

On the assessment of economic risk: factorial design versus Monte Carlo methods

Willem J. H. Van Groenendaal & Jack P. C. Kleijnen

Department of Information Systems and Auditing/Center for Economic Research (CentER), School of Management and Economics, Tilbury University, 5000 LE Tilburg, The Netherlands

The feasibility of large investment projects (such as gas transmission and power system projects) has many aspects. Usually, this problem cannot be modeled as a single optimization problem; instead, the multiple aspects (demand, supply, prices, investment costs) are modeled separately. Each aspect may require a large, nonlinear submodel. The results of such a submodel can often be summarized by one or a few variables, which combine all the submodel's information; for example, total demand is the sum of the demand per customer type, each type being modeled separately. Traditionally, the feasibility of the investment project is then judged by combining the results of the various submodels for the 'base case' values of all model inputs.

This base case information, however, is not sufficient for the decision makers; they also like to know the economic risk they are taking. To assess this risk on the project level (Hertz, D. B., Risk analysis in capital investment. *Harvard Business Review*, 1964, 95–106) developed a method known as *risk analysis*. This method is based on the estimated probability distribution of a project's net present value (*NPV*). This distribution is obtained by introducing distributions for the model inputs. The project's economic risk is then expressed as the probability of a negative *NPV* exceeding a critical value (say) α . Nowadays this approach is becoming popular, because many software packages (such as @RISK and Crystal Ball) facilitate such a risk analysis. Although Hertz's risk analysis is appealing, it has a number of theoretical and practical flaws, which may lead to wrong conclusions. These flaws are discussed in this paper.

From a modelling point of view, Hertz's risk analysis is similar to analysing the technological or operational risk of an investment. However, economic risk and technological risk are different concepts that require different analyses. In this paper these differences are discussed and it is shown that Hertz's risk analysis does not measure what is normally meant by a project's economic risk. Furthermore, the information requirements for the application of risk analysis to large investment projects are formidable; this makes the results of Hertz's investment analysis unreliable. Less information is required by sensitivity analysis based on the statistical design of experiments (such as 2^{k-p} designs); this analysis is more robust, and leads to results that better satisfy the information needs of decision makers. © 1997 Elsevier Science Limited.

1 INTRODUCTION

Large public investment projects are analyzed and evaluated through large complex models. Both in theory and in practice the criterion for a project's appraisal is often its net present value (*NPV*), defined as $NPV = \sum_{j=t}^T (1+r)^{-j} NR_j$, where t is the start and T is the end of the evaluation period, r is the discount rate, and NB_j stands for the net benefits in period j . Often *NPV* is estimated for the most likely or base case scenario for model inputs, which gives (say)

NPV^0 . NPV^0 , however, is not considered sufficient information by decision makers: they also require information to help them assess the uncertainty of the result, needed to support their assessment of the project's economic risk.

The role of uncertainty in *NPV* calculations has several aspects, from theoretical and practical points of view. This paper will show that risk analysis as introduced by Hertz [1], and based on work by Hillier [2], has a number of drawbacks. The method proposed by Hertz to assess economic risk is similar to those

used to evaluate a project's technological risk. However, technological risk and economic risk differ in nature. Technological risk depends only on the investment itself, in case international standards are applied in construction, operation, and maintenance. Technological risk is independent of other projects (except for catastrophes, such as being hit by a crashing airplane); this is especially true for well developed technologies. As we shall show, economic risk is defined in terms of other investment projects and national income.

Like technological risk analysis, Hertz's risk analysis uses Monte Carlo simulation; that is, it uses (pseudo-)random numbers. One of the reasons for introducing Monte Carlo simulation to evaluate the economic risk of an investment project is the presence of many factors (input variables, parameters, etc.). Another important argument is that Monte Carlo simulation takes into account interactions among factors; that is, it measures what happens if several factors change simultaneously. In case there are (say) n factors each with two possible values, practitioners argue that at least 2^n simulation runs are required to analyze all main effects and interactions [3]. Monte Carlo simulation reduces the number of runs actually executed [4]. But in order to cope with 'many' factors and interactions there are other approaches, especially the mathematical statistical theory on design of experiments (DOE).

In practice, changing one factor at a time is very popular in sensitivity analysis. However, this method is known to be inefficient and ineffective. There are designs that give accurate estimators of all main effects and certain interactions, but require fewer than 2^n simulation runs; see Box *et al.* [5]. We shall show that the application of DOE can help to meet the information needs of the decision makers.

Before we discuss methods for evaluating the riskiness of a project, we must answer the question: what is meant by the *risk* of an investment? An investment influences a country's economy in many ways and at several levels; see Sheng and Cho [6]. Our assessment of risk is restricted to the project level, that is, the riskiness of the *NPV* of a single project. But even on that level, several forms of risk can be distinguished: economic risk, technological or operational risk, regulatory risk, etc. We are interested in economic risk. We shall discuss the difference between technological and economic risk, and formulate a method to estimate the latter. We shall not discuss the effects of regulatory risk; that is, we assume that the government regulations will remain constant during the evaluation period.

