Q}_i \times U_i \right] \qquad (3.1.5)$$

where,

n = number of payable items in the contract.

$\tilde{Q}_i$ = quantity of item i.

U_i = unit price of item i.

The random variable $\tilde{TC}$ appears now as a function of the r.v.'s $\tilde{Q}_i$.

As explained before, the correlation that may exist between the quantities $\tilde{Q}_i$ of different items should not be neglected.

3.1.4 Lump Sum or Single Fixed-Price Contract

The contractor agrees to perform the work for a predetermined single price that includes profit. It provides the owner a guaranteed figure of what his cost is going to be. It must be used only when the end product desired by the owner is known in complete detail before the contract is awarded. The contractor is highly motivated to achieve the lowest possible costs and to a lesser extent to achieve a rapid completion, but not to perform high quality work.

This type of contract passes the maximum amount of risk to the contractor. Thus, he will provide in his bid price an allowance for contingencies. That is to say, the owner now has to pay a premium for the certainty of a final total cost.

This total cost TC to the owner is equal, therefore, to the agreed single price, i.e., it is fixed. Only in case of "major" changes of scope, outside those foreseen in the contract, may a new price be agreed to.

3.2 Owner-Contractor Risk Sharing

Unit price and lump sum contracts have been the traditional form of the competitively bid (sometimes negotiated) general and turnkey contracts, what does not mean that they are always the more appropriate and likely to be less costly or more effective.

As we have seen through the different approaches, from in-house to lump sum contracts, the risk has shifted from the owner toward the contractor (figure 3.1). On the other hand, the owner has to pay for his higher security and for his lesser involvement in the project.

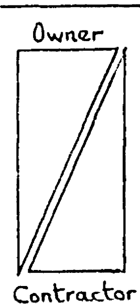
Risk Bearer	Arrangement
	. In-House Construction . Cost-Plus-Fee Contract . Unit Price Contract . Lump Sum Contract

Figure 3.1 Risk Sharing

The provision by the contractor in his bid of excessive allowance for contingencies may be overcome by an owner who is willing to assume some of those risks. Mostly, in unit price and lump sum contracts there are two ways of doing that. One is to include in the contract an escalation clause which provides for an increase or decrease in the contract price (or prices) if labor or material costs vary from an agreed base during the life of the contract. The other way is to include the standard "changed conditions" clause under which the owner accepts responsibility for increased costs if subsurface (or other type of) conditions differ materially from those shown in the plans or specifications or generally encountered in work of similar nature. In this way, the owner pays for contingencies if they actually take place, but he does not pay for contingencies that never occur.

We will deal with escalation in detail in a subsequent chapter.

CHAPTER 4

UNCERTAINTY PROPAGATION METHODS AND THEIR APPLICATION TO OUR CASE

4.1 Introduction

In the previous chapter, we have obtained the total cost to the owner, TC, in terms of more elemental components, such as the quantity Q_i of every item, the labor productivity P_i and so on, i. e.,

$$TC=f(Q_i,P_i,\ldots)$$

Given a project, the uncertainty about the components can be expressed in terms of probability distribution functions. The task now is to combine these distributions to obtain the distribution of TC for that project or, in other words, to "propagate the uncertainty".

In this chapter, we will explain three methods to accomplish this task: the Method of Moments, the Method of Discrete Probability Distributions and the Monte Carlo Method.

4.2 Uncertainty Propagation Methods

4.2.1 The Method of Moments

With this method we do not obtain the complete probability distribution of the total cost from the component distributions, but the moments of the total cost distribution from the moments of the component distributions.

We will start with the case of two variables. Let X and Y be independent variables having the probability density functions $f_X(x)$, $f_Y(y)$.

Addition. $Z=X+Y$

*The first moment or mean is

$$m_Z=E(Z)=\int_{-\infty}^{+\infty} z f_Z(z)dz,$$

where $f_Z(z)$ is the density function of Z, which can be expressed by the convolution integral

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx. \qquad (4.2.1)$$

Then

$$m_Z=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} z f_X(x)f_Y(z-x)dxdz$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f_X(x)f_Y(y)dxdy$$
$$=E(X)+E(Y).$$

Hence, the mean of a sum of two random variables is the sum of their means.

*The second central moment or variance is

$$\sigma_Z^2=\int_{-\infty}^{+\infty}(z-m_Z)^2 f_Z(z)dz$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[(x-m_X)+(y-m_Y)\right]^2 f_X(x)f_Y(y)dxdy$$
$$=\sigma_X^2+\sigma_Y^2.$$

Thus, the variance of a sum of two independent random variables is the sum of their variances.

Multiplication: $Z=XY$

The convolution now is

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(\tfrac{z}{x})\tfrac{1}{x}dx.$$

*From it, we can obtain the mean

$$m_Z=\bar{z}=\int_{-\infty}^{+\infty} zf_Z(z)dz$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)dxdy$$
$$=m_X\ m_Y.$$

*And the variance

$$\sigma_Z^2=\int_{-\infty}^{+\infty}(z-\bar{z})^2 f_Z(z)dz$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(xy-\bar{x}\bar{y})^2 f_X(x)f_Y(y)dxdy$$
$$=(x^2)(y^2)-(\bar{x})^2(\bar{y})^2$$
$$=\left[(x^2)-(\bar{x})^2\right]\left[(y^2)-(\bar{y})^2\right]+(\bar{x})^2\left[(y^2)-(\bar{y})^2\right]+$$
$$+(\bar{y})^2\left[(x^2)-(\bar{x})^2\right]$$
$$=\sigma_X^2\sigma_Y^2+m_X^2\sigma_Y^2+m_Y^2\sigma_X^2.$$

For the general case of several variables

$$Z=f(X_1,X_2,\ldots,X_n) \qquad (4.2.2)$$

we can expand the function f about the mean values of its arguments in a multivariate Taylor series,

$$Z=f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)+\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x_i-\bar{x}_i)+$$
$$+\frac{1}{2}\left(\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_i^2}(x_i-\bar{x}_i)^2+2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(x_i-\bar{x}_i)(x_j-\bar{x}_j)\right)+\cdots$$

where $\bar{x}_i$ is the mean of X_i and all the partial derivatives are evaluated at the mean values of the X_i variables.

Taking expectations of both sides of the previous equation, we can calculate the mean of Z:

$$\bar{z} = f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) + \frac{1}{2}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} \sigma_{X_i}^2 +$$

$$+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} E[(x_i-\bar{x}_i)(x_j-\bar{x}_j)] + \cdots \qquad (4.2.3)$$

where $E[(x_i-\bar{x}_i)(x_j-\bar{x}_j)]=\mathrm{Cov}[X_i,X_j]$, which is zero when the variables are uncorrelated.

The variance of Z can be obtained from:

$$\sigma_Z^2 = E(Z^2) - \bar{z}^2,$$

where $E(Z^2)$ can be determined by squaring both sides of equation (4.2.2), expanding in Taylor series and taking the expected value. The result is

$$\sigma_Z^2 = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_{X_i}^2 +$$

$$+ 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial f}{\partial x_i} E[(x_i-\bar{x}_i)(x_j-\bar{x}_j)] + \cdots \qquad (4.2.4)$$

Let us particularize for the sum and product.

<u>Sum of random variables</u>

$$Z = f(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i$$

*Mean. Equation (4.2.3) yields

$$\bar{z} = \sum_{i=1}^{n} \bar{x}_i. \qquad (4.2.5)$$

*Variance. Equation (4.2.4) gives

$$\sigma_Z^2 = \sum_{i=1}^{n} \sigma_{X_i}^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j). \qquad (4.2.6)$$

For uncorrelated variables, we have

$$\sigma_Z^2 = \sum_{i=1}^{n} \sigma_{X_i}^2 .$$

Product of random variables

$$Z=f(X_1,\ldots,X_n)=\prod_{i=1}^{n} X_i$$

*Mean. From equation (4.2.3) we have

$$\bar{z}=\prod_{i=1}^{n}\bar{x}_i+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\bar{z}}{\bar{x}_i\bar{x}_j}\,\mathrm{Cov}\,(X_i,X_j)+\ldots \qquad (4.2.7)$$

which, for uncorrelated variables, yields

$$\bar{z}=\prod_{i=1}^{n}\bar{x}_i.$$

*Variance. Equation (4.2.4) gives

$$\sigma_Z^2=\sum_{i=1}^{n}\left(\frac{\bar{z}}{\bar{x}_i}\right)^2\sigma_{X_i}^2+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E\left[(x_i-\bar{x}_i)^2(x_j-\bar{x}_j)^2\right]+$$
$$+\,2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\bar{z}^2}{\bar{x}_i\bar{x}_j}\,\mathrm{Cov}\left[X_i,X_j\right]+\ldots \qquad (4.2.8)$$

which, for uncorrelated variables, becomes

$$\sigma_Z^2=\sum_{i=1}^{n}\left(\frac{\bar{z}}{\bar{x}_i}\right)^2\sigma_{X_i}^2+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\sigma_{X_i}^2\,\sigma_{X_j}^2$$

4.2.2 The Method of Discrete Probability Distributions

Let X be an ordinary scalar variable and let $x_1,x_2,\ldots,x_n$ denote particular discrete values of X. Let $p_1,p_2,\ldots,p_n$ be associated probability values ($p_i=P(x_i)$) such that $\sum_{i=1}^{n}p_i=1$.

Then, the set of doublets

$$<p_1,x_1>,<p_2,x_2>,\ldots,<p_n,x_n>=\{<p_i,x_i>\}$$

represents the PMF (probability mass function) of X and may be thought of as a discrete approximation to a continuous PDF (probability density function) $f_X(x)$.

A discrete approximation to a continuous distribution can be done by simply dividing the total range of x into

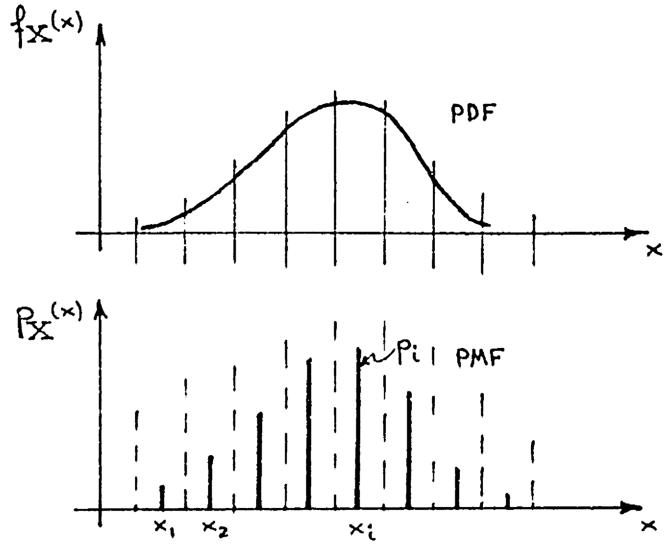


Figure 4.1 Discretization of a PDF.

intervals (figure 4.1). The probability that X falls in an interval is assigned to a single point x_i inside that interval. The point x_i may be, for example, the mean value of the points in each interval or the midpoint of the interval. Of course, the x_i do not need to be regularly spaced.

Sum of random variables

Let us suppose the following relationship among the random variables X, Y and Z:

$$Z = X + Y$$

and that our state of knowledge about them is expressed by the PMFs:

$$X \sim \{<p_i, x_i>\} \quad ; \quad Y \sim \{<q_j, y_j>\} .$$

Moreover, suppose that this state of knowledge implies interdependence between X and Y.

The PMF of Z may be defined as

$$\{<p_i, x_i>\} + \{<q_j, x_j>\} = \{<p_i q_j, x_i + y_j>\} \qquad (4.2.9)$$
$$= \{<r_{ij}, z_{ij}>\},$$

i. e., $r_{ij} = p_i q_j$

$z_{ij} = x_i + y_j$

Equation (4.2.9) may be regarded as a discrete analog to the convolution operation (4.2.1).

If X is "described" by n doublets

$$<p_1, x_1>, <p_2, x_2>, \ldots, <p_n, x_n>$$

and Y by m doublets

$$<q_1, y_1>, <q_2, y_2>, \ldots, <q_m, y_m>$$

then Z will have n×m doublets

$$<p_1 q_1, x_1+y_1>, <p_1 q_2, x_1+y_2>, \ldots, <p_1 q_m, x_1+y_m>,$$
$$<p_2 q_1, x_2+y_1>, <p_2 q_2, x_2+y_2>, \ldots, <p_2 q_m, x_2+y_m>,$$
$$\ldots$$
$$<p_n q_1, x_n+y_1>, <p_n q_2, x_n+y_2>, \ldots, <p_n q_m, x_n+y_m>.$$

<u>Product of random variables</u>

Similarly, in the case of multiplication:

$Z = XY$,

if X and Y are independent, we can define the PMF of Z by:

$$\{<p_i,x_i>\}\{<q_j,y_j>\}=\{<p_iq_j,x_iy_j>\}$$
$$=\{<r_{ij},z_{ij}>\}$$

with $r_{ij}=p_iq_j$

$z_{ij}=x_iy_j$

Here also Z is described by n×m doublets.

Correlation

A complication to that simple procedure appears when the variables X and Y are correlated.

