Practical Multifactor Approach to Evaluating Risk of Investment in Engineering Projects

Abraham Warszawski, F.ASCE,[1] and Rafael Sacks[2]

Abstract: Risk analysis is crucial in enabling management to make informed decisions regarding the economic viability of engineering projects. In most building construction projects, sophisticated risk assessment methods are not used because the detailed input information they require is unavailable to the average project owner or manager. As a result, risk assessment is often limited to simple sensitivity analyses. This paper presents a practical yet thorough method in which the economic risk inherent in a construction project can be calculated with input information of varying levels of detail. The proposed ''multifactor'' method includes consideration of interdependence between a project's risk factors. The principles of the method are explained, its application to a large construction project is illustrated, and the findings are discussed.

DOI: 10.1061/(ASCE)0733-9364(2004)130:3(357)

CE Database subject headings: Risk analysis; Investments; Economic factors; Probabilistic models; Project management; Construction industry; Decision making.

Introduction

The decision whether to proceed with any construction project (the go–no go decision) hinges on the project owner's assessment of the prospects for profit or loss. This decision is based on many factors, including some that cannot be predicted with certainty. '''Risk analysis'' is the process of assessing decision variables which are subject to risk and uncertainty' (Edwards and Bowen 2000). ''Uncertainty'' is defined as a 'state of incomplete knowledge about a decision variable.' ''Risk'' implies the occurrence of an undesirable result (i.e., incurring a loss)—it is 'the chance (probability) that an adverse event occurs during a stated period of time' (Edwards and Bowen 2000).

The larger the scale and complexity, and the longer the time frame of an engineering project, the greater will be the impact of risk, since the degree of uncertainty and severity are harder to assess. In the context of this paper, risk is defined as the probability that an owner's investment in a construction project might not attain the required measure of profitability. The risk profile is the distribution of the probability of different profitability levels being attained in the project. A construction project is idealized as a set of interdependent construction and commercial activities, each with an associated payment (a cost or a revenue) and duration. These are termed the project's risk factors. The economic risk in a construction project can be expressed in terms of the uncertainty of the risk factors; risk assessment is the process of evaluating the risk profile subject to the variability of the risk factors. Risk management includes risk assessment and consequent manipulation of the risk factors in such a manner that the project risk will not exceed an acceptable level.

In construction and real estate development, the sources of uncertainty in risk factors are many and varied. They may be technical (design changes and errors, estimation error), managerial (productivity, cost control), economic (resource supply, inflation, exchange rates, market fluctuations), political, financial, legal, and natural (weather and geology) to name but a few (Ashley and Bonner 1987; Yeo 1990; Edwards and Bowen 2000; Tah and Carr 2000).

Probabilistic treatments of both time and cost attributes of individual activities, and their impact on the overall duration and cost of a whole project, have received considerable attention in project modeling since the introduction of the program evaluation and review technique (PERT) in the late 1950s (Malcolm et al. 1959). For the sake of brevity, only the most characteristic contributions in this respect are mentioned in the following review.

The probabilistic aspects of project duration were introduced by Malcolm et al. (1959) in PERT. The results of the PERT technique were subsequently shown to be consistently optimistic (MacCrimmon and Ryavec 1964). More advanced analytical methods in this respect have been offered, such as the probabilistic network evaluation technique (PNET) (Ang et al. 1975); these were later modified and applied by Ranisinghe and Russell (1993) and Ranisinghe (1994). Keefer and Bodily (1983), Newton (1992), and others explored simulation techniques for risk analysis in construction. The controlled interval and memory (CIM) method, a decision-tree enumeration approach, has also been proposed (Cooper and Chapman 1987).

Risk in project cost has generally been treated using the same basic approaches—analytical, simulation and decision-tree enumeration—employed in the above-mentioned treatments of risk in project duration (Diekmann 1983; Newton 1992; Touran 1993; Ranisinghe 1994).

Correlations between different project activities, in terms of both costs and durations, have been incorporated in a limited

[1]Professor, Faculty of Civil and Environmental Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel.

[2]Lecturer, Faculty of Civil and Environmental Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel.

Note. Discussion open until November 1, 2004. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on June 5, 2001; approved on March 5, 2003. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 130, No. 3, June 1, 2004. ©ASCE, ISSN 0733-9364/2004/3-357–367/$18.00.

number of efforts (Ranisinghe 1994). For example, Touran (1993) used a correlation matrix to weight the random values for each correlated cost item in a Monte Carlo simulation. Wang and Huang (2000) proposed considering risk dependencies between cost and time, not only between cost items. In the CIM method (Cooper and Chapman 1987) risk dependencies are accounted for by explicitly expressing the discrete probabilities for a primary variable for each discrete possible value of another primary variable lower down in the tree. In general, if dependencies between factors are neglected, the resulting mean value of an assessment is not affected, but the variance— and therefore the risk—is underestimated (Diekmann 1983).

The main problem in risk assessment is the difficulty of obtaining information about the variability of the risk factors. Construction projects are generally unique, and design and construction teams change from project to project. Records are seldom kept in a retrievable manner; and when they are, the differences in project teams and the unique technical makeup of any new individual project, make the application of historic data difficult. Intensive investment in efforts to establish the nature and scale of the risk factors with any precision, as is common in large-scale capital projects, are typically well beyond the financial means of the majority of commercial or residential building construction project teams.

Conventional sensitivity analysis, still probably the most widely applied procedure for the examination of risk in building construction projects, has three important drawbacks:

1. It examines the risk impact of each risk factor separately, while in reality they vary simultaneously (with compound effect);
2. It considers risk factors to be mutually independent, while in reality, risk factors are very often interdependent; and
3. The change of each risk factor is assumed to be linear, i.e., the value of the parameter over all its range has the same probability of occurrence, while in most real life cases the values at mid-range have a higher probability of occurrence than those at the margins.

The aim of this study was to develop a straightforward method for the assessment of risk in building construction investment projects, without the drawbacks of sensitivity analysis, yet employing similar levels of input information and of management resources. Given the extremely limited resources typically allocated to risk assessment for such projects, the following objectives were emphasized:

- The ability to operate, if necessary, with only the basic information that is readily available to the decision makers;
- The ability to perform the assessment with standard and readily available computing hardware;
- The assessment procedures must be transparent and easily understandable to users, yet allow representation of realistically varying project circumstances; and
- The ability to assess the potential impact of management actions on the project risk, through controlled manipulation of the duration and cost attributes of individual activities.

This paper presents a risk evaluation method that achieves the above objectives. The authors first review the accepted profitability measure of a project based on its discounted cash flow, using deterministic "most likely" estimates of its parameters. Subsequently, they examine the behavior of the cash flow parameters as random variables and the effect of their variability on the risk of the total project. Two cases are examined: first, when the risks inherent in different activities are assumed independent of each other, and second, when interdependence is assumed to exist. Application of the method using purpose-built software is illustrated by a comprehensive example, and conclusions are drawn and discussed.