Note that regulatory risk is a form of risk which is becoming more and more important, and which can affect both the technical and the economic risk. Regulatory risk can be described as the risks emerging

from government action. Changes in regulation can affect both the technical risk and the economic risk, but in general only when the system is already in place. An example may illustrate this. If the government introduces stricter environmental regulations, the design of a system is affected and has to be adjusted. This calls for a redesign, which affects the technical risk. At the same time the increase in costs affects the cost structure and thus the profitability. Furthermore, investors might change their expectations of the company's environmental liability, and thus become more reluctant to invest in this company; this affects the company's position on the capital market; see also Kolbe *et al.* [7].

In compliance with standard economic theory we use the term *risk* for those situations in which the degree of ignorance about the data is expressed through probability density functions for the inputs ([8], pp. 427–9). In case no information about the probability distributions is available, we speak about decision making under *uncertainty*.

There are many methods for analysing the effects that changes in factors have on the *NPV* (or any other output variable of interest). Well known and widely applied techniques are break-even analysis, sensitivity analysis, scenario analysis, risk analysis, decision trees, and uncertainty analysis [9,10]. We discuss only sensitivity analysis and Hertz's Monte-Carlo risk analysis.

Sensitivity or *what-if* analysis is the assessment of the consequences of changes in inputs and model parameters, not taking into account information on the *probability* of these changes. The most popular form of sensitivity analysis is the one-factor-at-a-time approach; economists speak of 'ceteris paribus'. The advantage of this approach is that the results can be interpreted easily, so the decision makers have no problems understanding the results. There are also several graphical tools for visualizing the results, such as tornado graphs and spider plots [11]. Another form of sensitivity analysis is *scenario analysis*. Scenario analysis recalculates the model for a combination of simultaneous changes in input variables. Every scenario is considered a realistic future. In many cases an optimistic scenario and a pessimistic scenario are distinguished, besides the base case scenario. A main problem is that with many inputs the number of possible scenarios increases rapidly, and the scenarios actually used are selected somewhat arbitrary.

Risk analysis tries to assess the same effects as sensitivity analysis does, but it takes into account the (joint) probability distribution of the input variables. The main practical problem is: how to obtain the distributions of inputs, that is, the joint distribution, conditional, and marginal distributions per input type?

Currently there is a renewed interest in Hertz's risk analysis, because many software packages, such as

@RISK and Crystal Ball, make this risk analysis easily available for (inexperienced) users. These software developments seem to support Hertz and Thomas's [12] claim that at least twenty-five years of promotion are needed between the moment a concept is invented and the moment it is widely accepted. This is also confirmed by the survey study of Ho and Pikes [13]. These authors see as remaining problems for the widespread application of Hertz's risk analysis: (i) the assessment of the (subjective) probability distributions of the inputs, and (ii) making managers understand the approach (now they lack the required knowledge). We claim that it is questionable to qualify the managers' distrust of the concept as a lack of knowledge. The theory behind risk analysis is criticized too, as this paper will show. Maybe the managers' distrust is not only based on lack of knowledge, but also on sound gut feeling.

In economics there are many models that formalize risk and risk attitude. Hertz and Thomas ([12], p. 21) suggest that there is only a minor difference between their risk analysis and formal risk analysis based on utility or prospect theory. They state that 'it [risk analysis] differs from decision analysis in its use of risk simulation as a solution technique and by its intuitive, rather than formal (i.e. not using the utility function apparatus) incorporation of the decision-maker's risk attitude (i.e. his preference function for payoffs) into the decision making process'. In our opinion, however, their claim is incorrect. We shall show that they ignore the crucial fact that economic risk can be evaluated only in a broader context, in relation to other projects and total national income. Therefore, it is misleading to use the term risk analysis for their approach (yet their approach has become popular under this name).

This paper is organized as follows. Section 2 discusses the differences between technological risk and economic risk. Section 3 discusses Hertz's risk analysis as a tool for project evaluation, and compares it with sensitivity analysis. Section 4 improves sensitivity analysis through experimental design. Section 5 discusses a case study: investing in a gas transmission pipeline on the Indonesian island Java. Section 6 contains conclusions.

2 TECHNOLOGICAL RISK VS ECONOMIC RISK

With technological risk we do not mean the effects of possible technical distortions that may occur when *new* technologies are introduced; see Kunreuther *et al.* [14]. This paper is restricted to the many investment projects that do not require new technologies (this paper was inspired by work for the Dutch natural gas company Gasunie Engineering, performed in Indone-

sia). For these projects it is reasonable to assume that the technical risks are covered by the use of internationally accepted design standards.

As an example consider a gas transmission system. In its simplest form the technological safety requirement can be formulated as: the chance of a blow up during operation has to be smaller than (say) α : $P(\text{PIPELINE BLOW UP}) \leq \alpha$ with $0 < \alpha \ll 1$. This requirement leads to technical requirements for material and design quality (such as pipeline thickness for different areas, cathodic protection, and safety valves) for various parts of the system. For many technologies the technical requirements have been translated into international standards for the construction, maintenance, and operation, which are sometimes enhanced with additional local standards. If these international safety standards are met, the technological risk can be insured in the same way as a house can be insured against fire.