In this case we must deal with conditional probabilities.

*Sum:

$$\{<P(x_i),x_i>\}+\{<P(y_j),y_j>\}=\{<P(x_i,y_j),x_i+y_j>\}=\{<P(x_i)P(y_j/x_i),x_i+y_j>\}$$
$$=\{<P(y_j)P(x_i/y_j),x_i+y_j>\}$$

*Product:

$$\{<P(x_i),x_i>\}\{<P(y_j),y_j>\}=\{<P(x_i,y_j),x_i+y_j>\}=\{<P(x_i)P(y_j/x_i),x_iy_j>\}$$
$$=\{<P(y_j)P(x_i/y_j),x_iy_j>\}$$

Therefore, the conditional distributions -or the joint distribution- of the correlated variables (figure 4.2) must be defined.

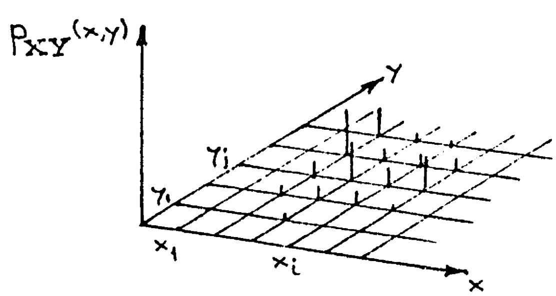


Figure 4.2 Joint distribution of X and Y.

Generalization

Given a function of several variables

$Z=f(X^1,X^2,\ldots,X^k)$,

where each X^k has a PMF defined as $\{<P(x^k_{i_k}),x^k_{i_k}>\}$, then the variable Z has a PMF given by

$$\{<P(x^1_{i_1})P(x^2_{i_2}/x^1_{i_1})P(x^3_{i_3}/x^1_{i_1}x^2_{i_2}) \ldots , f(x^1_{i_1},x^2_{i_2},\ldots,x^K_{i_K})>\}$$

which, if $X^1,X^2,\ldots,X^k$ are independent, becomes

$$\{<\prod_{k=1}^{K} P(x^k_{i_k}),\ f(x^1_{i_1},x^2_{i_2},\ldots,x^K_{i_K})>\}$$

As we can see, the procedure described is computer friendly. Note that in each operation the resulting number of doublets is the product of the numbers in the components, so that a condensation or aggregation may be necessary after each probabilistic operation in order to reduce the number of doublets.

4.2.3 The Monte Carlo Method

This is a numerical method of simulation widely used in many fields. Given a function

$$Z=f(X_1,X_2,\ldots,X_n), \qquad (4.2.10)$$

where each random variable X_i has a probability distribution, a value is generated for each X_i according to its distribution and a value of Z is obtained after performing (4.2.10). Repeating this procedure many times, a distribution of values of Z can be determined. The method lends itself typically to a computer, which can execute hundreds of simulations in a relatively short period.

If dependence exists between variables, conditional probability distributions must be defined in similar way to the previous method. The range of variation of correlated variables has to be partitioned into intervals. Then, if a value falling in a certain interval is generated for a variable, the value of its correlated variable must be generated from the conditional distribution given that interval for the first variable. The conditional distributions are created from statistical analysis of historical data or by judgement of the experienced estimator. In any case, the task is not easy and doubts can be raised about the reliability of the results.

There are simplifications, like the "discriminate sampling" of Eilon and Fowkes (1973), which is "a compromise between the two extremes of independent and conditional sampling". The procedure leans on the fact that managers prefer to consider relationships between variables in terms of the range of values which one variable might assume, given the values of other related variables, rather than in terms of conditional distributions. This means that the data to be input are: the probability distributions of the independent variables and the permissible ranges for the dependent variables, given the possible values for the related variables.

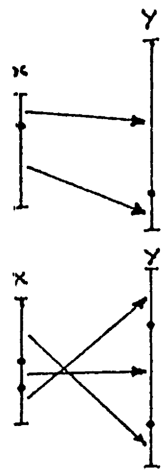


Figure 4.3

Correspondence between the variables X and Y.

The procedure, thus, consists in generating a value for X based on its probability distribution, and then generating a value for Y over a range corresponding to that of X and according to some density function. In this method, the following must be specified (figure 4.3) for two correlated variables X and Y: the respective truncation, the probability distributions assumed and the correspondence between ranges.

A further elaboration may be introduced in "discriminate sampling". It consists in considering that the correspondence between X and Y holds only for a given percentage of the simulations.

4.3 Distribution of the Components

The uncertainty about the components of the total cost TC of a project can be described in terms of a probability distribution. It is the purpose of this section to analyze which distributions are more suitable for our case.

The uniform, triangular, normal, beta and lognormal are the most common distributions. They offer wide possibilities and have been used in similar studies.

In the case of the variables with which we are concerned (quantity, unit cost and productivity), the estimator usually has in mind a certain value that is likely to occur more often than any other value and also a maximum and a minimum value. The beta and triangular distributions fit fairly well in such a way of estimating, and also can be skewed in any direction.

Beta Distribution

This distribution has two parameters and can adopt a wide variety of shapes. Since its CDF can not be evaluated in terms of elementary functions, it may offer some difficulty in the case of Monte Carlo simulation. It is, however, very appropriate for the method of moments.

The beta distribution is used in the PERT scheduling method, where the three estimates technique that we will describe briefly now was developed.

Three values are estimated (figure 4.4): an optimistic or minimum value (a), a pessimistic or maximum value (b) and a mode (m). Then, an approximation to the mean is:

$$\bar{x} = \frac{a + 4m + b}{6}$$

and to the variance:

$$\sigma_x^2 = \left[\frac{b - a}{6}\right]^2$$

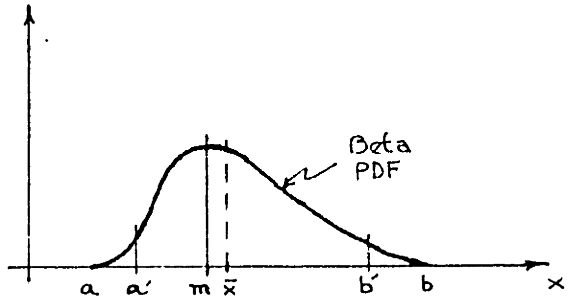


Figure 4.4 Beta Distribution.

Some authors (Moder and Rodgers 1968) have argued that the 5 and 95 percentiles (a' and b') should be used instead of the 0 and 100 (a and b), and therefore, the variance should be:

$$\sigma_x^2 = \left[\frac{b' - a'}{3.2}\right]^2$$

Two arguments are given: i) the difference (b' - a') varies from 3.1 to 3.3 standard deviations for a wide variety of limited range distributions (rectangular, triangular, beta), whereas the (b - a) varies from 3.5 to 6, so that a' and b' lead to a more robust estimator of σ_x; ii) the 0 and 100 percentiles are very difficult to estimate, since they would never have been experienced before by the estimator.

The two moments, $\bar{x}$ and σ_x^2, are related to the parameters, ρ and ν, through the following relations:

$$\bar{x} = \frac{\rho}{\nu}\ ; \qquad \sigma_x^2 = \frac{\rho(\nu - \rho)}{\nu^2(\nu + 1)}$$

Triangular Distribution

If the beta and uniform distributions are accepted as extremes of reasonable behavior, then the triangular may offer an acceptable compromise (figure 4.5).

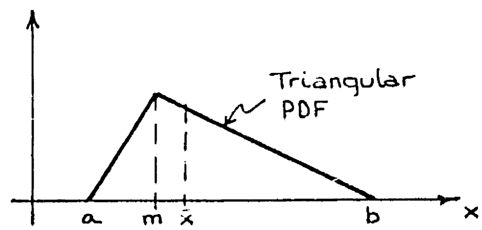


Figure 4.5

Triangular Distribution.

It is easy to handle and its mean and variance are:

$$\bar{x} = \frac{a + m + b}{3}\ ; \qquad \sigma_x^2 = \frac{(b - a)^2}{20}$$

4.4 Distribution of the Result

In previous sections we have obtained the total cost to the owner, TC, as an addition of products of random variables, that is to say, as an addition of random variables.

If the random variables to be added are independent, the central limit theorem is applicable. This theorem states that the sum of n independent random variables is a new random variable that, provided n is large enough, approaches the normal distribution.

And this holds regardless of the distributions of the variables summed (*).

In the case of independence, the uncertainty of TC can be easily defined. In fact, we know that it has a normal distribution, with a mean and a variance readily available through the method of moments.

If dependence exists between variables, the central limit theorem is less helpful. The total cost now will follow a different distribution. However, if the dependence is not strong, it can be approximated with a normal PDF.

The distribution of the output is not highly sensitive to different assumptions on the type of distributions of the components as long as their statistics (mean and variance) do not vary (Spooner 1974).

4.5 Discussion of our Case.

We will limit our study to the most representative cases of project arrangement: "in-house" construction and unit price contract. The others are analogous for our purpose.

Independent Components

As we have seen, the total cost TC will tend to be normally distributed. We only need to know the two parameters mean and variance, of such distribution. The method of moments provides the answer.

(*) For very skewed distributions, n must be larger.

• "In-House" Construction:

Mean. Equations (3.1.1), (4.2.5) and (4.2.7) yield:

$$\overline{TC} = \sum_{i=1}^{n} \left[\overline{Q}_i x(\overline{P}_i x\overline{LU}_i + \overline{M}_i x\overline{EU}_i + \overline{MU}_i)\right] + \sum_{j=1}^{m} \overline{OC}_j$$

Variance. Equations (3.1.1), (4.2.6) and (4.2.8) give:

$$\sigma_{TC}^2 = \sum_{i=1}^{n} \left[\overline{Q}_i^2 x(\overline{P}_i^2 \sigma_{LU_i}^2 + \overline{LU}_i^2 \sigma_{P_i}^2 + \sigma_{P_i}^2 \sigma_{LU_i}^2 + M_i^2 \sigma_{EU_i}^2 + EU_i^2 \sigma_{M_i}^2 + \sigma_{M_i}^2 \sigma_{EU_i}^2 + \sigma_{MU_i}^2) + \sigma_{Q_i}^2 x (\overline{P}_i \overline{LU}_i + \overline{M}_i \overline{EU}_i + \overline{MU}_i) + \right.$$

$$\left. + \sigma_{Q_i}^2 x (\overline{P}_i^2 \sigma_{LU_i}^2 + \overline{LU}_i^2 \sigma_{P_i}^2 + \sigma_{P_i}^2 \sigma_{LU_i}^2 + \overline{M}_i^2 \sigma_{EU_i}^2 + EU_i^2 \sigma_{M_i}^2 + \sigma_{M_i}^2 \sigma_{EU_i}^2 + \sigma_{MU_i}^2)\right] + \sum_{j=1}^{m} \sigma_{OC_j}^2$$

• Unit Price Contract:

Mean. From equations (3.1.5) and (4.2.5);

$$\overline{TC} = \sum_{i=1}^{n} \left[\overline{Q}_i \ x \ U_i\right]$$

Variance. From equations (3.1.5) and (4.2.6):

$$\sigma_{TC}^{2} = \sum_{i=1}^{n} \left[\sigma_{Q_i}^{2} \ x \ U_i^{2} \right]$$

Weak Dependence

In this case, a normal distribution can be assumed for TC and the mean and variance can be also obtained by the method of moments.

• "In-House Construction

Let us consider that two components within an item i are correlated.

Suppose that, for example, those are the productivities P_i and M_i. Then, P_i x LU_i and M_i x EU_i would be also correlated. The mean $\overline{TC}$ would be the same as in the independence case, and the variance σ_{TC}^{2} would be function of the covariance $Cov(P_i x LU_i, M_i x EU_i)$ according to equation (4.2.6). This covariance can be reduced to the $Cov(P_i, M_i)$ (*), which can be estimated from historical data of the joint behavior of P_i and M_i. The coefficient of correlation $\rho_{P_i M_i} = \dfrac{Cov[P_i, M_i]}{\sigma_{P_i} \sigma_{M_i}}$ may be useful for that purpose.

(*) $Cov(P_i x LU_i, M_i x EU_i) = E[(P_i x LU_i) x (M_i x EU_i)] - E(P_i x LU_i) E(M_i x EU_i) = \bar{P}_i \overline{LU}_i \bar{M}_i \overline{EU}_i + \overline{LU}_i \overline{EU}_i . Cov(P_i, M_i) - (\bar{P}_i^{2} \sigma_{LU_i}^{2} + \overline{LU}_i^{2} \sigma_{P_i}^{2} + \sigma_{P_i}^{2} \sigma_{LU_i}^{2}) . (\bar{M}_i^{2} \sigma_{EU_i}^{2} \ \overline{EU}_i^{2} \sigma_{P_i}^{2} + \sigma_{P_i}^{2} \sigma_{EU_i}^{2})$

If two components, for example Q_i and Q_j, belonging to different items are correlated, the variance of TC would be similarly function of the covariance of such components, Cov (Q_i, Q_j).

• Unit Price Contract.

In this case, correlation can exist only among the quantities.