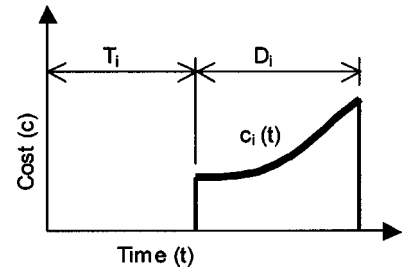

Fig. 1. Timing and cost of activity *i*

Profitability Measure—Net Present Value of Project Cash Flow

The net present value (NPV) of a project's cash flow is usually the preferred criterion of its profitability since it reflects the net contribution of the project to the owner's equity in due consideration of his cost of capital. The cash flow includes payments—negative and positive—pertaining to the project's activities. A negative payment is a cash disbursement associated with the cost of an activity. A positive payment is the cash receipt for the completion of a predefined project stage, associated with the completion of a certain activity. In the most general case, a payment associated with an activity may be distributed over the duration of the activity extending between the times of its start and its finish, respectively, as shown in Fig. 1. The payment can also be represented conveniently by a single equivalent payment at the start, finish or mid duration of the activity.

The profitability of a project is measured using the net present value of its cash flow, discounted by the investor's cost of capital. The net present value can be obtained from either Eq. (1), with continuous compounding, or from Eq. (2), with discrete compounding

$$\mathrm{NPV}=\sum_i e^{-rT}*\int_0^{D_i} c_i(t)e^{-rt}dt=\sum_i e^{-rT}*C_i$$

with

$$C_i=\int_0^{D_i} c_i(t)e^{-rt}dt \tag{1}$$

at the start.

$$\mathrm{NPV}=\sum_i C_i(1+r)^{-(T_i+D_i/2)}$$

with

$$C_i=\sum_0^{D_i} C_{ij}(1+r)^{D_i/2-j} \tag{2}$$

at mid-duration, where T_i = early start of activity i (considering the duration of preceding activities); $c_i(t)$ = cost of activity i as a continuous function of t; C_i = discrete equivalent cost at the beginning or mid-duration of activity i; D_i = duration of activity i; and r = interest rate per period.

In an ordinary cash flow analysis, the costs and durations of activities are considered to have deterministic values, based on

the owner's "most likely" realistic estimate. In reality, costs and durations vary. The wider the range of variation of any cost or duration, the greater its impact on the overall project risk.

Calculating Net Present Value Cumulative Probability Distribution

The accepted approach to expressing risks, to allow their mathematical combination into an overall project risk assessment, is to determine a probability distribution for the possible values of each risky decision variable. Traditionally, risky costs or durations have been expressed using standard distributions. These are commonly drawn from the Pearson system of distributions (Johnson et al. 1963), including normal, student's T, beta, and other distributions. The main reason for this is that such distributions can be completely described by two to four parameters (usually the mean μ, the standard deviation σ, the skewness β_1, and the kurtosis β_2), which has the advantage of allowing analytical combination of like distributions, using appropriate formulas (Kottas and Lau 1982; Keefer and Bodily 1983). The CIM method (Cooper and Chapman 1987) makes use of discrete distribution functions; the discrete possible results of each primary variable are laid out as branches of a tree, which then branch again for each additional primary variable.

In almost all cases, the analyst or a domain "expert" sets the type and values of the distributions subjectively. This is due to the lack of recorded data describing such risks (Ranisinghe and Russell 1993). Elicitation of coherent and consistent values from experts is extremely difficult (Cooper and Chapman 1987). It is the writer's experience (reported by Benjo 1999) that elicitation of any risk input values from practitioners, (beyond the most likely value and the range of its variation) is unreliable, and sometimes even misleading, unless there is sufficient recorded data to support it. The subjective assessments tend to be optimistic. The norm is that construction projects suffer both cost and schedule overruns compared to initial estimates. Morris and Hough (1987), for example, found consistent cost overruns of 40–500% in an extensive survey of major public sector projects in the United States, United Kingdom, and other countries.

Proposed Multifactor Method

The multifactor method proposed here consists of the steps described in Fig. 2. Before beginning, the project must first be described as a set of activities, each with duration and payment values, and the precedence relationships between them, as might be done for a critical path method network evaluation. The activities may be technological (with payment and duration), purely financial (with zero duration) or ancillary (with zero payment, e.g., permit review).

The risk evaluation procedure begins with identification of the significant risk factors and assessment of the range and pattern of their variation. A risk factor is considered significant if its impact on the variability of the result, either alone or because of the dependence of other risk factors on it, is likely to be profound. This subjective threshold is set by the risk evaluator in each case based on the overall accuracy required, the managerial resources allocated to the risk analysis, the information available, and the potential for risk management. Conventional sensitivity analysis may aid in examining the effect of any particular factor on the total project risk, but effective selection depends finally on the evaluator's experience and judgment. The validity of the selection is tested later in the procedure, and can be corrected if necessary.

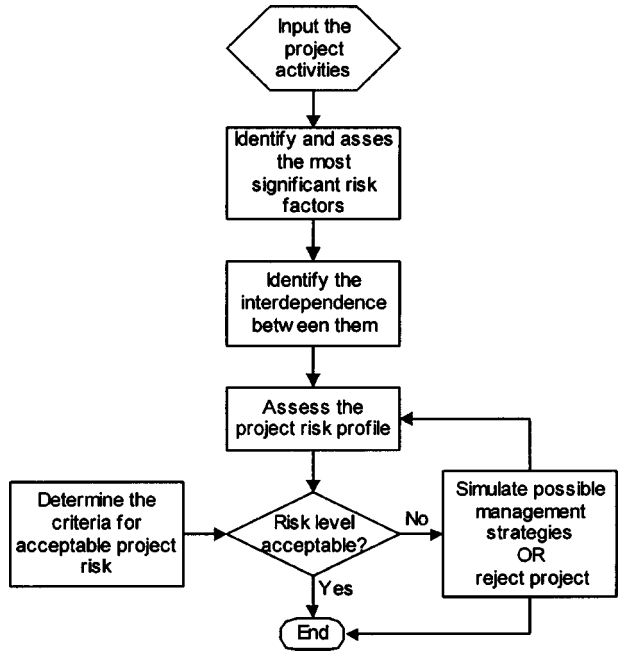

Fig. 2. Proposed risk assessment procedure

Next, if interdependence between any pair of significant risk factors is assumed, it must be defined on a linear scale from 0 to 1. Lastly, the evaluator should determine a criterion for acceptable project risk and express it as the probability that some measure of profitability of the project will be attained. A positive net present value (NPV$>$0) of the project cash flow is convenient, although other measures (such as NPV exceeding an arbitrary critical value, internal rate of return exceeding a desired value, etc.) can also be used. The risk profile is then evaluated using appropriate software.