Furthermore, the technological risk of the total system consists of a large number of smaller risks, which are related to particular parts of the system, but are mutually independent. For example, the risks involved in the construction of a citygate station are independent of the risks of the pipeline sections out in the field. If one of the risks materializes, the system can be repaired relatively easy.

Technological risk can be quantified as follows. Let a gas transmission system consist of n separate elements (the wellhead, pipeline sections, city-gate stations, etc.). For every element we can define an indicator function (say) x_i ($i = 1, \dots, n$) which indicates whether the segment operates ($x_i = 1$) or malfunctions ($x_i = 0$). Let the chance that the element x_i functions correctly be $P(x_i = 1) \geq 1 - \alpha_i$. The technological risk for any part x_i of the system depends on a large number of factors, which will not be independent. Assessing the technological risk for the separate elements is the subject of reliability analysis. The chance that the segment corresponding with x_i fails if international standards are applied, will be within the required limit: $P(x_i = 0) \leq \alpha_i$. Application of the international standards will also assure that the risks of two elements x_i and $x_{i'}$ ($i \neq i'$) can be considered to be independent. The risks of two different investment projects (say, a gas transmission pipeline and an oil pipeline) will also be independent. So a technically safe system remains safe, *independent of other systems*, provided it is well looked after. Good housekeeping, however, will not reduce the economic risk to a chosen level. As we shall show in Section 2.1, contrary to technological risk, economic risk is actually defined in terms of the covariances between projects.

The example of the two pipeline systems may clarify this point. The profitability of the gas and the oil transmission systems will depend on the development

of national and international energy prices. Even in a strongly regulated domestic market for natural gas (which for most countries or regions can be characterized as a technical and economic monopoly due to shape of the infrastructure) price changes for oil products affect the price for gas. The main reason is that every consumer (but especially large consumers) has access to technologies that use oil products as an alternative for natural gas. These technologies become profitable, whenever the price of gas differs much from the price of the alternative. Actually many large users have a dual-fuel (oil and gas) system installed permanently, to avoid gas peak-load prices. So the economics of the gas transmission system are constantly affected by the market for alternative fuels, and thus the alternatives are a permanent risk. The technological risk of the gas transmission system, however, is not affected by these alternatives.

2.1 How to measure a project's economic risk?

In this subsection we shall see what Hertz's risk criterion $P(NPV \leq 0) \leq \alpha$ actually measures, and we shall investigate if there are better ways to express this measurement. First we must define the exact meaning of economic risk, and we must determine if $P(NPV \leq 0)$ is an indication for this risk. Risk is an important issue in economic theory and finance, and it has resulted in subjective expected utility theory and prospect theory [15]. We do not discuss these theories, but limit our analysis to the role of risk in a project's NPV. Arrow and Lind [16] gave a formal proof of the sufficiency of $E(NPV)$ when evaluating a project from a group or national point of view. They used the framework of a small theoretical and analytical economic model, and took the increase in the variance of income as an indication for risk. In practice, however, such a model is not available for investment projects; descriptive simulation models are used instead. Arrow and Lind's argument, however, still holds under certain conditions. Suppose the project's NPV is shared by a large group of (say) M individuals with individual m 's income denoted by (say) Y_m . Then the effect on the variance of individual m 's income (that is, the increase in individual m 's risk) due to the project is

$$\begin{aligned} V(Y_m + M^{-1}NPV) - V(Y_m) \\ = M^{-2}V(NPV) + 2M^{-1}COV(Y_m, NPV) \end{aligned} \quad (1)$$

where $V(Y_m)$ denotes the variance of Y_m and $COV(Y_m, NPV)$ denotes the covariance between Y_m and NPV . Equation (1) shows that for large values of M , the effect of the project on an individual's risk may be neglected, since it will tend to zero when M is large. So if the risk is spread among many individuals,

then there is no significant risk to individuals. This is known as risk spreading.

A small risk per individual does not mean that the risk for a group or society as a whole may be neglected. The group's risk is obtained by summing the risk over the M individuals:

$$\begin{aligned} \sum_{m=1}^M [M^{-2}V(NPV) + 2M^{-1}COV(Y_m, NPV)] \\ = M^{-1}V(NPV) + 2M^{-1} \sum_{m=1}^M COV(Y_m, NPV) \\ = M^{-1}V(NPV) + 2M^{-1}COV(Y, NPV), \end{aligned} \quad (2)$$

where $Y = \sum_{m=1}^M Y_m$. To make the group's risk go to zero for large M , it is necessary to assume that total income Y and NPV are independently distributed, so $COV(Y, NPV)$ becomes zero; otherwise this term will in general not vanish (since Y increases with M). Note that the absence of group risk does not necessarily mean that there is no risk to individuals: $COV(Y_m, NPV) > 0$ may compensate $COV(Y_{m'}, NPV) < 0$, with $m \neq m'$, so that the overall covariance may be neglected. So the Arrow-Lind theorem (stating that $E(NPV)$ provides sufficient information for evaluating a project) has two conditions: (i) the returns on a public project are distributed independent of group income ($COV(Y, NPV) = 0$), and (ii) the benefits and costs of the project are spread over a sufficiently large group ($M \uparrow \infty$).