The mean would be the same as in case of independence and the variance is

$$\sigma_{TC}^2 = \sum_{i=1}^{n} (\sigma_{Q_i}^2 \times U_i^2) + 2 \text{ Cov } (Q_i x U_i, Q_j x U_j)$$

$$= \sum_{i=1}^{n} (\sigma_{Q_i}^2 \times U_i^2) + 2 U_i U_j \text{ Cov } (Q_i, Q_j)$$

Strong Dependence

A priori, the distribution of TC is unknown. Therefore, the method of discrete probability distributions and the method of Monte Carlo are more suitable for this case, since they supply the distribution of the output.

With the first method, sets of conditional distributions must be prepared to define the dependencies. With the second, we can use instead of conditional distributions the "discriminate sampling", which is simpler and may give more confidence to the estimator.

4.6 The Ratio R

In the previous Chapters, we have described how, for a particular project, we can arrive at the distribution of the actual total cost ($\tilde{TC}$) from the distributions of more elemental components of the project.

The owner is interested in comparing this result with the estimated total cost (ETC). This amount ETC is typically the authorized budget (the estimate upon which the project authorization was based) for "in-house" construction and cost-plus-fee contracts, and the winning bid total price for unit price and lump sum contracts.

The ratio R, also called "cost factor", of actual total cost TC to estimated total cost ETC is very useful for comparison purposes and, therefore, we will use it in the next chapters.

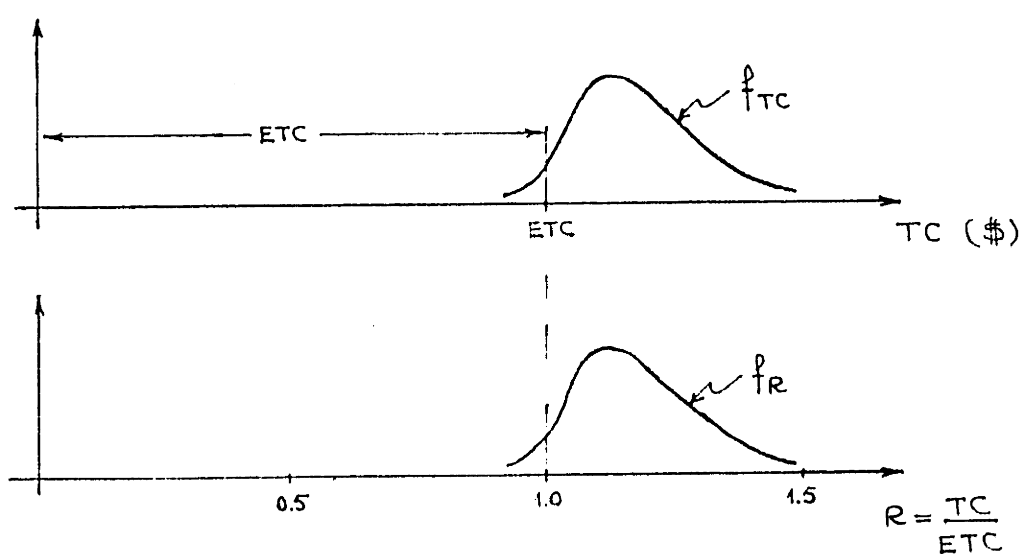


Figure 4.6

Distributions of the actual total TC and the ratio R = TC/ETC.

Since TC is a r.v. while ETC is deterministic, R is also a r.v. Given the distribution of $\widetilde{TC}$, the distribution of $\widetilde{R}$ is straightforward ($\widetilde{R} = \frac{\widetilde{TC}}{ETC}$; Figure 4.6).

Usually, the estimated total cost ETC does not include the effect of price changes or inflation in order to remain as reference. As far as $\widetilde{TC}$ is concerned, it may take into account the price changes, as we have done in the development of the equations, or may consider the original prices constant for the duration of the construction. Therefore, we take two possible views on R:

- One that includes inflation, i.e., it measures the overall deviation in the project total cost, or
- One that does not include inflation, and thus it expresses only the deviation imputable to other causes. This approach rests on the fact that inflation is out of the control of all the parties involved in the project and, therefore, nobody is 'responsible' for it.

Both approaches have been used in the literature. Clearly, both can be used together to express the uncertainty in the deviation from the estimated total cost, so that the effect of the inflation can be considered separately.

CHAPTER 5

INFLATION

5.1 Introduction

Inflation is one of the major causes of project cost overrun. Even if it has been taken into account in preparing the cost estimate, the difference between the actual and the foreseen inflation may produce an important deviation in the cost (in nominal or current dollars) of the project to the owner. One may argue that the real cost (in real or constant dollars) remains unaffected. As we will see, however, that only occurs when the inflation of construction costs is equal to the general inflation.

Since it can not be controlled either by the owner or by the contractor, inflation belongs to the category of exogenous or uncontrollable causes, and, therefore, it is beyond the project management accountability.

In this chapter, first, we will describe inflation in general and its relationship with other macroeconomic variables. Then, we will analyze inflation in construction projects. And finally, ways of further explaining the uncertainty of the owner's total cost by including the inflation will be given.

5.2 General Inflation

Inflation is the overall average increase in the level of prices. Like other macroeconomic variables, it has a pattern over time that rather follows the "business cycle" of economic expansions and recessions. The "business cycle" consists of the several-year-long irregular cyclical movements of output (GNP) relative to the trend of output and the associated movements of other economic variables (figure 5.1). These cyclical fluctuations have been observed for many years, although they have been much less severe since World War II.

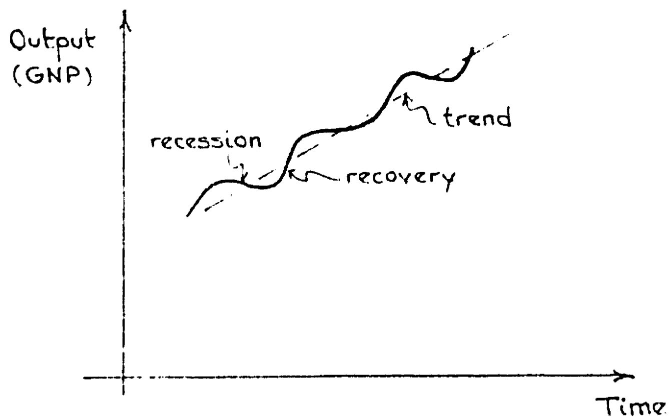


Figure 5.1

The economic "business cycle."

Shifts in aggregate supply and demand in the economy generate changes in the level of output and prices and, therefore, in the inflation rate. As the level of output changes, so does the unemployment (another major macroeconomic variable). Shifts in aggregate supply and aggregate demand are thus the underlying source of the business cycle.

The aggregate demand may shift due to:

- The fiscal policy. The government may reduce taxes so as to "expand" the economy (increase the GNP). The tax cuts, in the other hand, increase the government budget deficit. A fiscal expansion may be caused also by a higher government spending.

- The <u>monetary policy</u>. The money supply by the central bank, which in the U.S. is called the Federal Reserve (Fed), produces changes in interest rates and, therefore, in investment. Tight money tends to reduce inflation (figure 5.2). In the U.S., the Treasury can not create or "print" money, but if its deficit is financed by selling bonds to (borrowing from) the Fed, the result is the same since the quantity of high-powered money (*) increases.

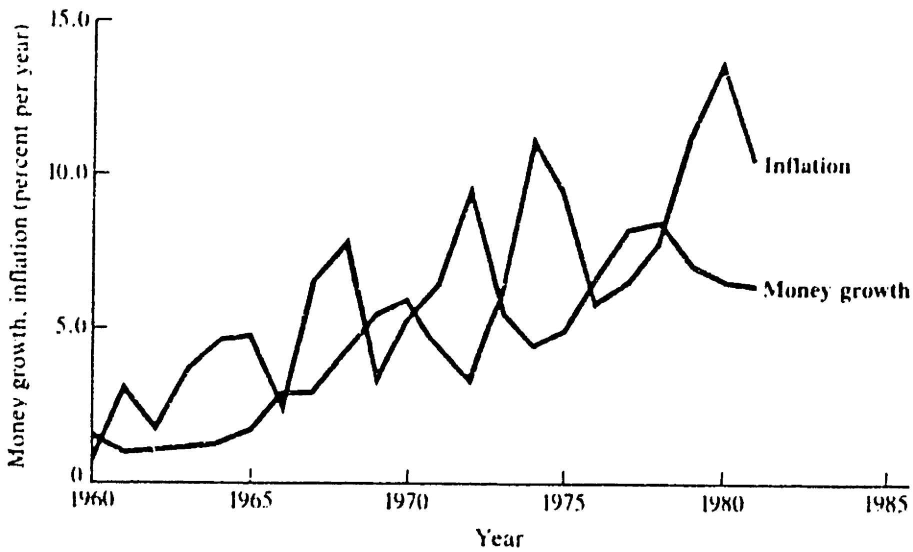


Figure 5.2

Money growth (M1) (**) and inflation (CPI), United States.
(Source: Fischer and Dornbusch, 1983)

- Changes in demand for goods in the private sector affect also the aggregate demand.

(*) High-powered money: currency plus bank reserves at the Fed.

(**) Money M1: currency and checkable deposits.

The aggregate supply may shift due to unusual changes in the price of certain goods or production factors, such as, rises in the price of oil or high wage settlements.

As the economy is hit by all these different changes, the level of output and the price level will be changing continuously and irregularly. Wars, for example, are a major source of disturbance to the economic system that can help set off booms and later recessions.

Inflation and deflation tended to alternate before the 1950s. But since then, the United States and most industrial countries have had a period of continually rising prices (persistent long-term inflation). Although inflation moves up and down, on balance it shows a steady upward trend.

Figure 5.3 shows the inflation and unemployment rate evolution in the U.S. We can note that the traditional 1960s concept of inflation process (Phillips curve) as a trade-off between inflation and unemployment - the less the unemployment, the more the inflation, and vice versa - is no longer valid.

Inflation and government budget deficit, which as we have seen are related, appear in figure 5.4.

The rate of inflation is calculated as the growth rate of the consumer price index (CPI) or the GNP deflator. The resulting value differs slightly.

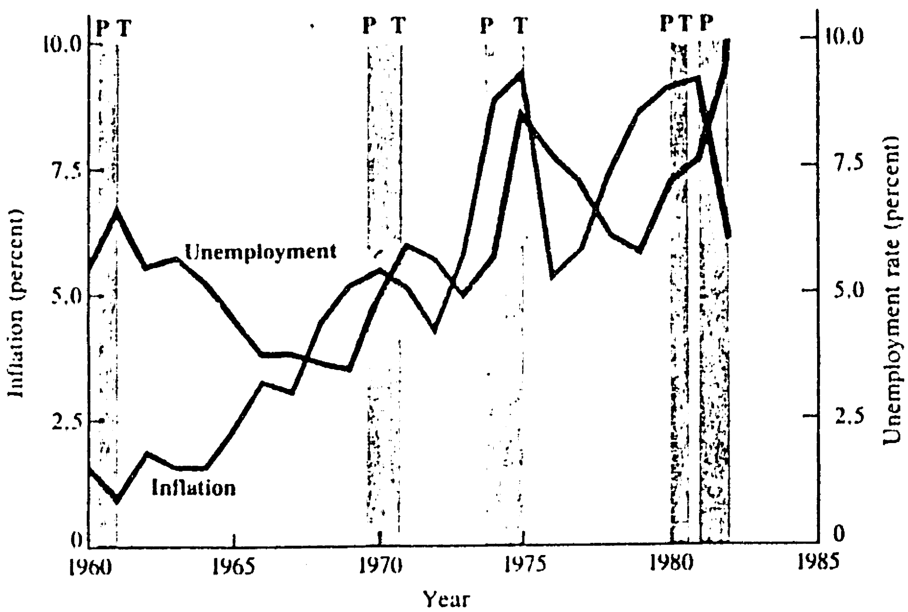


Figure 5.3

Inflation and unemployment in the United States.
Inflation rate is that of GNP deflator.
(Source: Fischer and Dornbusch, 1983)
(The shaded areas correspond to periods of recession. Recessions slow inflation but do not stop it for good).

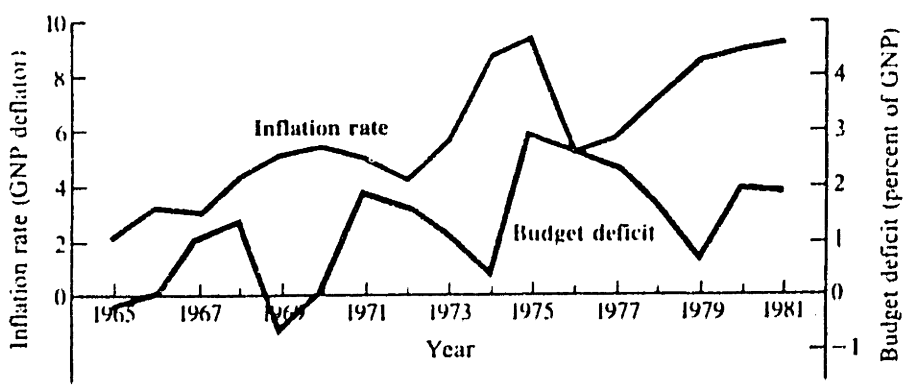


Figure 5.4

Budget deficits (as percentage of GNP)
and inflation, United States, 1965-1981
(Source: Fischer and Dornbusch, 1983)

The consumer price index (CPI) measures the cost of a typical basket of goods that households consume. Currently, it includes the prices of 224 different goods and services, from food and housing to entertainment and medical care. Therefore, it gives a good idea of how prices in general are changing. It is published every month.