The resulting project risk profile indicates the probability of the NPV exceeding different values in the range of NPV_{min} and NPV_{max} for the project, as shown in Fig. 3 for the case of NPV >0. If the criterion for project approval is not met, the evaluator can simulate possible risk management strategies by manipulating the values of any of the significant risk factors and rerunning the evaluation until an acceptable risk level is achieved.

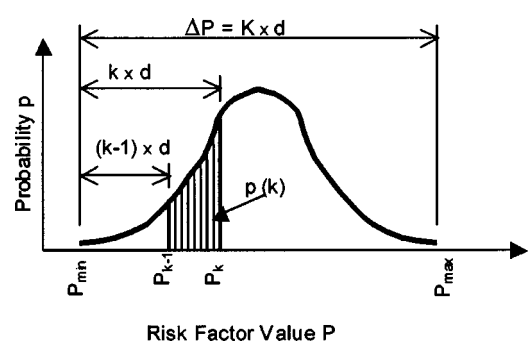

Fig. 3. Probability density function interval

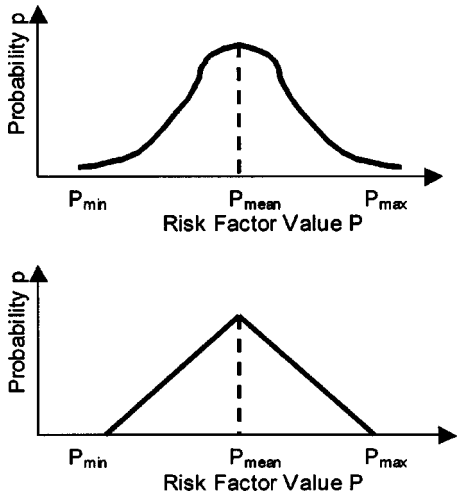

Fig. 4. Normal and triangular probability distributions

In the following sections, various aspects of the procedure are explained in detail. The general method is presented first, and is then extended to include consideration of interdependence between risky activities.

Assessing Variability of Risk Factors

Evaluation of the risk profile of a project requires that one first determine the pattern of variation of its significant risk factors. The uncertainty inherent in a factor—either the cost or the duration of any activity—is represented by a probability distribution through its range of variation. The user can select different probability distributions for different uncertain parameters. Four cases are possible, depending on the degree of detailed knowledge the user has regarding the risk parameter.

Case 1

Detailed information is available. In cases where reliable and pertinent data can be obtained from past records (or from experts), discrete or continuous distributions can be determined. The probability of the parameter being within a given interval can be derived directly from the distribution as shown in Fig. 3.

Case 2

Only the range of variation (i.e., maximum and minimum values) and the most likely outcome can be elicited. Two courses should be distinguished. If the most likely value is perceived at the mid range, a symmetrical distribution will be applied. A normal distribution can be then fitted with its mean μ as the most likely value, and its standard deviation σ as approximately 1/6 (16.6%) of the range. Alternatively, a symmetrical triangular distribution can be selected (with σ approximately 21% of the range) if a less centered distribution is assumed. This distribution implies that the probability of an outcome declines linearly with its distance from the most likely value. Both distributions are shown in Fig. 4. The sensitivity of the project risk to the choice among these distributions will be examined later in this paper.

If the most likely outcome is not at the mid range, a different course may be pursued: one of two available skewed distributions could be used, beta or skewed triangular. For both distributions, the most likely value sets the position of the peak of the distribution. The elicited "most likely" value is assumed to be the mode, not the mean as is consistent with the assumption of simplicity and with most previous projects (Malcolm et al. 1959; Kottas and Lau 1982; Keefer and Bodily 1983; Ranisinghe and Russell 1993).

The triangular distribution is used, as in the symmetric case, for representation of less centered outcomes. For the case of the Beta distribution, two additional parameters must be set: α and β. The analytical expression for the beta function (with α and β both integer) is

$$f(x)=C(x-a)^{\alpha}(b-x)^{\beta}$$

with

$$C=(\alpha+\beta+1)!/\alpha!\beta!(b-a)^{\alpha+\beta+1} \tag{3}$$

where a,b = minimum and maximum values of the range, respectively. The mode and the standard deviation of the function may be calculated from Eqs. (4) and (5), respectively

$$x_o=a+(b-a)[\alpha/(\alpha+\beta)] \tag{4}$$

$$\sigma_x^2=(\alpha+1)*(\beta+1)/[(\alpha+\beta+2)^2*(\alpha+\beta+3)] \tag{5}$$

Because simplicity of use is a priority in the proposed method, the user may simply accept a value for the standard deviation of the function, defaulting to approximately 16% of the range (which is compatible to its value in the Normal distribution). Since a,b and X_o are known, the values of α and β can be selected accordingly.

Case 3

Only the range of variation between maximum and minimum values can be elicited, or the cost of elicitation of the most likely value is not justified (in terms of either its impact or its accuracy). Most simply, one of the two symmetric distributions can be preferred as in the first course of the former case. Alternatively, the user may select an asymmetric distribution and be helped by qualitative descriptions of the skewness to choose a suitable value for the most likely outcome on a scale between 0 and 1, as set out in Table 1.

Case 4

The user is unable to estimate the range of variation. In such case two or three "bounding" ranges for all or part of the variable parameters can be tested. Based on a survey of range variations (Benjo 1999), the maximum deviation from the most likely estimate of cost in a regular construction project should not exceed 10–15% of its value. Consequently, examination of $\pm 10-\pm 20\%$ ranges should suffice. The sensitivity of the project risk profile to the variation range assumed will be examined later in this paper.

Assessment of Project Risk Profile

The risk profile of the project, as noted earlier, indicates the probability of the NPV attaining the stated profitability criterion (Fig. 5). Two alternative approaches can be adopted for the evaluation of the project risk profile, based on the probability distributions of the risk significant variable parameters: enumeration, using a procedure similar to CIM (Cooper and Chapman 1987) or simulation (Monte Carlo).

The enumeration procedure begins with expression of the probability distribution of each risk factor as a set of K discrete

Table 1. Characteristic Values of Beta and Triangular Distributions in Range between 0 and 1

		Beta (normal) distribution				Triangular distribution	
Location of most likely estimate	X_o	α	β	μ	σ_x	μ	σ_x
Very close to minimum	0.20	1	4	0.31	0.15	0.37	0.23
Closer to minimum	0.33	2	4	0.37	0.16	0.40	0.22
Somewhat closer to minimum	0.42	3	4	0.44	0.16	0.47	0.21
At mid range	0.50	4	4	0.50	0.17	0.50	0.21
Somewhat closer to maximum	0.57	4	3	0.56	0.16	0.53	0.21
Closer to maximum	0.67	4	2	0.62	0.16	0.60	0.22
Very close to maximum	0.80	4	1	0.69	0.15	0.63	0.22

probability values, each covering equal subdivisions of the full range. The procedure then enumerates all K'' combinations of the n cost and/or duration risk factors, and evaluates the project NPV and the associated probability of each combination. The results are accumulated and ultimately a cumulative distribution function can be plotted for the project. The full procedure is set out in the Appendix.