However, when there are *several* projects, risk spreading is not sufficient to make $E(NPV)$ give sufficient information for evaluating a project. If all projects are perfectly positively correlated (that is $(COV(NPV_i, NPV_j) / \sqrt{V(NPV_i)V(NPV_j)}) = 1$), all projects together might be rejected, whereas each individual project is accepted when evaluated independently. This can be explained as follows. The variance of a project may be ignored, but the covariances between the project and other projects and income must be considered. If there are N projects, the expected average revenue is

$$E\left(N^{-1} \sum_{n=1}^N NPV_n\right) = N^{-1} \sum_{n=1}^N E(NPV_n), \quad (3)$$

and the variance of the average revenue is

$$\begin{aligned} V\left(N^{-1} \sum_{n=1}^N NPV_n\right) &= N^{-2} \sum_{n=1}^N V(NPV_n) \\ &+ 2'N^{-2} \sum_{n=1}^N \sum_{n'=1}^{n-1} COV(NPV_n, NPV_{n'}) \\ &= N^{-2} \sum_{n=1}^N \sigma_n^2 + 2N^{-2} \sum_{n=1}^N \sum_{n'=1}^{n-1} \sigma_{nn'}, \end{aligned} \quad (4)$$

where $\sigma_{nn'}$ denotes the covariance between the projects n and n' . With $\sigma_n^2 < \infty$ for all $n \in N$ the first term goes to zero, as the number of projects (N) increases. There are $N(N-1)$ covariance terms, and

since all expectations and variances are finite, the covariances are finite too (Cauchy-Swartz theorem). Denote the average covariance between projects by $\bar{\sigma}_{nn'} = [N(N-1)]^{-1} \sum_{n=1}^N \sum_{n'=1}^{n-1} \sigma_{nn'}$. Then the second term can be rewritten as

$$\begin{aligned} 2N^{-2} \sum_{n=1}^N \sum_{f=1}^{n-1} \sigma_{nn'} &= 2N^{-1}(N-1)\bar{\sigma}_{nn'} \\ &= 2\bar{\sigma}_{nn'} - 2N^{-1}\bar{\sigma}_{nn'} \approx 2\bar{\sigma}_{nn'}. \end{aligned} \quad (5)$$

So for decision makers confronted with a portfolio of investment projects, only the covariances between profitable projects ($E(NPV) > 0$) are of real importance. (This argument is similar to the one in the theory on portfolio diversification and individual asset risk (Copeland and Weston [17], pp. 184–8).) As Wilson ([18], pp. 205–6) pointed out, projects with negative covariances might even have a premium, because they make the portfolio of projects less risky. The chance of a negative NPV plays no role in eqn (5); it is no indication of a project's risk.

Nevertheless, there is a disclaimer. The model in eqns (3)–(5) has been further developed for easy-to-trade securities. But the investment projects we discuss are bulky all-or-nothing type of investments, which can not be traded in the way securities are. The model is difficult to apply for securities; applying it for a constant stream of bulky investment projects is impossible. The data requirements would be formidable. Furthermore, once construction has started, it will be difficult to stop; so contrary to a portfolio of securities, reshuffling the portfolio of projects is difficult.

We showed that a project's risk depends on the way the project result is influenced by the other projects (see eqn (5)) and income (see eqn (2)). We also stated that from a practical point of view it is not possible to measure these covariances. However, this does not mean that we cannot assess the riskiness of a project. The basic idea proposed in Hertz's risk analysis is to analyze how the variability in input variables affects the NPV . If it is reasonable to assume that those projects that influence each other have input variables in common, then $P(NPV \leq 0)$ can be interpreted as an indication of the effect of the covariances. It is equivalent to what econometricians call a final form analysis. $P(NPV \leq 0)$ could then be seen as an overall measure for the effect of the covariances on the NPV . However, there are better and more direct ways to measure this effect.

But even if we reject all objections against risk analysis based on stochastic simulation, and we do construct a probability distribution for the NPV , we still have a problem. Suppose we can tell the decision maker that with ten percent chance the NPV of the project will be negative. The first question the decision maker will then ask is: *when* will this negative NPV

happen? Stochastic simulation in itself does not give an answer to this question. One could, of course, store all simulation runs that yield $NPV < 0$, but information on what causes $NPV < 0$ will not be readily available and will require further analysis. The same information can be obtained more easily through DOE.