The GNP deflator is the ratio of nominal GNP (production measured in current dollars) to real GNP (production measured in constant or "base year" dollars) expressed as an index.

5.3 Inflation in the Construction Industry

Inflation adds another element of risk in the construction industry, which, by its nature, is rather changing and unpredictable.

A small inflation rate does not usually interfere much with economic planning. A high and irregular rate, like that existing in many countries nowadays, cause distortions and inability to forecast return on one's investment. That problem hits in similar ways owners and contractors.

Owners may find government restrictions to the increase of the prices of the goods or services produced by their constructed plants or facilities, and thus may not be able to compensate the increased construction costs. Contractors suffer severe uncertainties in estimating and financing the work, and charge premiums on construction prices to offset them.

Each good or factor of production that participate in construction will have a price change which may differ from the general inflation rate and usually will do.

This means that the overall price change in the construction sector often will differ from that of the whole economy. And the same will happen for different kinds of projects and facilities. These changes differ also among areas.

The price changes are calculated as the growth rates of indexes prepared "ad hoc" for every specific type of facility and area. An index I_k, for a period k, is determined as a weighted average:

$$I_k = W_1 \frac{C_{1k}}{C_{10}} + W_2 \frac{C_{2k}}{C_{20}} + \dots + W_n \frac{C_{nk}}{C_{n0}}$$

where W_i = relative weight of component or item i

$$\left(\sum_{i=1}^{n} W_i = 1 \right).$$

C_{ij} = cost of item i in period j (C_{i0} is therefore the cost of item i in the base period).

The items are the materials, equipment and labor that take part in the construction of the facility.

Builders' construction cost indexes

Name, area & type	1981 AVG.	1982 AVG.	1983 AVG	1983 June	July	Aug.	Sept.	Oct.	Nov.	Dec.	1984 Jan.	Feb.	Mar.	Apr.	May	June
GENERAL PURPOSE COST INDEXES																
ENR 20-cities: Construction cost	329	356	379	379	383r	385	386	384	385	383	382	383	383	385	386	387
ENR 20-cities: Building cost	310	331	353	353	357	359	360	356	358	356	355	356	357	358	358	358
U.S. Commerce Department	322	327	334	327	333	336	337	337	335	338	339	340	343	344	—	—
BuRec, Denver, bldgs.	307	325	335	—	339	—	—	346	—	—	346	—	—	346	—	—
Engleeman General Bldg.	na	256	266	267	269	269	269	269	270	270	271	272	273	na	na	na
Federal Construction Co. (Florida)	na	na	na	212	212	212	212	212	213	213	213	213	213	213	na	na
Dodge bldg. cost	390	415	435	—	—	—	440	—	213	213	213	213	213	213	na	na
Factory Mutual indus. bldg.	319	349	367	—	372	—	—	—	—	—	—	—	na	—	—	—
Lee Saylor Inc.: labor/material	335	339	371	362	364	366	368	368	370	371	371	375	378	377	377	—
Means: const. cost	293	316	337	—	341	—	—	341	—	—	342	—	—	343	—	—
CONTRACTOR PRICE INDEXES—BUILDING																
Austin: central & eastern U.S., indus.	314	338	349	348	—	—	349	—	—	351	—	—	353	—	—	—
Lee Saylor Inc.: subcontractor	331	343	351	345	347	349	354	354	350	351	345	346	354	357	357	—
Turner: general	301	325	342	—	344	—	—	348	—	—	353	—	—	358	—	—
Smith, Hinchman & Grylls: general†	282	299	312	311	313	315	316	317	316	317	318	318	318	319	—	—
VALUATION INDEXES																
Boeckh index: 20-cities, comm'l & mfg††	311	337	354	355	357	361	359	359	359	360	363	364	364	365	367	--
Marshall & Swift: industrial	302	315	326	325	328	329	330	332	333	334	335	336	337	337	338	339
SPECIAL PURPOSE INDEXES																
Nelson Refinery Cost: "inflation" index	315	348	357a	356	359	361	364	364	365	365	366	na	na	—	—	—
Chemical Engineering plant cost	270	286	286	288	288	289	290	290	290r	291p	292	293	na	—	—	—
Port Authority of N.Y. & N.J., cost	311	338	356	353	353	360	360	360	360	360	364	362	362	362	—	—

Base. 1967 = 100

†Smith, Hinchman & Grylls is an A-E firm ††Prepared by E.H. Boeckh Co., div. of American Appraisal Assoc., Inc a-eleven mo. avg., r-revised; p-preliminary; na-not a
*Latest recorded period.

Table 5.1

The most common construction cost indexes.
Source: Engineering News Record (June 21, 1984)

There are many indexes available. In building construction, for example, at least 11 major indexes are compiled in the United States covering from 1 to 17 types of buildings from 1 to over 200 different locations. Some of these indexes are available at a charge from their compiler, while others are freely available. Table 5.1 shows the most common indexes in the U.S. There are also firms that develop their own indexes.

The average periodical (annual or monthly) price changes for a certain type of facility over a time span of m periods may be calculated from the indexes I_k and I_{k+m} at the beginning and the end periods of the span, respectively. The rate s is calculated as a geometrical average which implies a constant percentage change of prices from period to period:

$$s = \left(\frac{I_{k+m}}{I_k}\right)^{1/m} - 1 \qquad (5.3.1)$$

and the price in current dollars, or p_k in prices of the k+m period, will be:

$$p_{k+m} = p_k(1+s)^m$$

Sometimes, it is preferable to deal with the differential inflation δ, i.e., the price increase over and above the general rate of inflation f. Thus, $1 + s = (1 + f)(1 + \delta)$, and solving for δ:

$$\delta = \frac{1+s}{1+f} - 1$$

5.4 Escalation Clauses

As we commented in chapter 3, one way of overcoming the risk premium or allowance for price contingencies (inflation) that the contractors will include in their bid prices is to shift such risk to the owner by including an escalation clause in the contract. This clause is a contractual agreement by which the owner compensates the contractor for price escalation during construction.

The compensation is done by linking the owner's progress payments to an appropriate price index.

The payment to the contractor for work performed in period k is, therefore,

$$R_k^* = R_k \frac{I_k}{I_c}$$

in which R_k = value of work performed in period k in contract prices of (period c); R_k^* = adjusted value of the payment for work performed in period k; and I_k and I_c = price index values in the period k, and the reference period c, respectively. If the interval between owner payments allows for publication of several index values, then I_k would be an average of all the index values measured in the period k.

The index reflects an average composition of a large population of projects of a similar type. Although there are many indexes available, the one chosen may not reflect accurately the relative weights of the various items of a specific project, thus distorting the compensation for price escalation, mainly in the case of high inflation.

The solution to this problem is to build and index for each particular project. This index must represent the composition of works in the project and may be constructed as a combination of published indexes of basic components, such as labor, steel, concrete, asphalt, and so on.

In unit price contracts, each payable item may be associated to a different index, or each trade (carpentry, HVAC, etc...) or group of trades may have their own index.

There are some problems in the use of escalation clauses. For example, in the case of long-term contracts where the relative weights change over time. Also, the delay or lag between the work performance and the payment with compensation to the contractor may reduce substantially the value of the payment under conditions of high inflation.

5.5 Inflation in the Uncertainty of Owner's Cost

Here, we will explain further the uncertainty in the project's total cost (TC) to the owner by considering the change of prices over time. The types of project arrangement that are affected by price changes, as far as owner's TC is concerned, are "In-House" Construction, Cost-Plus-Fee Contract, Unit Price Contract with Escalation Clause, and Lump Sum Contract with Escalation Clause.

We will analyze each of them except Cost-Plus-Fee Contract, which is analogous to the "In-House" Construction for our purpose.

- "In-House" Construction

The unit cost of labor, equipment and material for every item (equation 3.1.1) may be disaggregated as follows:

$$\tilde{LU}_i = LU_{i_o} \times (1 + \tilde{s}_{L_i})^{\tilde{T}_i}$$

$$\tilde{EU}_i = EU_{i_o} \times (1 + \tilde{s}_{E_i})^{\tilde{T}_i} \qquad (5.5.1)$$

$$\tilde{MU}_i = MU_{i_o} \times (1 + \tilde{s}_{M_i})^{\tilde{T}_i}$$

where — LU_{i_o}, EU_{i_o}, MU_{i_o} = unit costs of labor, equipment and material, respectively, of item i at the time of making the estimate (period 0). Therefore, they may be taken as constant.

— $\tilde{s}_{L_i}$, $\tilde{s}_{E_i}$, $\tilde{s}_{M_i}$ = price change of labor, equipment and material, respectively, of item i per period (year, month, etc.). They are random.

— $\tilde{T}_i$ = period in which the item or activity i is performed. It is also random.

Actually, the value of s_i (s_{L_i}, s_{E_i}, s_{M_i}) will change from period to period. In order to avoid unpractical complications, the value of s_i considered in equations (5.5.1) is the average in the T_i periods.

Several models can be used to predict s_i. These statistical methods forecast, using "moving average" techniques, future movements of a variable based on its past performance over a representative time span.

The analyst can now combine these predictions with his judgement to estimate a distribution for s_i. The behavior of the general inflation and other macroeconomic variables may help in the estimation. The distribution of T_i can be determined taking into account the work schedule. As we commented in Chapter 4, those distributions can be obtained from the estimation of the most likely, the pessimistic and the optimistic values of the variables.

It is likely that correlation exists among $\tilde{s}_{L_i}$, $\tilde{s}_{E_i}$, and $\tilde{s}_{M_i}$ within an item and between items, since in general those price changes will tend to follow the pattern of the general inflation.

The uncertainty propagation methods more suitable in this case are the Monte Carlo method and the method of discrete probability distributions. The method of moments becomes very cumbersome here.

A complication arises in estimating T_i when the duration of the activity (or item) i is very long. A way of overcoming such a problem is to split that activity into shorter activities and to introduce the proper correlation between their T_i to keep the continuity. Another way is to distribute the work (or quantity to execute) uniformly over the foreseen duration $\tilde{D}$ (*). The value (constant $) of work to be executed in the

---

(*) As in previous chapters, the symbol "~" over a variable means that such variable is random.

period k (*) would thus be affected by the factor $(1+s)^k$. Then, the total cost of labor, for example, for a certain activity which starts in the period $\tilde{\tau}$ and lasts $\tilde{D}$ periods would be

$$\frac{\tilde{Q}}{\tilde{D}} \times \tilde{P} \times LU \times \sum_{k=\tilde{\tau}}^{\tilde{\tau}+D-1} (1-\tilde{s})^k \quad (**)$$

(in current dollars), where $\tilde{s}$ is the average periodical price change rate up to the period $\tilde{\tau}+\tilde{D}-1$. The last term is a geometric series that can be summed

$$\sum_{k=\tilde{\tau}}^{\tilde{\tau}+\tilde{D}-1} (1+\tilde{s})^k = \frac{(1+\tilde{s})^{\tilde{\tau}+\tilde{D}} - (1+\tilde{s})^{\tilde{\tau}}}{1+\tilde{s}-1} = \frac{1}{\tilde{s}}(1+\tilde{s})^{\tilde{\tau}}\left[(1+\tilde{s})^{\tilde{D}}-1\right] \qquad (5.5.2)$$

The same may be done with the equipment and material cost.

Of course, the duration $\tilde{D}$ is a direct function of the quantity $\tilde{Q}$ and the productivity $\tilde{P}$, i.e., $\tilde{D} = k \times \tilde{Q} \times \tilde{P}$.

(*) We can consider, for example, that these periods are quarters.

(**) τ is the time span between now (when the study is being undertaken) and the beginning of the activity. Therefore, LU is the present labor unit cost (current dollars).

- <u>Unit Price Contract with Escalation Clause</u>

Since the periodical progress payments are linked to an index, we may write:

$$\tilde{TC} = \sum_{i=1}^{n} \sum_{j=1}^{m} [(\tilde{Q}_{ij} \times U_i)(1 + \tilde{\bar{s}}_{ij})^j]$$

in which,

Q_{ij} = quantity of item i to be executed in period j.

U_i = contractual unit price of item i.

$\bar{s}_{ij}$ = average price change of item i in the j periods (geometric average).

Instead of $(1+\tilde{\bar{s}}_{ij})^j$, we can write $(1+\tilde{s}_{i1})(1+\tilde{s}_{i2})\ldots(1+\tilde{s}_{ij})$, but this disaggregation is unnecessary, usually due to the obvious difficulties in estimating the $\tilde{s}_{ij}$. Sometimes, it may be useful to consider an s_i for a certain number of periods and then a different s'_i for the rest of them, i.e., $(1+s_i)^p(1+s'_i)^q$, $(p+q=j)$.

The price change rates S_{ij} (or $\bar{s}_{ij}$) can be obtained from the predictions of the index adopted as we explained before (equation (5.3.1)).