A significant drawback of the enumeration approach is that the number of different combinations to evaluate may be very large. For example, for a project with $n=7$ random variables P, varying in a specified range with $K=5$ intervals, the number of possible combinations will be $5^7=78{,}125$. This limits the practical application of the risk analysis to 8–11 parameters that can be evaluated simultaneously. The number of parameters to be evaluated simultaneously can be substantially increased by selective enumeration, as explained later. In practice however, evaluation of those 7–10 parameters that are to be manipulated subsequently for the possible reduction of the project risk, will usually suffice: focusing on a limited number of significant parameters allows a more thorough analysis of the risk reduction alternatives. The impact of the number of risk factors on the reliability of the risk profile is discussed in the context of the case study.

The second alternative for generation of the project risk profile is through inferring from a sample drawn from the whole population of possible outcomes. The Monte Carlo procedure employed for this purpose simulates each ''event'' by assigning a random value of k (or any arbitrary interval in its range) to each variable risk factor. A sample of 5,000–10,000 ''events'' usually suffices for practical purposes.

Both of the above methods have been implemented for this research in a computer program in *Visual Basic*, based on *Microsoft Project* and *Microsoft Excel*. The enumeration procedure is implemented using a recursive algorithm to set the k risk factor values at each level of the probability tree. The function is repeatedly called from within its own body at each progression up the tree, first for all duration risks, at which point the activity precedence network is solved, and then for cost risks, until the last risk factor is reached, at which point the NPV is calculated, and its result registered. The Monte Carlo procedure sets random values for each risk factor at every step, and so must solve the precedence network and calculate NPV for every calculation. The software allows any distribution type to be selected for any risk factor, with any range of values of k for each. Risk dependency may be set between any pair of risk parameters and may be positive or negative.

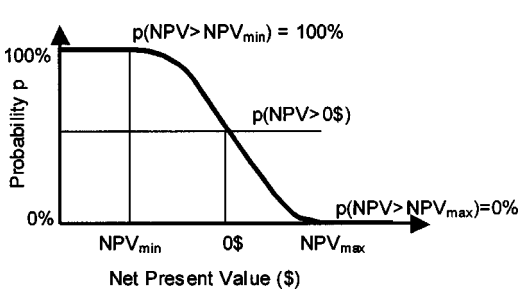

Fig. 5. Risk profile of project

Computational Efficiency

It was pointed out earlier that full enumeration and evaluation of all possible events might involve a large calculation effort. The various algorithmic options for reducing the computational effort are as follows:

1. The number of intervals may be reduced from $K=5$ to $K=3$ for all or most variables. This is reflected in the case study, and will be discussed later. It is also possible to select $K=5$ for more prominent parameters and $K=3$ for others. In general, when using the enumeration approach (see Appendix), the sequence of the computation should be structured so that the risky durations are placed at the base of the tree. In this way, the precedence network need be solved only K^j times, where j is the number of risky durations, rather than K^n times. This strategy was adopted in the software implementation.
2. The project may be decomposed into ''packages'' of activities that may be represented by a single risk pattern. A package in this context is a set of interdependent activities that emanate from a single node in the project network, and terminate at a single node of the project. A package will have its own risk profile derived from the variability of its composite parameters. For the purpose of risk assessment, the package may then interact with the rest of the project in the same manner as a single activity with a defined variability, while it may still be seen as a set of many distinctive activities for managerial planning and control.
3. If the acceptable risk criterion requires NPV>0 at a specified cumulative probability (say 80 or 90%) a Branch and Bound technique may be employed for ''pruning'' all branches which originate from unpromising nodes of the decision tree. Any node in a decision tree (see Fig. 8) may be viewed as a compound event Z' which consists of risk factors which precede the node and whose value has already been determined, and the risk factors below the node whose value has yet to be determined. The upper or lower bounds of a node indicate the minimum and maximum values, respectively, of NPV pertaining to all the branches that originate in this node. If the lower bound of a node is higher than 0, then the sum of probabilities of all the branches originated

in this node is added to $p(\mathrm{NPV}>0)$ and no further evaluation of their NPV is needed. In a similar fashion, if the upper bound on NPV of branches originating in the node is less than 0—then the total probability of the node's branches is added to $p(\mathrm{NPV}<0)$ and no further computation is needed. The bounds are assessed by assigning least favorable (for lower bound) or most favorable (for upper bound) k values to all variables below the node level. The procedure can reduce the amount of computation manifold.

4. A heuristic algorithm can be used, under which the outcome of the evaluation process at each level of m evaluated risk factors is represented by its risk profile. The profile is obtained from variation of these m parameters while the remaining $n-m$ parameters are held constant at their most likely values. The risk profile pertaining to these m parameters is represented by K' intervals for the different values of $\mathrm{NPV}(m)$ in the range of $\mathrm{NPV}_{\min}(m)$ and $\mathrm{NPV}_{\max}(m)$. Each interval k' has a representative vector $G[Pk']$ of its m parameter values attached to it. As the additional variable $Pm+1$ is added, K' interval sets (for the m parameters) times K values for the new parameter are evaluated. The process is continued until $m=n$ with $K'*K$ iterations at each stage or the total of merely $n*K'*K$ iterations instead of K^n in a complete enumeration.
5. A Monte Carlo simulation strategy can be adopted.

The effectiveness, efficiency, and computational aspects of Methods 2–4 have yet to be explored in detail. In the research reported here, simulation proved adequate and efficient for all cases where enumeration was computationally impractical.

Dependence between Risk Factors

Up to this point, it was implicitly assumed that the different risk factors were independent of each other in their variability. It can be expected however that in many cases some dependence will exist between certain activities—whether they affect each other, or whether they depend on some common external risk factor. Dependencies may be linear or nonlinear, positive or negative, or mixed (Cooper and Chapman 1987). For the purpose of the model developed here, it is assumed that dependencies are linear, may be positive or negative, and that a dependent risk factor is correlated to just one other factor. Two particular cases are identified:

1. One or more of the activities are correlated with an external risk factor N (such as weather, availability of labor, political stability). The probability of any internal risk factor having a particular value, $p(k)$, depends on the probability of the external factor having a probability $p(N)$.
2. One or more of the activities are risk dependent on a particular preceding activity in the project. This means that the probability of the result (cost or duration) of activity i_1, is dependent on the result of activity i_2, either entirely or to some extent. Since the external factor N can be viewed as an ancillary activity without duration or cost attributes, case (1) reduces to case (2).