Note: Besides the economic risk, there are risks related to the analysis, namely changes in the model's parameters, constraints, and model structure [19]. Not all these aspects are equally important. It is easier to analyze changes in parameters than it is to analyze the effects of an alternative model structure. The structure of the model is normally not included in risk analyses, but is tested through model validation; also see Kleijnen *et al.* [20]. However, model validation assumes that there is information on the real system. For an investment project the real system does not exist yet, so there are no data on its past behaviour. This paper assumes that the simulation model is an adequate description of the investment project; see Ref. [21] for validity in general. So we restrict our analysis to factor changes, such as changes in input variables, parameters, and constraints.

3 SENSITIVITY ANALYSIS VS HERTZ'S MONTE CARLO RISK ANALYSIS

In this section we restrict our analysis to the question: does the existence of economic risk require stochastic project analysis, deterministic project analysis, or both? The answer to this question is not straightforward. There are practical as well as theoretical considerations in support of applying both analyses. Practical guidelines for project evaluation discuss only sensitivity analysis based on one-factor-at-a-time experimental design (Duvigneau and Prasad [22], pp. 22–3) and scenario analysis (UNIDO [23], pp. 188–9). Others state that sensitivity analysis is the only way to estimate the effects of a change in assumptions about input variables, whereas *the likelihood of the effect has to be judged by the decision makers* (Gittinger [24], p. 369). As we shall show, this view can indeed be supported theoretically; moreover, it is more practical, and it better fulfills the information requirements of decision makers.

3.1 Advantages and disadvantages of Monte Carlo risk analysis

In Section 2 we showed that stochastic simulation that yields a probability distribution of the NPV as proposed by Hertz [1] and Hillier [2], does not indicate a project's risk: for risk assessment only the relation with other projects and income is of interest. This conclusion raises the question: does our analysis

imply that stochastic simulation has no value in project evaluation? The answer is: no, stochastic simulation can certainly have merits. Myers [25] suggested to adjust risk simulation, and to apply it only to the distribution of the cash flow per year, in order to obtain the correct $E(NPV)$. Hertz and Thomas [12] admitted that the distribution of the NPV has no clear meaning, but they defended its use with the remark that similar problems arise with other approaches. To overcome problems when applying the probability distribution of the NPV to measure risk, they suggested to use the probability distributions of the $NPVs$ of several projects in order to categorize the riskiness of different projects. However, we showed that the drawbacks of risk analysis for one project cannot be overcome in this way, because it are the covariances and not the marginal distributions that are important. They further stated that stochastic simulation should be used to develop the understanding of a project. However, we doubt the usefulness of stochastic simulation in case the distributions of the factors are unknown, and have to be based on assumptions. To explain this doubt, we review the main arguments in favour of stochastic simulation: (i) interdependencies among variables, (ii) uncertainty in project life, and (iii) asymmetries in cash flow distributions.

Sub (i): Interdependencies among factors are often used as an argument in favour of stochastic simulation, when assessing the correct $E(NPV)$ of an investment. It is well known that the expected value of a nonlinear relation among stochastic variables is not equal to the value obtained when replacing those variables by their expected values. So this substitution may lead to an incorrect interpretation of the expected value of the cash flow [1]. But does this problem call for stochastic simulation? In case stochastic simulation is applied to obtain the probability distribution of the cash flows, we have to make the (joint) distribution of all factors explicit. Then the expected cash flows can be calculated. These expected cash flows may be introduced into the NPV calculation to obtain the expected NPV , since the NPV calculation is a linear function in the cash flows.

In case the expected cash flows can not be obtained explicitly, stochastic simulation is required. However, the reliability of the result obtained through simulation depends strongly on the knowledge of the distribution of the factor values. Assuming an arbitrary distribution does in general not improve knowledge about the mean cash flows. Furthermore, to avoid costs and to meet time requirements [26], none of the advocates of stochastic simulation suggest the introduction of a distribution for *all* input variables; instead they propose a distribution for the most important ones (no more than ten inputs).

Hertz and Thomas ([12], pp. 306–7) argued that a

thorough training of decision-makers and staff members, in the concepts and meaning of probability is a prerequisite for risk analysis. Only then, better probability distributions can be obtained. Whether this is ever achieved, is questionable, as we know from empirical research in the field of expected utility theory and rational choice [27]. Cooper and Chapman ([28], Chapter 6) discussed methods for eliciting probabilities in risk analysis; they review possible biases. The procedure for minimizing biases is very labour intensive and requires much skill from both the analysts and the experts in a field. We wonder if these results add to understanding the problem. The techniques are certainly not easy to apply. Also the costs involved may be prohibitive. Furthermore, the tails of the resulting distributions are the most unreliable part; but for Hertz's risk analysis, these are the parts of the distribution that matter most, when the project is evaluated.

Sub (ii): Uncertainty in project life is used as an argument for stochastic simulation. However, Lewellen and Long [26] showed that the effect of project life on the NPV is small, especially when project life is long (as is the case for the investment projects we investigate here in this paper). We add that in many evaluations the project life (planning horizon) is fixed by the decision makers; in our case it is twenty years.

Sub (iii): Asymmetries in cash flow distributions is a problem that can disturb results. In case the experts state the modes (most likely values) of the cash flows instead of the means, the calculated NPV is no longer the mean (assuming independent variables). This can cause serious problems [1]. However, if we can obtain the mean cash flow per period instead of the mode, then we can use this mean to compute $E(NPV)$; so stochastic simulation is not necessary.