As in the previous case, the uncertainty propagation methods that are computer friendly apply better here.

- <u>Lump Sum with Escalation Clause</u>

The expression of the owner's total cost now is:

$$\tilde{TC} = \sum_{j=1}^{m} \tilde{P}_j \, (1+\tilde{\bar{s}}_j)^j \qquad \left(\sum_{j=1}^{m} \tilde{P}_j = \text{contractual TC} \right)$$

in which,

P_j = progress payment to be made in period j.

$\bar{s}_j$ = average price change (according to the index of the project) in the j periods.

The comments about $\bar{s}_j$ and the uncertainty propagation methods made in the unit price contract with escalation clause hold also here.

CHAPTER 6

UPDATING AND EVOLUTION OF THE COST UNCERTAINTY

6.1 Description

Our aim is to find the distribution of the ratio $\tilde{R}$ of a project from all the information available. Therefore, the procedure of determination of the distribution of $\tilde{R}$ has to integrate and synthesize all the factors and causes that may affect the uncertainty about $\tilde{R}$.

In the Chapters 3, 4 and 5, we explained how to arrive at a distribution of $\tilde{TC}$ and thus $\tilde{R}$ from the distribution of its more elementary components, i.e., from specific knowledge and information of that particular project. The estimators, when supplying their most likely, pessimistic and optimistic values of the different variables (quantities, costs, etc.), are taking into account causes such as: accuracy and completeness of the design, possible (minor) changes of scope, subsurface conditions, variability in productivities, and inflation.

Other relevant factors, however, are not captured in that approach. These non-explicit factors or causes are, for example: organizational aspects (cooperation owner-project manager - contractor, decision delays, adequacy of change procedures, and the like), major changes of scope, and complexity and size. Clearly, these factors derive from the characteristics and peculiarities of the owner (government agency, private firm,...) and from intrinsic nature of every type of project (dams, highways, buildings,...). In other words, it is likely that the owner and his organization behave in similar ways and use analogous procedures in each project. It is likely also that projects of the same type carry similar uncertainty. To take into account these factors, we must look at the owner's historical data of past experience.

The owner may construct his own "historical" distribution of $\tilde{R}$, based on observed ratios of actual to estimated final cost of his completed projects. Ideally, a distribution can be constructed for every group of projects of similar type and size. The task may be simple in the case of agencies (owners) that specialize in a certain type of facilities. Most of the records of a State Department of Transporation, for example, would be on highways. A proper size discretization would be required based on length, earthworks volume or total cost. Since also an agency often uses one or two types of contracts and these, as we know, are related to the owner's cost uncertainty, separate distributions would be required for each type.

Let us summarize. The information available to the owner in order to obtain the distribution of the ratio R for a certain project is:

- the "historical" distribution of R for similar projects, which incorporates, through the past owner's performance, organizational and project-type-intrinsic aspects, and
- the "specific" distribution of R for the project. This distribution integrates factors related to the particular project under analysis.

If personal historical records are not available to the owner, general data of past experience in the state or region may be useful. Really, the latter is another source of potentially valuable information in any case, and should not be ignored (*).

6.2 Bayes' Theorem

The tool that allows us to combine various types of information, or to update the previous knowledge with the new specific information, is the Bayes' theorem.

Let A_i, i=1,...,n be simple events (mutually exclusive and collectively exhaustive) and let B be an event of the same sample space (figure 6.1). The Bayes' theorem states that the conditional probability of A_i given B is:

(*) In fact, the information can be incorporated with double updating (using Bayes' Theorem twice).

$$P[A_i|B] = \frac{P[B|A_i]P(A_i)}{\sum_{i=1}^{n} P[B|A_i]P(A_i)} \quad (*)$$

in which

$P[A_i]$ is the "prior" probability of A_i.

$P[A_i|B]$ is the updated or "posterior" probability of A_i.

$P[B|A_i]$ is the likelihood of observing B if A_i is true.

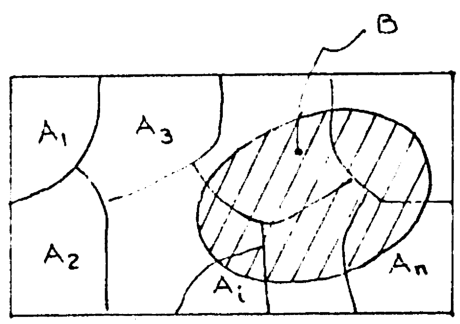


Figure 6.1

Venn diagram showing the simple events A_i, i=1,...,n and the event B.

(*) Proof:

$$P[A_i|B] = \frac{P[A_i \cap B]}{P[B]} = \frac{P[A_i \cap B]}{\sum_{i=1}^{n} P[A_i \cap B]} = \frac{P[B|A_i]P[A_i]}{\sum_{i=1}^{n} P[B|A_i]P[A_i]}$$

6.3 Updating R

In our case, the "historical" distribution of the ratio R has to be updated based on the evidence E, leading to the "posterior" distribution of R for our particular project. Discretizing both distributions, we may write, according to Bayes' theorem:

$$P[\tilde{R} = R_i | E] = \frac{P[\tilde{R}=R_i]P[E|\tilde{R}=R_i]}{P[E]} \qquad (6.3.1)$$

where $P[E|\tilde{R}=R_i]$ is the likelihood of the evidence if $R = R_i$ (figure 6.2). The term P[E] is a normalization factor:

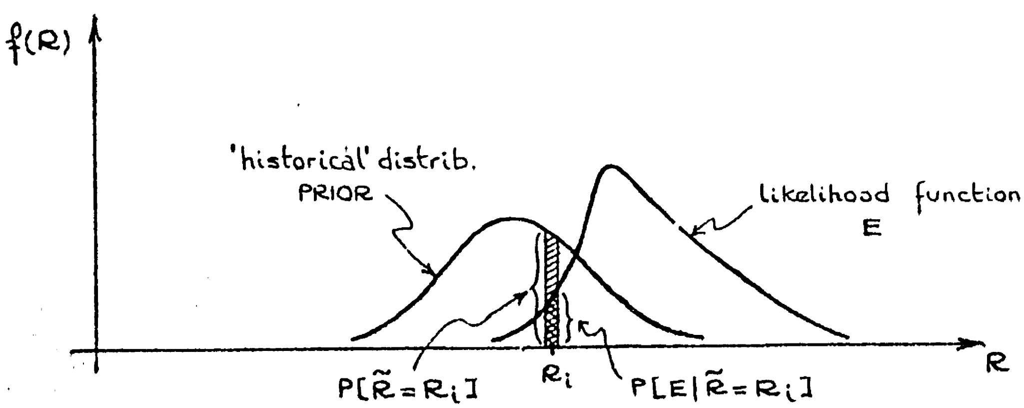


Figure 6.2

Bayesian updating.

Usually, the "historical" and the likelihood function will be available in discrete form. The updating procedure described lends itself very well to a computing machine.

Equation (6.3.1) may be expressed in continuous form:

$$f''(R)\ dR = \frac{f'(R)dR \times f^{s}(R)dR}{k.dR}$$

Hence,

$$\underline{f''(R) = C.\ f'(R).f^{s}(R)} \qquad (6.3.2)$$

where C is a normalizing constant:

$$C = \left[\int_{-\infty}^{+\infty} f'(R) f^{s}(R) dR \right]^{-1},$$

$f'(R)$ is the "prior" or"historical" distribution of R,

$f^{s}(R)$ is the likelihood function given R.

Specifically, the beta distribution can adopt a wide variety of shapes by appropriately choosing its parameters ρ and σ. Its PDF is:

$$f(x) = \frac{1}{B(\rho,\sigma)}\ x^{\rho-1}(1-x)^{\sigma-1} \qquad \begin{array}{l} 0 \le x \le 1 \\ \sigma = \nu - \rho \\ \rho,\ \sigma > 0 \end{array}$$

where $B(\rho,\sigma)$ is the complete beta function (*).

If we can fit satisfactorily a beta distribution to the "historical" and "specific" distributions of R, then the "updated" (a posteriori) distribution of R will be another beta distribution. Let

$$f'(R) = C_1 . R^{\rho'-1} . (1-R)^{\nu'-\rho'-1} \equiv \text{beta } (\rho',\nu')$$

and

$$f^s(R) = C_2 . R^{\rho^s-1} . (1-R)^{\nu^s-\rho^s-1} \equiv \text{beta } (\rho^s,\nu^s)$$

where C_1 and C_2 are suitable constant. By equation (6.3.2), we then have

$$f''(R) = C . R^{(\rho'+\rho^s-1)-1} (1-R)^{(\nu'+\nu^s-2)-(\rho'+\rho^s-1)-1}$$

which is a beta $(\rho'+\rho^s-1, \nu'+\nu^s-2)$.

The fitting of the beta distribution may be done by moments or fractiles (**). An χ^2 test, then, may be performed to test the goodness of the fit.

(*) $$B(\rho,\sigma) = \int_0^1 t^{\rho-1}(1-t)^{\sigma-1}\, dt, \quad \rho > 0, \ \sigma > 0$$
$$= \frac{\Gamma(\rho)\ \Gamma(\sigma)}{\Gamma(\rho+\sigma)}$$

(**) Pratt, J.W., et al., "Introduction to Statistical Decision Theory", McGraw-Hill Co., Inc., 11.5.2-11.5.3

In most cases, however, it will be preferable to operate with discrete distributions through the help of computer routines.

From equation (6.3.2) it is possible to draw qualitative conclusions about the shape of the posterior density function in two extreme cases. If $f^s(R)$ is gently or slowly changing where $f'(R)$ is rapidly changing, then the shape of $f''(R)$ will be determined mainly by $f^s(R)$, i.e., $f^s(R)$ does not contain very much information. Similarly, if $f'(R)$ is gentle and $f^s(R)$ is rapidly changing, then $f''(R)$ will look like $f^s(R)$ (figure 6.3).

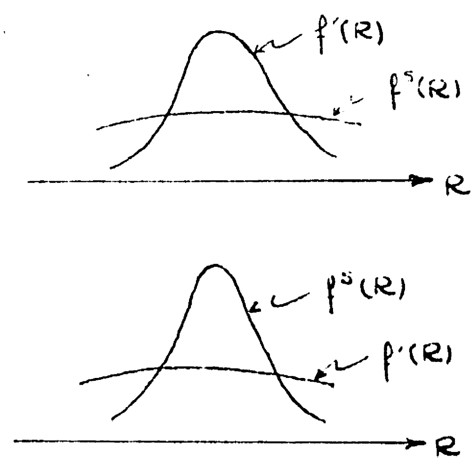


Figure 6.3

Different shapes of $f'(R)$ and $f^s(R)$.

6.4 <u>Evolution of the Uncertainty as the Project Progresses</u>

The construction process, generally speaking, consists of developing a project from its abstract definition in the owner's mind to reality. It involves different stages. The most relevant are: a) preliminary design, which usually forms the basis of a feasibility study; b) design, which fully defines the elements of the project; and c) construction, in which the project is executed.

The owner, concerned about the cost, usually requires various estimates to be made at different points in time. Typically, the estimates are:

- Initial or proposal estimate. It is a "rule of thumb" estimate, which gives the order of magnitude of the total estimated cost to the owner. This estimate is made at the very beginning and it is based on experience of similar projects, often using published literature sources, such as cost vs. capacity curves.
- Preliminary design estimate. It is made upon completion of the preliminary design and, therefore, it is based on it. Emphasis is given to major elements.
- Detailed design estimate. After the design is completed and the project thus is defined in detail, this estimate is prepared. All the elements and items are disaggregated and priced in detail.

In unit price and lump sum contracts, the bidding takes place after the design estimate is available. Bids are received from the invited contractors and the owner awards the contract to a certain bidder according to the owner's procedures. The winning bid becomes the contractual price of the project, and the owner will use it as a reference against which he can compare the actual cost.

In "in-house" construction and cost-plus-fee contract, the design estimate, after the corresponding revisions, is approved or authorized by the owner. The authorized estimate becomes, therefore, the reference cost of the project in these contracts.

It is clear now when we can obtain, for the first time, the "specific" distribution of the rates $\tilde{R} = \frac{\tilde{TC}}{ETC}$ for our project. The distribution of $\tilde{TC}$ (owner's total cost) may be determined upon completion of the design, i.e., at the time of the detailed or design estimate. As far as ETC (estimated total cost) is concerned, its value is known after the contract award or the project budget approval if the arrangement is unit price and lump sum contract or "in-house" construction and cost-plus-fee contract, respectively. In the first case, ETC is the winning bid, and in the second, ETC is the authorized or approved budget.

The first updating of the distribution of $\tilde{R}$ can thus take place upon contract award or budget approval.

As the construction progresses, more information becomes available to the estimator. Then, in further stages of construction it is possible to obtain a new distribution of $\tilde{TC}$ and therefore, a new "specific" distribution of $\tilde{R}$, which may be updated with the last (itself updated) distribution of $\tilde{R}$ now acting as a prior.

The process can be viewed graphically in figure 6.4.