The difficulty in eliciting dependencies between project activities and costs from project owners or managers is even more difficult than eliciting probability distribution densities. In this model, the user is asked to rate the dependence factor ρ between any two risk factors A and B on a simple scale of five levels, and to state whether the dependence is positive or negative. During calculation, the probability that the dependent parameter B will have a certain value of k, is then increased (in proportion to the dependence level ρ) when the parameter A has the same value, or decreased otherwise. The corrected probability of event B having a value in its interval range k_B, $p'(k_B)$, given that event A has a value in its interval range k_A, is calculated using Eq. (6) when $0\leqslant\rho\leqslant 1$, and Eq. (7) when $-1\leqslant\rho<0$

$$p'(k_B)=p(k_B)+\rho*[1-p(k_B)]=\rho+(1-\rho)*p(k_B)$$

when $k_B=k_A$;

$$p'(k_B)=(1-\rho)*p(k_B) \quad \text{when } k_B\neq k_A \tag{6}$$

$$p'(k_B)=-\rho+(1+\rho)*p(k_B) \quad \text{when } K+1-k_B=k_A$$

$$p'(k_B)=(1+\rho)*p(k_B) \quad \text{when } K+1-k_B\neq k_A \tag{7}$$

where k_B, k_A=current interval indices of the dependent parameter B and the independent parameter A, respectively; and $p(k)$, $p'(k)$=prior and revised probabilities, respectively, of parameter B having a value in the interval k.

For sake of simplification, a project owner or manager might select the dependence factor ρ from one of five basic values listed in Table 2. For example, if risk factors A and B each have four ranges, and the ranges of B have the probability values $p_B(1)=p_B(4)=0.1$, $p_B(2)=p_B(3)=0.4$, the corrected values of p_B when B is highly dependent on A (i.e., 75%) and A has its lowest value in the current calculation iteration (i.e., $k_A=1$), then p_B is corrected as in Fig. 6. Note that the condition $\Sigma\, p(k)=1$ is maintained in all cases of Eqs. (5) and (6). The derivation of the total probability distribution for the project proceeds as before, with the correction of probability factors being performed before each NPV calculation.

Table 2. Discrete Risk Dependence Factor Values

Value of dependence factor ρ	Level of dependence
1	Total $p'(k_B)=1$ when $k_B=k_A$ and 0 otherwise
0.75	High dependence (75%)
0.50	Some dependence (50%)
0.25	Slight dependence (25%)
0	Independent

Multifactor Risk Assessment Procedure Summarized

In light of the approach described, the risk assessment procedure can be summarized as follows (refer to Fig. 2):

1. Input the project activities, with the most likely values for the duration and payment for each one, and fix the technological sequential dependencies between them (such as finish–start).
2. Identify the significant cost and duration risk parameters for the project cash flow. For each, set the expected maximum and minimum values, and the required number of discrete sections ($k=3-5$). If the most expected value is known, and is different from the average of the minimum and maximum values, choose a skewed distribution (beta or triangular); if not, choose a symmetrical distribution (normal or triangular). If a detailed distribution is available, enter k discrete values to describe it.
3. Determine whether any risk dependencies exist between the parameters. If so, set the dependence factor for each such relationship (from −1 to +1).

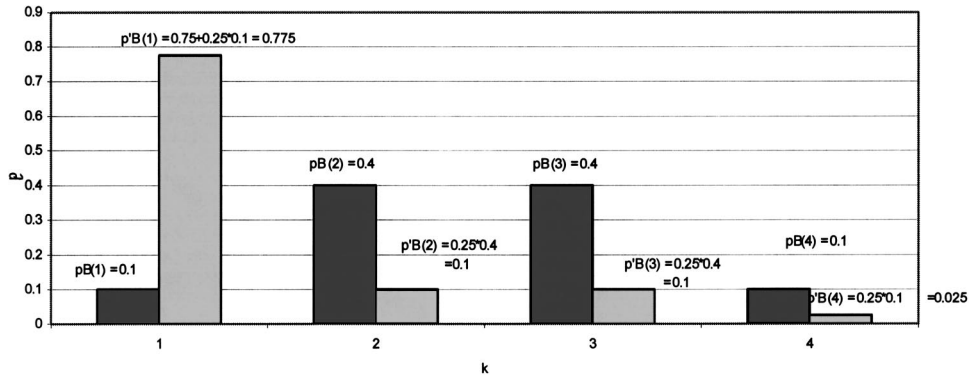

Fig. 6. Probability distribution corrected for risk dependence

4. Generate the risk profile of the project.
5. Using the risk profile, determine whether NPV at the required level of certainty is sufficient to justify the investment.
6. If the risk is unacceptable, allocate resources to selected risk significant activities to reduce their variability, until the risk level is sufficiently reduced. If a satisfactory result cannot be achieved, abandon the project.

Validation Experiment Case Study

The project chosen for validation of the multi-factor method consisted of three office buildings, 10–12 stories high, and the associated site development tasks. It comprised 28 activities: 19 construction ''packages,'' each composed of various works (such as all structure or finish works in a building), with its cost, duration, variation range, and dependence attributes; six ''financial activi-

Table 3. Project Activities and Their Attributes

ID No.	Activity	Precedent activity	Minimum duration (weeks)	Maximum duration (weeks)	Minimum income/(cost) ($ 000)	Maximum income/(cost) ($ 000)
1	Procurement of land	—	0	0	0	0
2	Preliminary design and permit	1	39	65	0	0
3	Earthwork and setup	2	4.33	8.67	0	0
4	Detailed design-buildings	2	13	21.67	(1,035)	(765)
5	Detailed design civils and landscaping	2	8.67	13	0	0
6	Foundations and substructure	4	8.67	13	(2,875)	(2,125)
7	Structure A	6,5	17.33	26	(2,415)	(1,785)
8	Finish A	7	13	21.67	(2,645)	(1,955)
9	Structure B	7,6	17.33	26	(2,875)	(2,125)
10	Finish B	8,9	13	21.67	(3,105)	(2,295)
11	Structure C	6	21.67	30.33	(3,795)	(2,805)
12	Finish C	11	21.67	30.33	(4,600)	(3,400)
13	Shelter	9,5	13	17.33	(1,207)	(892)
14	Retaining walls	11,5	8.67	13	(1,380)	(1,020)
15	Site works	14	4.33	8.67	(805)	(595)
16	Installation of special equipment in A	8	4.33	8.67	(2,012)	(1,487)
17	Furniture and fixtures	16,10,12	4.33	8.67	(2,587)	(1,912)
18	Site civils	6,14,15,13	8.67	13	(2,530)	(1,870)
19	Landscaping and gardening	18,13	8.67	13	(1,437)	(1,062)
20	Running in	17,19	4.33	8.67	0	0
21	Payment 1 (Deposit A)	6,27	0	0	0	0
22	Payment 2 (Deposit B)	9,28	0	0	5,525	7,475
23	Payment 3 (Deposit C)	11,29	0	0	6,375	8,625
24	Payment 4 (Occupation A)	27,18	0	0	5,950	8,050
25	Payment 5 (Occupation B)	28,19	0	0	10,200	13,800
26	Payment 6 (Occupation C)	29,20	0	0	11,050	14,950
27	Marketing and sales-A	2	0	52	0	0
28	Marketing and sales-B	2	0	130	0	0
29	Marketing and sales-C	2	0	104	0	0