The arguments in favor of stochastic simulation are not very convincing; they strongly rely on the quality of the probability distributions used. Stochastic simulation for public investment projects requires much judgement by the analysts and other people involved, on topics that are generally not well understood, even by trained people. The NPV results depend on the input distributions. In case reliable input distributions are available, these distributions should be used. Otherwise, further research on the usefulness of stochastic simulation for the determination of expected cash flows is required, before stochastic simulation should be applied in practice.

Brealey and Myers ([10], p. 227) argued that a realistic stochastic simulation is very complex. The model builders understand their own creation, but this will not be the case for the decision makers. Therefore the latter will not rely on the output. This conclusion is supported by Hertz and Thomas [12] and Bower and Lessard [29]. The latter analyzed the investment

decision process in seven large companies. Their conclusion was that the cause of not using sophisticated measures for project evaluation 'was the inability to translate results into simple measures executives could reconcile with their intuition and experience and use, with other measures, to make a judgement'. What they did find is that the analysts put great effort in finding the correct figures from which the simple measures were calculated; as a result these measures were quite reliable.

3.2 Hertz's risk analysis and the discount rate

Consider a situation with a single project. We can then ask whether or not Hertz's risk simulation is a reliable instrument for project evaluation. Even in this case there are arguments against the application of stochastic simulation (to assess a project's risk in the form of a probability distribution of the project's *NPV*). The definition of the *NPV* implies that the expected cash flows are discounted at a discount rate that expresses valuation of future money in terms of the present. This discount rate is the (economic) opportunity cost of capital. But that rate is adjusted for risk, because it is usually defined as the rate that would have been obtained in the next best acceptable project. So to obtain a distribution of the *NPV* that expresses uncertainty about the *NPV* correctly, we should discount by the risk-free rate (normally set equal to the long-term government bond rate). The risk-free rate, however, is not independent of the future. It can be shown that in case there are several possible futures, the risk-free rate is independent of these future states of the world after period one, which is the same as not treating uncertainty at all (Brealey and Myers [10], p. 228). For a formal proof in the context of a time-state preference model we refer to Keeley and Westerfield [30]; also see Copeland and Weston ([17], pp. 116–9). Only if the *NPV* is independent of total income (and thus the future), the risk-free rate is the appropriate rate; see [16, 31].

A popular approach to account for risk is the capital asset pricing model (CAPM), which proposes to increase the risk-free rate to account for risk. The argument is that if there is one project with certain returns (long term state bonds) and one project with risky returns, then the decision maker will prefer the risky project only if $E(NPV)$ is sufficiently high compared with the certainty equivalent *NPV*. This implies that for a risky project a higher rate of return is required. Lind ([31], pp. 65–7), however, states that discounting with one rate in a multi-period setting requires risk to be such that the ratio of certainty equivalent net benefits and project net benefits decreases in an unrealistic way. Indeed Hertz and Thomas ([12], p. 296) agreed that the meaning of the

NPV distribution at a risk-free rate is unclear. They defended their use of the distribution of the *NPV* with the argument that other practical approaches to risk (such as the adjusted discount rate) are problematic too, since the certainty equivalent argument shows that no single adjustment to the discount rate can be made that correctly accounts for risk; see [32] and [31] (p. 67).

Our critique on Hertz's Monte Carlo risk analysis might suggest that the effects on the *NPV* of changes in factors (input variables and parameters) are of no interest to decision makers. Decision makers, however, are certainly interested in such effects, and sensitivity analysis is normally required in a feasibility study. However, these results should not be given in terms of a probability distribution of the *NPV*, but in terms of the relation between the project's *NPV* and the factors. This is what decision makers want to know, since these relations indicate the riskiness of the project.

Obtaining this information is exactly what traditional sensitivity analysis tries to do. Modelling makes the investment problem no longer a black box. Decision makers want to know how the economic prospects interact with the investment project, and what the most important factors are. These factors deserve special attention during project execution. Decision makers will use the information, when monitoring the progress of the project, and when designing adjustments when progress is not as expected.

We therefore believe that decision makers are better supported by knowledge about (i) which individual factors are important, and (ii) which interactions are important. This knowledge gives the decision makers insight into their ability to react in time, and to adjust the project when unfavourable developments occur. Sensitivity analysis based on changing one factor at a time gives some answers to the first question, but not to the second question. Neither does stochastic simulation proposed in Hertz and Hillier's risk analysis, answer these questions.

The use of DoE, based on statistical theory, is an alternative for both the one-factor-at-a-time design and stochastic simulation analysis. To the best of our knowledge this alternative has never been tried in investment analysis. A well chosen design allows the assessment of the link between the project and the rest of the economy, which indicates the project's riskiness.

In the next sections we shall show how sensitivity analysis can be improved through using DOE.