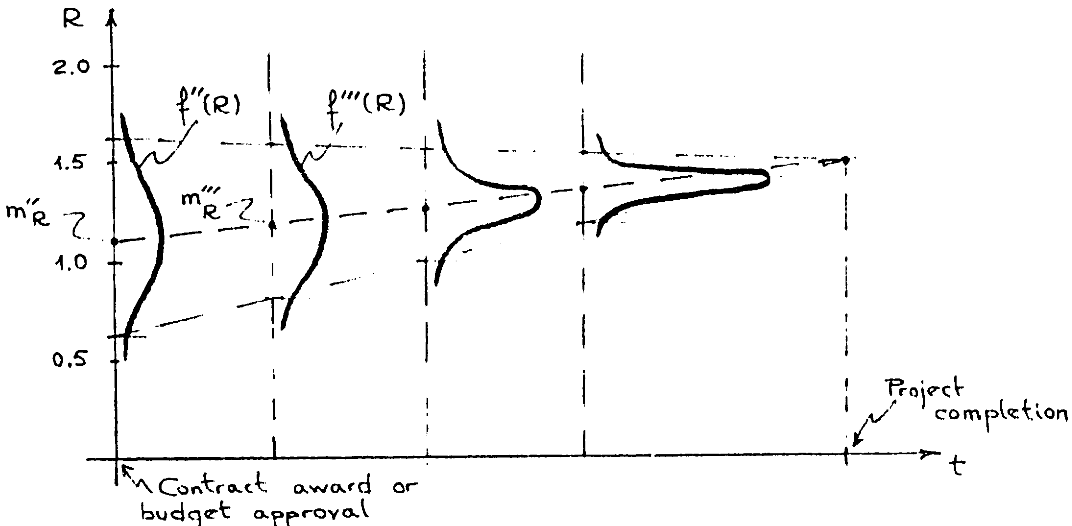


Figure 6.4

Evolution of the uncertainty of the ratio R.

In each new stage, the range of variation of the component variables will be smaller and thus the uncertainty of $\tilde{R}$ will decrease.

The random variable $\tilde{R}$ is function of the time, $\tilde{R}(t)$. Its mean $m_R(t)$ and variance $\sigma_R^2(t)$, and other moments also will be. As we can see, the variance $\sigma_R^2(t)$ decreases over time; and if we calculate the percentiles 5 and 95, for example, in each stage, it is possible even to predict the final value of R and the project completion time. Of course, this time can generally be predicted better through other scheduling methods; our main purpose is to study the cost (rather than the time) uncertainty.

Note that, as construction progresses, uncertainty decreases, and so do the possibility and ease of avoiding cost overrun.

6.5 Risk of Cost Overrun. Contingency

Given the distribution of the ratio $\tilde{R}$ at any point in time, it is easy to determine the risk of not meeting the authorized budget (ETC). This risk is the probability that R exceeds unity (figure 6.5).

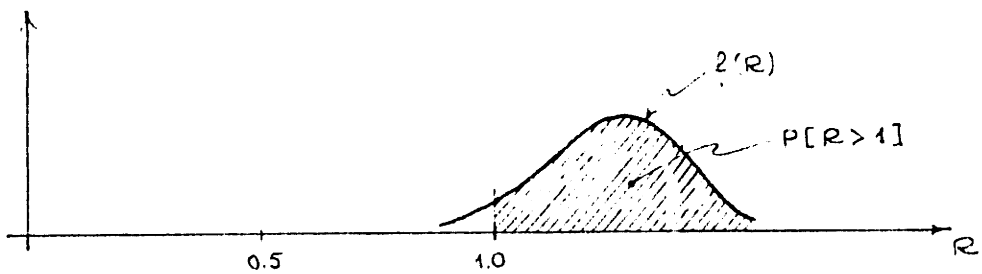


Figure 6.5

Risk of cost overrun.

$$P[\tilde{TC} > ETC] = P[\frac{\tilde{TC}}{ETC} > 1] = P[\tilde{R} > 1]$$

In the particular case that the distribution f(R) is normal, we have:

$$P[\tilde{R} > 1] = 1 - P[\tilde{R} \leq 1] = 1 - P[\frac{\tilde{R} - \bar{R}}{\sigma_R} \leq \frac{1 - \bar{R}}{\sigma_R}] =$$

$$= 1 - \Phi(\frac{1 - \bar{R}}{\sigma_R})$$

where $\mathbf{U}$ is the cumulative standard normal distribution, which is tabulated in many probability books.

Contingency is a monetary provision to cover cost increases in a project due to unforeseen factors. Usually, a contingency is allocated at the beginning of the project construction and is gradually depleted as the construction progresses. Typically the initial contingency may amount to 25% of the authorized budget.

From the distribution of $\tilde{R}$, it is possible to obtain the contingency necessary to reduce the risk to a certain level r, (figure 6.6).

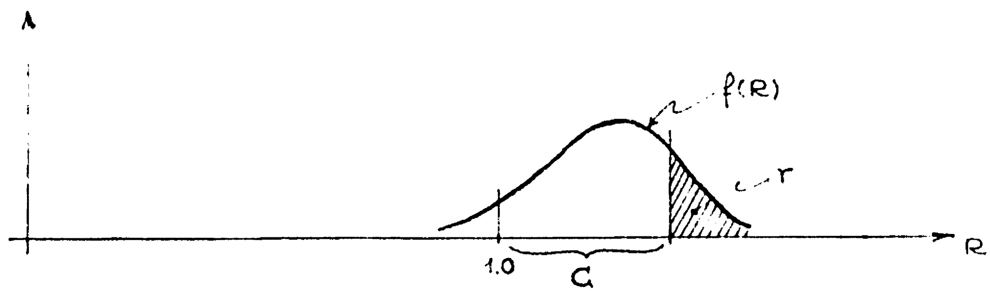


Figure 6.6

Contingency (C).

The contingency is obtained as a fraction C of the authorized budget (ETC):

$$\text{Contingency} = \text{ETC} \times C$$

If f(R) is normal, then C and r are related as follows:

$$C = \mathbf{U}^{-1}(1-r) - 1$$

CHAPTER 7

CASE STUDY AND CONCLUSIONS

7.1 Description and Inputs

- o Project: A 100 km. highway in a rural area.
- o Owner: Government agency.
- o Type of construction arrangement: In-house (*).
- o Status: Definitive design is completed, budget just approved and construction ready to start.
- o Approved budget (estimated cost): $64,350,000 (in present $).
- o Contingency allocated: 30% .
- o Work breakdown structure: As shown in table 7.1.
- o Project data: The "Input List" in table 7.2 shows, for each direct cost item, the expert estimates of the optimistic, most likely and pessimistic value of the following variables:
  - Quantity Q (number of units)
  - Productivity P (h/unit) (**)
  - Labor price change rate s_L (% per year)
  - Equipment price change rate s_E (% per year)
  - Material price change rate s_M (% per year)

(*) Although "in-house" construction for large projects is rare, we have chosen this case because it is the most complete and illustrative.

(**) As in most of the actual cases, here we consider that labor and equipment productivities are equal (P_i=M_i) for all the items.

- Time span up to activity beginning (quarters)

and the present value of:

- Labor unit cost LU ($/h)
- Equipment unit cost EU ($/h)
- Material unit cost MU ($/h)

That list also shows, for the overhead cost items, the optimistic, most likely and pessimistic cost in present or constant dollars and current dollars.

Correlation was supposed to exist between some variables. It was defined by correspondences between ranges of variation ("discriminate sampling") as shown in table 7.3.

The historical distribution for similar projects, after smoothing, is given in table 7.4 (inflation not considered) and in table 7.5 (inflation considered).

7.2 Method

Given the number of items and the assumed correlation, the most suitable procedure to propagate uncertainty is the Monte Carlo Method. A computer program was prepared, with which 1000 simulations were carried out for both cases (inflation and no-inflation considered).

In order to compare the results, another two runs were done considering all the variables independent.

Triangular distributions were assumed for all the input variables. The optimistic and pessimistic values were considered as the 5 and 95 percentiles.

In considering inflation, we have used the procedure described in section 5.5, based on distributing the work of each item uniformly over its duration.

Once the specific distribution is calculated, the updating with the historical distribution is done, and the combined final distribution of the ratio R (= actual cost/ estimated cost) is then obtained.

7.3 Discussion of the Output

Figures 7.1 and 7.2 show the specific, historical and updated distribution for the cases in which inflation is not and is considered, respectively.

Due to the correlations involved, the specific distribution is skewed to the right. The skewness is then further strengthened by the updating.

The dispersion, measured by the standard deviation, is higher in the inflation case (figure 7.2) as it can be expected, due to the additional variability introduced by that factor.

The results can be summarized as follows:

- Inflation not considered:

Expected total cost of the project

$= \mu_R \times ETC$

$= 1.102 \times 64,350,000$

$=$ \$70,913,700 (constant dollars)

standard deviation $= \sigma_R \times ETC$

$= 0.050 \times 64,350,000$

$=$ \$3,217,500

Risk of exceeding approved budget = P[R > 1]

= 97.5%

- Inflation considered:

Expected total cost of the project

= μ_R x ETC

= 1.330 x 64,350,000

= $85,585,500 (current dollars)

Standard deviation = σ_R x ETC

= 0.070 x 64,350,000

= $4,504,500

Risk of exceeding approved budget = P[R > 1]

= 100%

Risk of exceeding approved budget + contingency

= P[R > 1.30]

= 67.6%

Required contingency to reduce the risk up to a 25%

= 38% of approved budget.

When all the variables are assumed independent, the specific distribution is symmetrical and close to the normal, as shown in figures 7.3 and 7.4. The updating produces also a little skewness to the right due to the particular shape of the historical distribution considered.

Table 7.1

Work Breakdown Structure

- Direct Cost Items -

#	Description	Unit
1	Clearing & Grubbing	sy
2	Earth Excavation	cy
3	Rock Excavation	cy
4	Embankment	cy
5	Subgrade	cy
6	Base Course	cy
7	Binder Course	ton
8	Wearing Course	ton
9	Granular Fill	cy
10	Concrete Class A	cy
11	Concrete Class B	cy
12	Reinforcing Steel	lb
13	Lateral Drains	lf
14	Line Painting	lf

- Overhead Cost Items -

1'	Mobilization
2'	Indirect Salaries
3'	Office Expenses
4'	Taxes & Insurance

Table 7.2

Input list

* DIRECT COST ITEMS *

ITEM 1

$Q_1 =$	750000.0	800000.0	900000.0	(m^2)
$P_1 =$	0.00140	0.00160	0.00180	(h/m^2)
$s_{L_1} =$	8.0	10.5	12.0	(% per year)
$s_{E_1} =$	7.0	8.0	9.5	(% per year)
$s_{M_1} =$	7.0	8.0	9.0	(% per year)
$T_1 =$	0.0	0.5	1.0	(quarters)
$LU_1 =$	84.00	($/h)		
$EU_1 =$	241.00	($/h)		
$MU_1 =$	0.00	($/h)		

ITEM 2

1600000.0	1750000.0	1950000.0
0.00285	0.00293	0.00305
8.0	10.5	12.0
7.0	8.0	9.5
7.0	8.0	9.0
0.7	1.0	1.2
410.00		
1205.00		
0.00		

ITEM 3

200000.0	450000.0	550000.0
0.00538	0.00568	0.00598
8.0	10.5	12.0
7.0	8.0	9.5
10.0	11.0	12.0
3.0	3.5	4.0
540.00		
1940.00		
0.00		

ITEM 4 .

1150000.0	1300000.0	1500000.0
0.00384	0.00394	0.00405
8.0	10.5	12.0
7.0	8.0	9.5
7.0	8.0	9.0
0.7	1.0	1.2
370.00		
870.00		
1.10		

ITEM 5

490000.0	520000.0	560000.0
0.00300	0.00308	0.00321
8.0	10.5	12.0
7.0	8.0	9.5
10.0	11.0	12.0
3.5	4.0	5.0
440.00		
1260.00		
1.45		

ITEM 6

400000.0	435000.0	495000.0
0.00540	0.00610	0.00680
8.0	9.0	10.0
8.0	8.5	9.5
10.0	11.5	13.0
5.5	6.2	6.9
450.00		
760.00		
5.40		

ITEM 7

110000.0	115500.0	122000.0
0.01900	0.01940	0.02000
8.0	10.0	11.0
8.0	8.5	9.5
11.0	12.5	14.0
7.2	7.8	8.1
210.00		
200.00		
20.10		

ITEM 8 .