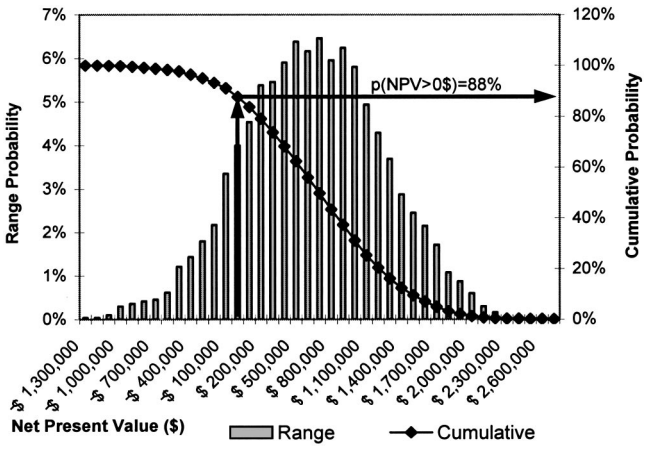

Fig. 7. Typical program output; 5,000 iterations, normal distributions, all risk factors vary

ties'' with payment and dependence attributes; and three marketing activities, as shown in Table 3. In absence of other information, a variation range of 30% of the mean value has been assumed, with the most likely value equaling the mean. Several asymmetric cases and several risk dependences between the risk significant parameters were also assumed, as explained later.

The ''expected'' (i.e. without consideration of risk) net present value of the project, calculated from the expected values of all the factors amounted to $650,000, given the cost of capital to the owner of 12% per year. The risk profile of the project was calculated and expressed as a cumulative distribution curve, which indicates the probability of any value of NPV being exceeded (Fig. 7 shows a typical program output). The owner's objective was to attain a positive NPV of the project's cash flow with a confidence level of 90%. It will be seen that in order to attain this confidence level some changes were necessary in the basic factors.

The following issues had to be resolved within the context of the risk analysis.

Choice of Risk Significant Parameters

Sensitivity analysis showed that the ten most significant parameters were, in the order of their impact on NPV (C=cost, D=duration, activity IDs as designated in Table 3): C-26, C-25, C-23, C-22, D-28, D-29, C-12, C-11, D-2, and D-4. Use of all N from $N=5$ to $N=10$ parameters was examined. It is evident from Table 4 that the increase of the number over $N=7$ has almost no effect on the project risk profile.

Table 4. Influence of Number of Varying Risk Factors on Project Risk (at $X_o=0.5$, $R=30\%$)

Enumeration	p (NPV>0) $K=5$	p (NPV>0) $K=3$
$N=5$	83.4%	—
$N=6$	82.5%	—
$N=7$	81.9%	81.0%
$N=8$	—	80.4%
$N=9$	—	80.2%
$N=10$	—	80.0%
Monte Carlo	2,500 iterations	5,000 iterations
p (NPV>0)	78.1%	79.1%

Table 5. Influence of Range of Estimate on Project Risk (at $N=7$, $K=5$, $X_o=0.5$)

Measure	$R=20\%$	$R=30\%$	$R=40\%$
p (NPV>0)	88.0%	80%	73.5%
NPV (p>90%)	−$50,000	−$350,000	−$650,000

Range of Variation

It was assumed that the project's parameters vary within the range of ±15% from their most likely values ($R=30\%$). The variation of the parameters within $R=20$ and 40% ranges has also been examined. The impact of the variation range on the confidence level of the project profitability is shown in Table 5. As expected, it shows that the increase in range directly and adversely affects the risk of the project—a roughly 7–8% decline in the confidence level with each increase of 10% of the range.

Probability Distribution over Variation Range

It was assumed as the leading alternative that the parameters follow the normal pattern with the mean coinciding with the most likely value at the mid range of variation with a standard deviation of 1/6 of the range. The impact on risk of the triangular distribution (with a standard variation of 0.21 of the range) was also examined. Table 6 shows that the triangular distribution, with a wider dispersion over the variation range, results in the symmetric case in higher risk—a decrease of the confidence level by 6%. In the asymmetric case, however, (''Influence of Skewness'') it has a mitigating effect on the very strong impact of skewness.

Number of Sections of Parameter Variation Interval

The options of $K=3$ and have been examined. As Table 7 shows, the increase in K resulted in a very small change in the risk profile. In this particular case $K=3$ produces reliable results of the risk analysis, differing by only 0–1.5% from the results obtained when setting $K=5$.

Influence of Skewness

The location of the most likely value away from the middle of the variation range was also examined. The most likely values of the income payments was tested at $X_o=0.37\ \Delta P$ and at $X_o=0.20\ \Delta P$ of the range, while the values of the cost payments were located at $X_o=0.57\ \Delta P$ and $X_o=0.80\ \Delta P$, respectively. As expected (Table 7), the skewness of the distribution has a strongly adverse effect on the project risk level if the most likely value is closer to the ''unfavorable'' end of the distribution and vice versa. In this case the wider dispersion of the triangular distribution considerably mitigates the results.

Interdependence between Activities

Three cases of dependence were examined. In each, different factors were related to the most significant parameter, C-26, which

Table 6. Influence of Probability Distribution Type on Project Risk (at $N=7$, $K=5$, $R=30\%$)

Distribution	$X_o=0.5$	$X=0.57$	$X=0.80$
p (NPV>0) for normal/beta	80.0%	48.5%	1.5%
p (NPV>0) for triangular	74.0%	57.0%	17.0%

Table 7. Influence of Number of Interval Sections K and of Skewness on Probability of NPV>0

Range (%)	N	X_o	Distribution	$p(K=3)$	$p(K=5)$	Δp
20	7	0.5	Normal	87.0	88.0	−1.0
30	7	0.5	Normal	78.5	80.0	−1.5
30	7	0.5	Triangular	73.0	74	−1.0
30	7	0.57	Beta	48.5	48.5	—
30	7	0.57	Triangular	58.5	57.0	1.5
30	7	0.80	Beta	1.0	1.5	−0.5
30	7	0.80	Triangular	15	17	−2.0
40	7	0.5	Normal	73.5	73.5	—

was assumed to be independent. In the first case, only the second most significant parameter (C-25) was related; in the second case, the second through fifth most significant parameters (C-25,C-23,C-22,D-28) were related; and in the third case, six parameters (C-25,C-23,C-22,D-28,D-29,C-12) were used. The results are presented in Table 8.