4 IMPROVED SENSITIVITY ANALYSIS THROUGH DOE

Information on all main effects and possibly some interactions can be obtained by applying an adequate

design. Suppose for the time being, that there are k factors with no interactions at all. Then the *full factorial* design needs 2^k observations to estimate only $k + 1$ effects, namely k main effects β_k plus the overall mean β_o . In principle $k + 1$ observations suffice to obtain unique estimates of the $k + 1$ effects. In other words, for the $k + 1$ effects it may suffice to simulate only a fraction, namely 2^{-p} , of the 2^k observations such that $2^{k-p} \geq k + 1$. For example, if k equals 7 then $n = 2^{7-4}$ suffices. These designs are called 2^{k-p} designs.

However, 2^{k-p} designs have a number of runs equal to a power of two. A more general type of design is the *Plackett-Burman* design type [33]. These designs require a number of runs equal to a multiple of four. For example, for ten factors the Plackett-Burman design with twelve runs can be used; the transposed Plackett-Burman design matrix (say) D^T , where T stands for transpose, is given in Table 1. Every column of + and - signs in that table represents a combination of factors for a simulation run. A plus sign (+) stands for the base case value of the corresponding factor, and a minus sign (-) for the value that has a negative influence on the base case result. The assignment of a minus is based on a one-factor-at-a-time sensitivity analysis, of which the results are not reported here. Identifying the base case with (+, +, ..., +) means that the other runs focus on those conditions that will jeopardize the investment project (the base case has a positive value for the *NPV*). Hence all results are expected to be worse than the base case result. It is easily checked that the columns of the Plackett-Burman design matrix are orthogonal; thus $(D^T D)^{-1} = 12^{-1} I$, where + is interpreted as +1 and - as -1, and I denotes the identity matrix of proper dimension (here 10×10). It has been proved that orthogonal designs satisfy many optimality criteria, such as minimum variance of the estimated effects (say) $\hat{\beta}$. Furthermore, the design in Table 1 satisfies one linear constraint: the sum of the first eleven rows of D equals minus row twelve. The augmented matrix $X = (e; D)$, with $e = (1, 1, \dots, 1) \in R^{12}$ corresponding with β_o , has the same properties as the matrix D .

Consider the following first order regression metamodel (main effects only), which is an approximation of the input/output behavior of the simulation model:

$$Y_i = \sum_{h=0}^{10} \beta_h x_{ih} + \epsilon_i, \tag{6}$$

where ϵ denotes the approximation error. We use OLS to estimate the eleven main effects; OLS gives Best Linear Unbiased Estimators (BLUE) if ϵ is white noise. However, if the assumption of no interactions is false, the estimators are biased; see [34].

Let $\beta_M = (\beta_o, \beta_1, \dots, \beta_{10})^T$ be the vector of coefficients in model eqn (6). An unbiased estimator of the main effects β_M can be achieved by applying the Box-Wilson *foldover* theorem; see [35]. A foldover is obtained by adding $-D$ to the original design matrix D ; in the example twenty-four instead of twelve simulation runs are executed. The foldover of a Plackett-Burman design is a resolution IV design; that is, no main effect is confounded with any other main effect or any two-factor interaction; the two-factor interactions, however, are confounded with each other [36]. Obviously, with only roughly $2k$ runs no unbiased estimators of all k main effects and $k(k - 1)/2$ two-factor interactions are possible; in the example there are $1 + 10 + (10 \times 9/2) = 56$ effects and only 24 runs. Let $V = (V_1, \dots, V_9) \in R^{12 \times 45}$, with $V_i = (x_i x_{i+1}, \dots, x_i x_{10})$ and $x_i x_{i+1} = (x_{i1} x_{i+1,1}, \dots, x_{i,12} x_{i+1,12})^T$, be the matrix corresponding to the two factor interactions. Because V consists of combinations of the columns of D , the rank of V cannot exceed the rank of D ; V has the same rank as D . This rank is eleven, because the sum of the first eleven rows of this design matrix is equal to minus row twelve. So only up to eleven individual two-factor interactions can be estimated. Hence eleven independent columns from V have to be selected to form the matrix (say) V_f . The remaining columns of V are combined in the matrix (say) V_A . The resulting alias matrix for the eleven interactions is $(V_f^T V_f^{-1} V_f^T V_A)$; also see Kleijnen ([37], pp. 295-301). Note that if unbiased estimators of

Table 1. Plackett-Burman design for ten factors

Combination factor	1	2	3	4	5	6	7	8	9	10	11	12
1 investment costs	+	+	-	+	+	+	-	-	-	+	-	-
2 construction time	-	+	+	-	+	+	+	-	-	-	+	-
3 reserves West Java	+	-	+	+	-	+	+	+	-	-	-	-
4 real GVA	-	+	-	+	+	-	+	+	+	-	-	-
5 energy prices	-	-	+	-	+	+	-	+	+	+	-	-
6 relative gas/oil price	-	-	-	+	-	+	+	-	+	+	+	-
7 purchase prices	+	-	-	-	+	-	+	+	-	+	+	-
8 coal prices	+	+	-	-	-	+	-	+	+	-	+	-
9 other costs	+	+	+	-	-	-	+	-	+	+	-	-
10 discount rate	-	+	+	+	-	-	-	+	-	+	+	-

more than eleven interactions are needed, the design has to be further augmented (beyond the foldover), and additional simulation runs are required.