113000.0	118000.0	125000.0
0.01900	0.01940	0.02000
8.0	10.0	11.0
8.0	8.5	9.5
11.0	12.5	14.0
8.2	8.6	9.4
190.00		
200.00		
18.50		

ITEM 9

1500.0	1700.0	1900.0
0.46000	0.56000	0.66000
8.0	10.5	12.0
7.0	8.0	9.5
10.0	11.0	12.0
1.0	1.5	2.0
38.00		
40.00		
4.40		

ITEM 10

5600.0	6150.0	6900.0
0.28000	0.31000	0.37000
6.0	8.0	11.0
7.0	8.5	9.0
13.0	14.0	15.0
0.0	0.3	0.5
310.00		
85.00		
240.00		

ITEM 11

860.0	930.0	1000.0
1.90000	2.06000	2.20000
6.0	8.0	11.0
7.0	8.5	9.0
13.0	14.0	15.0
0.0	0.3	0.5
16.00		
20.00		
195.00		

ITEM 12

340000.0	372000.0	410000.0
0.00505	0.00515	0.00530
6.0	8.0	11.0
7.0	[illegible].0	[illegible].0
12.0	13.5	15.5
0.0	0.3	0.5
120.00		
10.00		
0.24		

ITEM 13

710000.0	722000.0	742000.0
0.00290	0.00310	0.00350
8.0	10.5	12.0
7.0	[illegible]	[illegible].5
13.0	14.0	15.0
7.0	8.0	9.0
280.00		
885.00		
4.50		

ITEM 14

1050000.0	1090000.0	1190000.0
0.00775	0.00787	0.00800
8.0	10.0	11.0
10.0	11.0	12.5
12.0	14.0	16.0
9.5	10.2	11.0
16.00		
10.00		
0.07		

* OVERHEAD COST ITEMS * (constant $)

ITEM	1′	C ($)=	[illegible]30000.0	870000.0	900000.0
ITEM	2′	C ($)=	6200000.0	6550000.0	6900000.0
ITEM	3′	C ($)=	46[illegible]0000.0	5350000.0	5700000.0
ITEM	4′	C ($)=	680000.0	720000.0	770000.0

* OVERHEAD COST ITEMS *(current $)

ITEM	1′	C ($)=	920000.0	967000.0	990000.0
ITEM	2′	C ($)=	6900000.0	7256000.0	7700000.0
ITEM	3′	C ($)=	5300000.0	5805000.0	6200000.0
ITEM	4′	C ($)=	750000.0	800000.0	850000.0

Table 7.3

Assumed correlations

Q_2	Q_3	
1600000 - 1750000	450000 - 550000	(Negative dependence)
175000 - 1950000	200000 - 450000	

Q_7	Q_8
110000 - 115500	113000 - 118000
115500 - 122000	118000 - 125000

Q_7	Q_6
110000 - 115500	400000 - 435000
115500 - 122000	435000 - 495000

P_2	P_4
.00285 - .00293	.00384 - .00394
.00293 - .00305	.00394 - .00405

And full positive dependency (correlation +1)

between

s_{L_1} and s_{L_2}, s_{L_3}, s_{L_4}, s_{L_5}, s_{L_9}, $s_{L_{13}}$

s_{L_7} and s_{L_8}

$s_{L_{10}}$ and $s_{L_{11}}$

s_{E_1} and s_{E_2}, s_{E_3}, s_{E_4}, s_{E_5}, s_{E_9}, $s_{E_{13}}$

s_{E_7} and s_{E_8}

$s_{E_{10}}$ and $s_{E_{11}}$

s_{M_7} and s_{M_8}

$s_{M_{10}}$ and $s_{M_{11}}$

T_2 and T_4

T_{10} and T_{11}, T_{12}

Portions of the text
on the following page(s)
are not legible in the
original.

Table 7.4

* HISTORICAL DISTRIBUTION OF "R" *
(inflation not considered)

INTERVAL	%
0.800-0.825	0.26
0.825-0.850	0.66
0.850-0.875	1.05
0.875-0.900	1.58
0.900-0.925	2.11
0.925-0.950	2.64
0.950-0.975	3.16
0.975-1.000	3.82
1.000-1.025	4.48
1.025-1.050	5.40
1.050-1.075	6.32
1.075-1.100	7.38
1.100-1.125	8.17
1.125-1.150	8.70
1.150-1.175	8.70
1.175-1.200	8.43
1.200-1.225	7.64
1.225-1.250	6.72
1.250-1.275	5.40
1.275-1.300	3.82
1.300-1.325	2.24
1.325-1.350	1.05
1.350-1.375	0.26

Table 7.5

* HISTORICAL DISTRIBUTION OF "R" *
(inflation considered)

INTERVAL	%
0.950-0.975	0.21
0.975-1.000	0.53
1.000-1.025	0.74
1.025-1.050	1.05
1.050-1.075	1.26
1.075-1.100	1.68
1.100-1.125	2.00
1.125-1.150	2.42
1.150-1.175	2.84
1.175-1.200	3.16
1.200-1.225	3.68
1.225-1.250	4.21
1.250-1.275	4.63
1.275-1.300	5.05
1.300-1.325	5.58
1.325-1.350	5.89
1.350-1.375	6.21
1.375-1.400	6.53
1.400-1.425	6.74
1.425-1.450	6.63
1.450-1.475	6.32
1.475-1.500	5.68
1.500-1.525	4.95
1.525-1.550	4.00
1.550-1.575	3.05
1.575-1.600	2.11
1.600-1.625	1.47
1.625-1.650	0.95
1.650-1.675	0.42

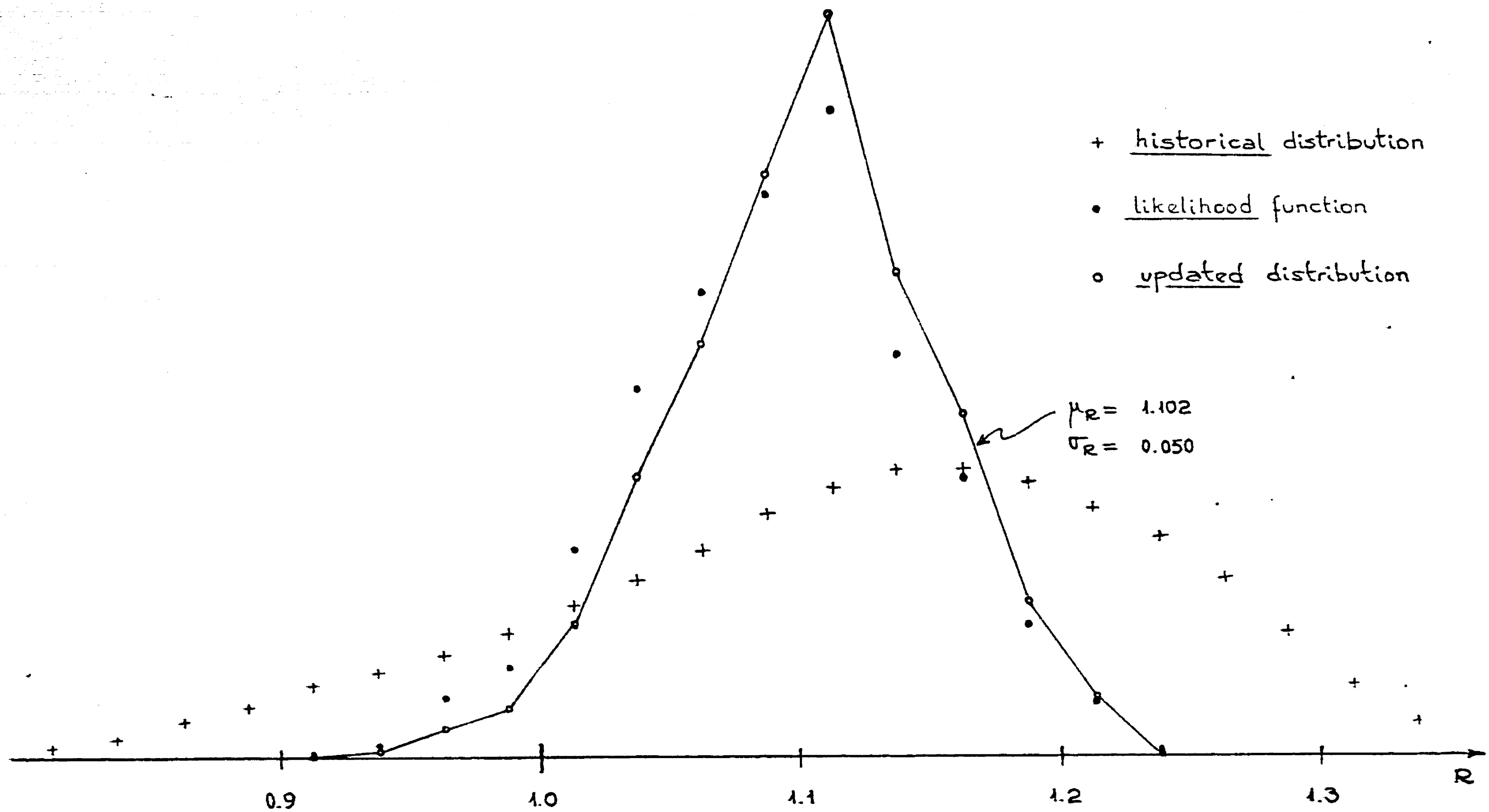


Figure 7.1

Distribution of the ratio R

(inflation not considered)

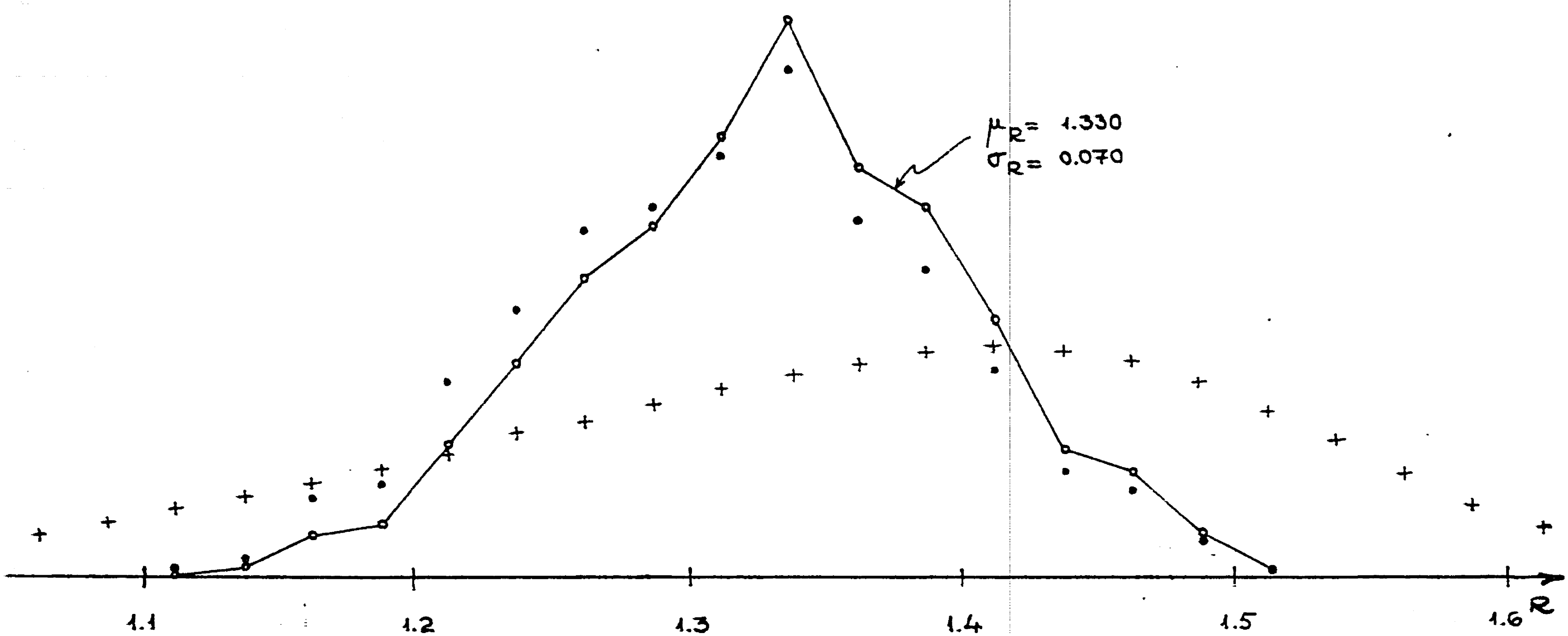


Figure 7.2

Distribution of the ratio R

(inflation considered)