Two conclusions can be drawn from the results. In general, the positive dependence between the significant parameters adversely affects the risk of the project. The coordinated variation of various parameters amplifies its effect. The risk increases both due to stronger interdependence of the main parameters as expressed by the growth of the coefficient ρ, and due to the growing number of parameters that vary interdependently.

Risk Management

As evident from the results, the project does not attain its desired level of risk performance if the variability of the significant parameters is perceived to vary within ±15%. If however the owner were to invest funds to obtain additional data, in better design, or in marketing, the risk exposure could be reduced. Checking a hypothetical increased investment of $30,000 in design and $50,000 in marketing, to increase value and shorten sales period to within ±10% of the mean, showed that the owner's risk goal (NPV>0 with 90% certainty) can be achieved, while the "deterministic" NPV of the project is reduced by less than $80,000.

Discussion

The multifactor analysis described can be applied in relatively few steps, which depend on the extent and reliability of information available about the significant risk factors. The most important aspects and benefits of its use are as follows.

Definition of Risk Objective and Risk Tolerance

The risk objective is defined as the desired probability level of obtaining a specified NPV of the project cash flow. The desired NPV may be $0, which indicates that the project is feasible at the applied discount rate, or it may be set at a higher value. The user can arbitrarily set the acceptable level of risk, usually within the range of 85–95%. For example, in the United Kingdom a 90% probability level of profitability is required for public sector projects [Central Unit on Procurement (CUP) 1993]. The same level of confidence has also been employed in Ye and Tiong (2000), and it is comparable with the 90–95% confidence levels required from physical performance of various building elements (Harr 1987).

The desired probability level must be set in light of the computational tolerance associated with the empirical assessment method. The tolerance may be set arbitrarily at 2–3%, which is in fact very small when compared with the accuracy of elicited minimum, maximum (and expected) values for the individual activities.

Selecting "Significant Risk Factors"

While it is recognized that it is not necessarily the case for every possible project, the risk impact for a majority of real life projects is concentrated in just a few cost or duration parameters. The significant factors are chosen according to their impact on the project NPV, using conventional sensitivity analysis. It is recommended that the minimum number of significant parameters be used without compromising accuracy: from seven to ten parameters may suffice in projects composed of 25–30 activity packages. The final number of significant parameters to be examined should be corroborated by their marginal impact on the project risk, as shown in the Table 4 of the case study. A smaller number of parameters allows a more thorough analysis of the feasible alternatives when it is necessary to reduce the project risk.

Elicitation of Risk Characteristics of Selected Parameters

The elicitation of information for the assessment of a parameter's variation pattern may be considered the most difficult task of the risk evaluator. The information available for this purpose can be classified, as explained earlier, into four levels. The amount of information varies between complete performance records from which one can derive the variation pattern of any parameter, at one extreme, to a lack of any information about the variation range, at the other. The proposed procedure can handle the available information at any one of these four operational levels.

Skewed Probability Distribution over Variation Range

If there is reason to believe that the distribution is asymmetric, i.e., that the most likely estimate is not at the middle of the variation range, the Beta distribution with standard deviation to range ratio as close as possible to the Normal distribution is recom-

Table 8. Impact of Interdependence between Risk Factor

		Confidence level for NPV>0, with $X_o=0.5$, $R=30\%$, $K=5$, and ρ, where			
Independent factor	Dependent factor	ρ=0 (%)	ρ=0.25 (%)	ρ=0.50 (%)	ρ=0.75 (%)
C-26	C-25	80	78.5	77	73.5
C-26	C-25,C-23,C-22,D-28	80	72.5	73	73
C-26	C-25,C-23,C-22,D-28,D-29,C-12	80	76.5	74	72.5

mended. Alpha and beta varying between 1 and 4 for different values of the most likely estimates (Table 2) perform well in this respect. The triangular distribution provides an additional option for this case. Placement of the most likely estimate closer to an extreme affects the mean of the distribution in a similar manner, i.e., it may effectively increase all the costs in the project, all durations etc., and thus significantly affect both the profitability and the reliability of the project. For this reason the specific estimate of the most likely outcome should be handled with care. The determination of the most likely value in the asymmetric case is possibly the most difficult task for the project owner or manager. When in doubt it is better to use a symmetric distribution or the closest skewness option (at 40 or 60% of the range).

Handling of Dependencies

The effect of risk dependencies on the results has been shown to be significant. Since significant dependencies exist in most engineering projects, the ability to model dependencies between any risk factors in a straightforward manner, is an important feature of the proposed model. Dependence can be positive or negative.

Generation of Risk Profile of Project

The risk profile of the project can be assessed either with numerical enumeration or with the Monte Carlo sampling method. The latter is more efficient in term of computation time and space if a large number of variables are involved. The former gives a more accurate insight into the effect on risk of the change in the attributes of the various parameters. The number of interval sections in each variable parameter should preferably be $K=5$ for up to 7–8 parameters examined simultaneously, and $K=3$ can be used when 9–11 parameters are to be evaluated. Selective enumeration methods can handle many more parameters simultaneously. This number can be considerably increased if more efficient computation methods are employed, although the Monte Carlo sampling method presents an efficient and effective alternative.

Risk Management

Risks can be managed and modified if necessary, using the procedure proposed here, by introduction of technological or organizational changes into risk significant project activities. In the final analysis, risk management involves investment in technology, organization and information sources in order to reduce the variability of these factors.

Conclusions

A practical and transparent method for risk analysis of investment in building construction projects has been developed, implemented in software, and tested. It enables a building project owner/developer to:

1. Assess the economic risk associated with their project very quickly and with relatively little input information and
2. Identify the ways in which he/she can intervene to ameliorate the risks, and assess the effectiveness of such interventions.

The multifactor method can operate with risk factors at any available level of information detail. This is an important advantage, because the difficulties associated with problem formulation and input data elicitation are the major obstacles to effective risk analysis (Ranisinghe and Russell 1993; Ye and Tiong 2000). "Forced" elicitation of increasingly more detailed risk distributions from owners/developers generally requires a level of effort that may be impractical for most executives. It is also likely to be counterproductive from a managerial perspective, by distracting attention from the risks that should be dealt with, and focusing instead on computational accuracy (which in any case is already far greater than the accuracy of estimation of input data). The method could be both enhanced and simplified—from the point of view of the nonprofessional user—if default probability distributions for typical activities were made available. Such distributions should be based on historic duration and cost distribution data for completed projects. The feasibility of providing such distributions for commercial housing projects is currently being investigated.

In summary, the multifactor method should be compared with the standard sensitivity analysis method, which is customarily applied today in most building investment projects. The multi-factor method produces more realistic results, because of the following advantages:

- All significant risk factors are handled simultaneously, rather than in isolation;
- It assumes a normal (and can use any other) probability distribution for each risk factor value over its range of variation, rather than simple linear variation;
- It accounts for dependencies between the risk factors in a straightforward manner; and
- It allows for easy evaluation of the tradeoffs between the cost and the risk inherent in a project.