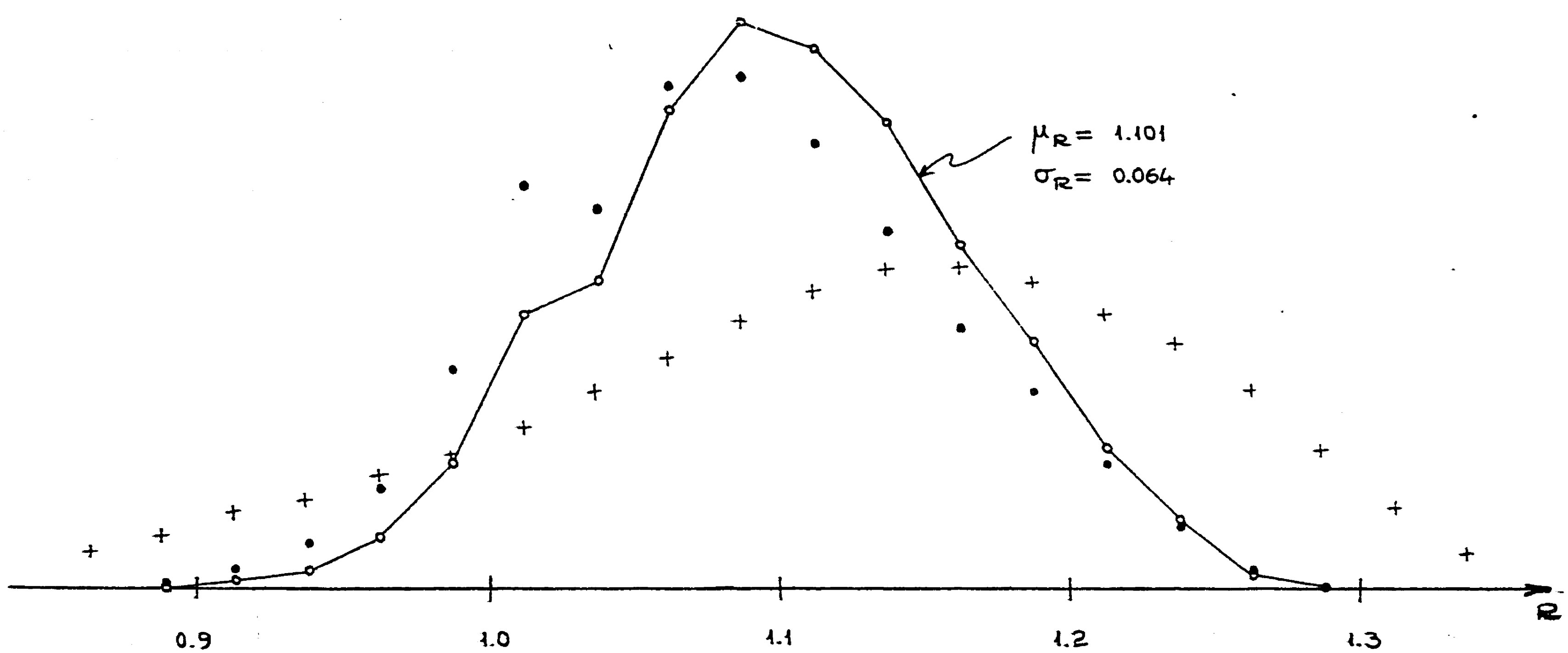


Figure 7.3

Distribution of the ratio R

(inflation & correlation not considered)

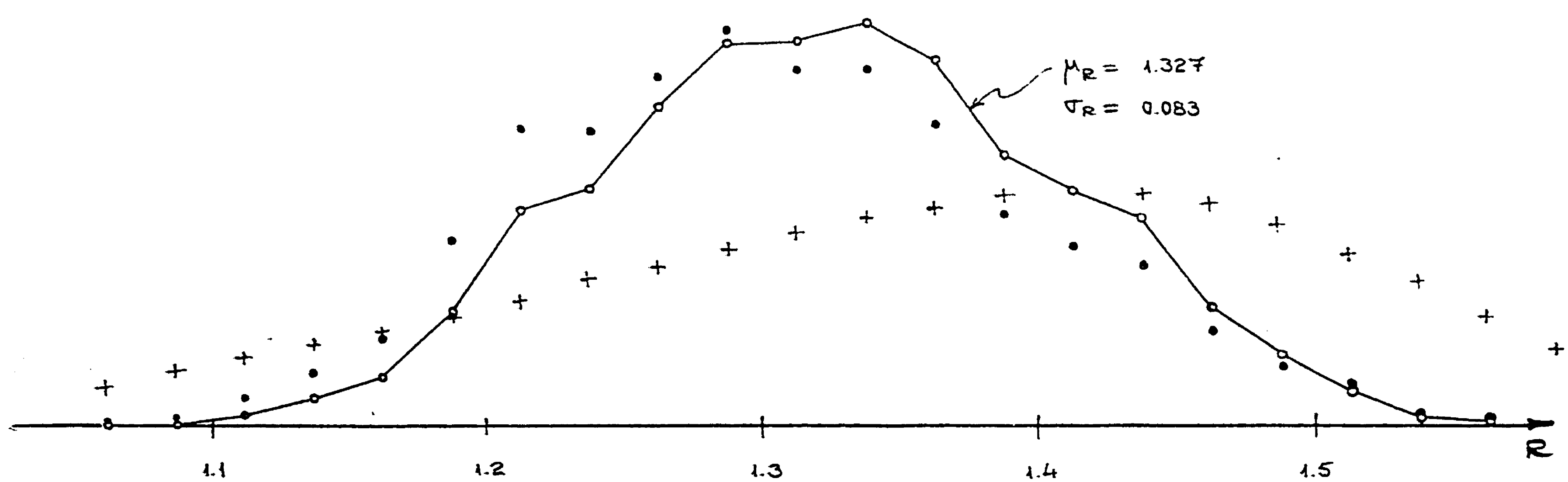


Figure 7.4

Distribution of the ratio R

(inflation considered, correlation not considered)

7.4 Conclusion

We have shown that cost overruns are associated with almost every construction project. Underlying causes behind them are both exogenous and endogenous, and are subject to different degrees of control. Previous studies attempting to quantify risk of cost overruns include assumptions and simplifications that make the results generally unreliable and of doubtful practical application.

The risk borne by the owner depends on the contractual arrangement established for the project. The highest uncertainty in the final cost lies in the "in-house" construction arrangement and the negotiated cost-plus-fee contract type. The smallest resides in the unit price and single fixed-price contract types. The smaller risk is offset by the premium charged by the contractor.

The "total" uncertainty results from contributions due to every component, that is, the quantity, productivity, price of each item. There exist three common methods to "propagate" uncertainty: the method of moments, the method of discrete probability distributions and the Monte Carlo method. When the components are independent or weakly correlated, the method of moments is perhaps best suited--one just has to calculate the mean and the variance of the resulting total cost using "ad hoc" formulae, and then adopt the normal distribution to obtain probabilities. In case of strong dependency, it is advisable to use either the method of discrete probability distributions (after defining the corresponding sets of conditional probabilities) or, better yet, the Monte Carlo method together with the so called "discriminate sampling" technique.

Inflation is an important factor behind cost overruns and, of course, it should not be ignored. Inflation introduces high variability (risk) in the final cost figure.

The question of choice for the distribution of input components is quite irrelevant. Several studies using Monte Carlo simulation have shown that the output distribution depends very little on the assumed input distributions, provided the number of components is reasonably large (as is common in construction projects) and no single contribution dominates over all others. In the case study, the triangular distribution was chosen for simplicity.

It is likely that the owner and his organization behave in a similar way and use analogous procedures in each project. It is likely also that projects of the same type "carry" similar uncertainty. Hence, past experience is a valuable source of information. Bayes' Theorem allows combining historical data of similar projects (which incorporate, through past owner's performance, organizational and project-type intrinsic aspects) and specific data about the particular project under analysis.

The resulting distribution of the cost factor or ratio R (= actual cost/estimated cost) can and should be updated at different points in time as construction progresses and new and more accurate data becomes available. As the case study shows, the distribution of the ratio R is dependent upon the correlations considered. Given the different positive and negative correlations, the distribution shape and width are generally unpredictable. In our particular case, the distribution was sharper than the normal and skewed. Inflation produced an important impact by spreading the distribution significantly.

The calculated risk of not meeting the estimated or budgeted cost and the contingency to be allocated depend thus, to a large extent, on whether inflation is included and whether the correlations are correctly stated.

The procedures described have practical application to real projects as the case study suggests. A company may construct its own "historical" distribution of the ratio R, based on observed actual to estimated final costs of completed projects. Ideally, a distribution should be constructed for every group of projects of similar type and size. A default distribution would be one based on data published for projects completed within that geographical area (indeed, this information should always somehow be incorporated). The "likelihood" function may be obtained through expert estimates at various points in time during project construction.

Private entities and public agencies may find this methodology useful. It can be used as an internal tool to know the project risk so as to take corrective steps on time, thus avoiding late "surprises". In addition, it may help support the petition of funds from financial institutions.

Further research is needed on the following:

. Reliability of the optimistic, most likely and pessimistic estimates (avoidance of subjective biases).

. The appropriate level of disaggregation in the break-down structure, mainly regarding the overhead items, in order to achieve a good balance between accuracy and simplicity.

. The "premium" an owner (with a certain risk aversion) is willing to pay for shifting part of the risk to the contractor (using a different contractual arrangement).

. An integrated and computer-based approach to cost and time overrun risk quantification.

BIBLIOGRAPHY

Alexander, A.J., and Nelson, J.R., Measuring Technological Change: Aircraft Turbine Engines, The Rand Corporation, R-1017-PR, June 1972.

Barrie, D.S., and Paulson, B.C., Professional Construction Management, McGraw-Hill Book Co., Inc., 1978.

Bjornsson, H.C., "Risk Analysis of Construction Cost Estimates", Transactions American Association of Cost Engineers, 1977.

Carrier, K.C., "A System for Managing Escalation and Contingencies", Transactions American Association of Cost Engineers, 1977.

Dickmann, J.E., Mayer, R.H., and Stark, R.M., "Coping with Uncertainty in Unit Price Contracting", Journal of the Construction Division, ASCE, Vol. 108, No. CO3, September 1982.

Eilon, S., "Sampling Procedures for Risk Simulation", Operational Research Quarterly, Vol. 24, No. 2, 1973.

Fischer, S., and Dornbusch, R., Introduction to Macroeconomics, McGraw-Hill Book Co., Inc., 1983.

Gates, M., "Bidding Contingencies and Probabilities", Journal of the Construction Division, ASCE, Vol. 97, No. CO2, Nov. 1971.

Harman, A.J., and Henrichsen, S., A Methodology for Cost Factor Comparison and Prediction, The Rand Corporation, RM-6269-ARPA, August 1970.

Harrison, F.L., Advanced Project Management, John Wiley & Sons, New York, 1981.

Hufschmidt, M.M., and Gerin, J., "Systematic Errors in Cost Estimates for Public Investment Projects", in J. Margolis (ed.), The Analysis of Public Output, Columbia University Press, New York, 1970.

Kaplan, S., et al., Methodology for Probabilistic Risk Assessment of Nuclear Power Plants, Picard, Lowe and Garrick, Inc., Irvine, California, 1981.

Kohen, E., "Cost of Delays in Construction", Journal of the Construction Division, ASCE, Vol. 104, No. CO3, Sept. 1978.

Marshall, A.W., and Meckling, W.H., Predictability of the Costs, Time and Success of Development, The Rand Corporation, p. 1821, December 1959.

Mead, W.J., Rogers, G.W. and Smith, R.Z., Transporting Natural Gas from the Artic, American Enterprise Institute for Public Policy Research, Washington, D.C., 1977.

Mendel, O., "Risk Management in Construction: The Contractor's Viewpoint", Transactions American Association of Cost Engineers, 1977.

Merewitz, L., "Cost Overruns in Public Works", in Benefit-Cost and Policy Analysis 1972, Aldine Publishing Co., Chicago, 1973.

Merewitz, L., and Sparks, T.C., "A Disaggregated Comparison of the Cost Overrun of the San Francisco Bay Area Rapid Transit District", Working Paper No. 156/BART 3, Institute of Urban and Regional Development, University of California, Berkeley, May 10, 1971.

Merrow, E.W., Chapel, S.W., and Worthing, Ch., A Review of Cost Estimation in New Technologies, The Rand Corporation, R-2481-DOE, July 1979.

Moder, J.J., Phillips, C.R., and Davis, E.W., Project Management with CPM, PERT and Precedence Diagramming, Van Nostrand Reinhold Co., Inc., New York, 1983.

Moder, J.J., and Rodgers, E.G., "Judgment Estimates of the Moments of PERT Type Distributions", Management Science, Vol. 15, No. 2, October 1968.

Murphy, D.C., Baker, B.N., and Fisher, D., Determinants of Project Success, N-74-30392, National Technical Information Services, Springfield, VA, Sept. 15, 1974.

Ostwald, P.F., Cost Estimating for Engineering and Management, Prentice-Hall, Englewood Cliffs, N.J., 1974.

Pratt, J.W., Raiffa, H., and Schlaifer, R., Introduction to Statistical Decision Theory, McGraw-Hill Co., Inc., New York, 1965.

Smith, S.E., et al., "Contractual Relationships in Construction", Journal of the Construction Division, ASCE, Vol. 101, No. CO4, Dec. 1974.

Spooner, J.E., "Probabilistic Estimating", Journal of the Construction Division, ASCE, Vol. 100, No. CO1, March 1974.

Spooner, J.E., "The Management of Contingency for Multiple Capital Projects", 1975 Proceedings of the Project Management Institute.

Stukhart, G., "Inflation and the Construction Industry", Journal of the Construction Division, ASCE, Vol. 108, No. CO4, Dec. 1982.

Summer, R., Cost Estimates as Predictors of Actual Weapon Costs: A Study of Major Hardware Articles, The Rand Corporation, RM-3061-PR, March 1965.

Tucker, J.F., "Cost Estimation in Public Works", Master of Business Administration thesis, University of California, Berkeley, Sept. 1970.

Warszawski, A., "Risk Element in Profit Planning Under Inflation", Journal of the Construction Division, ASCE, Vol. 108, No. CO4, Dec. 1982.

Williamson, R.S., Donnel, W.R., and Browder, J.F., "Estimate Contingency, Risk, and Accuracy--What do they mean?", Transactions American Association of Cost Engineers, 1980.