Appendix. Detailed Enumeration Procedure

In the enumeration procedure, the range of variation of any risk factor P (cost or duration of an activity) is divided into $k=1,\ldots,K$ discrete and equal intervals of size d, such that $\Delta P=K*d$. Each interval k has a mean value P_k and a probability $p(k)$ of the factor value falling within this interval. If the probability distribution of a parameter for a given activity is known, and represented either as a function or as discrete values, then the associated values P_k and $p(k)$ can be determined for any value of k.

A particular event Z can be defined by assigning a specific value P_i^z to each parameter in its range ΔP_i, which is the same as defining an appropriate value for k for each of the risk factors (for simplicity k is not indexed). These combinations can be enumerated using a decision tree, with the values of k for the first risk factor at the first level below the root (see Fig. 8), the values of k for the second risk factor on the second level, and so on. Any event Z may be depicted by a complete branch in the decision tree in Fig. 8 with a distinct selection for the value k_i of each risk factor i. The number N of all possible combinations of events Z for a project consisting of n variable parameters P will be $N=K^n$.

For each event Z, once values of k for each risk factor P_i have been assigned, the activity network can be solved, and both the net present value NPV^z, and the probability of its occurrence $p(P_{ik}^z)$ can be calculated using Eqs. (1) or (2). Assuming initially that all the parameters are independent of one another, the probability that an event z with NPV^z will occur, can be derived from

$$p(\mathrm{NPV})=p(P_1^z)*p(p_2^z)*\ldots*p(P_n^z)$$
$$=p(k_1^z)*p(k_2^z)*\ldots*p(k_n^z) \quad (8)$$

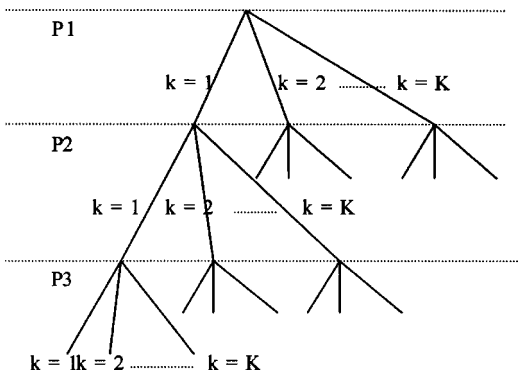

Fig. 8. Enumeration decision tree

where k_i^z = probability interval of a particular parameter P_i in the event z; and $p(k_i^z)$ is derived from its characteristic distribution, as shown in Fig. 3.

The different values of NPV^z, each resulting from a different event (a combination of values for risk factors $|P|^z$), can be grouped into K' distinct intervals; the probability of the project NPV falling in any interval k' is simply the sum of the probabilities of the individual values of NPV falling in that interval. In this way, the probability distribution of NPV for the whole project can be calculated. Given this distribution, and its cumulative expression, both the risk of the project not attaining any desired NPV, and the NPV associated with any specific risk level, can be assessed.

References

Ang, A. H.-S., Abdelnour, J., and Chaker, A. A. (1975). ''Analysis of activity networks under uncertainty.'' *J. Eng. Mech. Div., Am. Soc. Civ. Eng.*, 101(4), 373–387.

Ashley, D. B., and Bonner, J. J. (1987). ''Political risks in international construction.'' *J. Constr. Eng. Manage.*, 113(3), 447–466.

Benjo, M. (1999). ''Risk estimation in construction projects.'' MS thesis, Technion, Israel Institute of Technology, Haifa, Israel.

Central Unit on Procurement (CUP). (1993). ''Managing risk and contingency for works projects.'' *Guidance Manual No. 41*, London.

Cooper, D., and Chapman, C. (1987). *Risk analysis for large projects: models, methods and cases*, Wiley, New York.

Diekmann, J. E. (1983). ''Probabilistic estimating: Mathematics and applications.'' *J. Constr. Eng. Manage.*, 109(3), 297–308.

Edwards, P. J., and Bowen, P. A. (2000). ''Risk and risk management in construction projects: Concepts, terms and risk categories re-defined.'' *J. Constr. Procure.*, 5(1), 42–57.

Harr, M. E. (1987). *Reliability based design in civil engineering*, Dover, New York.

Johnson, N. L., Nixon, E., Amos, D. E., and Pearson, E. S. (1963). ''Table of percentage points of Pearson curves, for given $\sqrt{\beta_1}$ and β_2 expressed in standard measure.'' *Biometrika*, 50(3,4), 459–498.

Keefer, D. L., and Bodily, S. E. (1983). ''Three-point approximations for continuous random variables.'' *Manage. Sci.*, 29(5), 595–609.

Kottas, J. F., and Lau, H.-S. (1982). ''A four-moments alternative to simulation for a class of stochastic management models.'' *Manage. Sci.*, 28(7), 749–758.

MacCrimmon, K. R., and Ryavec, C. A. (1964). ''Analytic study of the PERT assumptions.'' *Oper. Res.*, 12, 16–37.

Malcolm, D. G., Roseboom, J. H., Clark, C. E., and Fazar, W. (1959). ''Application of a technique for research and development program evaluation.'' *Oper. Res.*, 7, 646–669.

Morris, P. W. G., and Hough, G. H. (1987). *The anatomy of major projects*, Wiley, New York.

Newton, S. (1992). ''Methods of analyzing risk exposure in the cost estimates of high quality offices.'' *Construction management and economics*, Vol. 10, E&FN Spon, London, 431–449.

Ranisinghe, M. (1994). ''Contingency allocation and management for building projects.'' *Construction management and economics*, Vol. 12, E&FN Spon, London, 233–243.

Ranisinghe, M., and Russell, A. D. (1993). ''Elicitation of subjective probabilities for economic risk analysis: An investigation.'' *Construction management and economics*, Vol. 11, E&FN Spon, London, 326–340.

Tah, J. H. M., and Carr, V. (2000). ''A proposal for construction project risk assessment using fuzzy logic.'' *Construction management and economics*, Vol. 18, Taylor and Francis Ltd., London, 491–500.

Touran, A. (1993). ''Probabilistic cost estimating with subjective correlations.'' *J. Constr. Eng. Manage.*, 119(1), 58–71.

Wang, C.-H., and Huang, Y.-C. (2000). ''A new approach to calculating project cost variance.'' *Int. J. Proj. Manage.*, 18, 131–138.

Ye, S., and Tiong, R. L. K. (2000). ''NPV-at-risk method in infrastructure project investment evaluation.'' *J. Constr. Eng. Manage.*, 126(3), 227–233.

Yeo, K. T. (1990). ''Risks, classification of estimates, and contingency management.'' *J. Manage. Eng.*, 6(4), 458–470.