In the next section we discuss a case study to illustrate our approach to assess economic risk.

5 A CASE STUDY: JAVA'S GAS TRANSMISSION SYSTEM

Indonesia is one of the fast growing economies in the Far East. In the 1970s Indonesia's industrial exports were mainly natural gas and oil. Over the last fifteen years, however, other industries have emerged, which now account for over fifty percent of the exports. Many of these industries are located on Java, and in particular West Java. To produce heat, these industries use fuel oil and diesel oil. Indonesia's oil reserves, however, are almost exhausted, and without additional new findings Indonesia will become an oil importer instead of an oil exporter, early next century. Indonesia's gas reserves on the other hand are large. Furthermore, there are a number of small gas reserves in the vicinity of Java, containing gas which cannot be exported as liquefied natural gas, but can be used to develop the market for natural gas on Java. What the Indonesian government wanted to know is: how can a gas transmission system for Java be established, and under what circumstances will it be a positive contribution to the economy?

To analyze these questions a pre-design for Java's gas transmission system was developed, together with investment and operating cost estimates for each section of the system. At the same time the potential markets for gas were analyzed. The potential market and the transmission system have to be developed in phases, because a complete system without sufficient demand is too expensive. First, a gas transmission and distribution system is constructed in West Java to supply existing and new industries with gas. If the relative prices of gas and oil products are set according to their economic value (that is, gas is priced at fuel oil parity), new industrial investments will use gas instead of oil because the use of gas has a number of advantages over fuel oil. For many existing industries the fuel oil parity price is approximately the break-even price; that is, at this price conversion from oil to gas will not improve profits. Therefore, the gas company will initially have to give a discount to make conversion to gas profitable for existing industries. With or without conversion, the growth of the economy will result in a sufficiently large market after ten to fifteen years to utilize all gas reserves in the vicinity of Java.

When the utilization of the West Java gas reserves reaches its maximum, an investment in a transmission system to gas reserves in Kalimantan (1,000 miles

away) is required. This is the most expensive part of the investment project; it is only profitable when the gas market on Java is large enough to guarantee substantial demand. The profitability depends on two factors: market growth and sufficient local reserves to develop the market. After the connection to Kalimantan, the introduction of compressor stations can further boost the capacity of the system.

The gas reserves in the vicinity of Java play a crucial role in the development of the market. The content of gas reservoirs, however, is always an estimate, and the reliability of the estimate depends on the number of drills to estimate it. But even the most reliable estimate, known as proven reserves, states that there is a 90% chance that the amount of gas estimated is actually available.

To design and analyze this investment plan, a decision support system for analysing all possible options was build. The information on the different sub-problems (investment, markets, reserves, etc.) was combined in a cost-benefit analysis according to the World Bank standards for project analysis; see [22]. The ten factors in the first column of Table 1 indicate the main threats to the project. Each factor represents a multitude of different threats, which from an economic point of view can be treated as one single threat. For example, investment cost overruns can be caused by higher prices for materials or specialized equipments, or by the fact that there are more river crossings or more urbanized areas than expected. From an economic point of view it all results in more investment costs, and can therefore be treated as a single factor. From a technical point of view the different causes for investment cost changes do matter. Higher prices do not affect the technological risk, whereas changes in construction conditions do affect the technological risk of a pipeline section.

The Plackett-Burman design and its foldover are applied to these ten factors. The results of the twenty-four simulation runs for the *NPV* of the gas transmission project are shown in Table 2. Combinations 1 to 12 are identical to the twelve columns of Table 1. Because only negative factor values are simulated, there are many negative entries in the columns denoted by *D* and $-D$. Since each minus value of a factor lowers the *NPV*, the negative effects of the ten factors are mitigated only if the interactions among factors have positive influences. In case these interactions are negative, the results of the simulation will be even more negative. In the foldover part ($-D$) the combination corresponding with all factors at their 'plus' is the base case. So in the base case the *NPV* is maximal (1,750.9 billion Indonesian Rupiah).

The analysis of the *NPV* data of Table 2 starts with the OLS estimation of the first order approximation eqn (6). The OLS estimates $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_3, \hat{\beta}_5, \hat{\beta}_7,$ and $\hat{\beta}_g$ are significant at the level $\alpha = 0.05$ ($t_{13}^{0.05} = 1.77,$

Table 2. NPV results for the Plackett-Burman design (D) and its foldover (-D)

Comb.	Factor										NPV	
	1	2	3	4	5	6	7	8	9	10	D	-D
1	+	-	+	-	-	-	+	+	+	-	-1,252.8	-1,268.0
2	+	+	-	+	-	-	-	+	+	+	-3,033.8	-67.9
3	-	+	+	-	+	-	-	-	+	+	-1,132.5	-1,006.0
4	+	-	+	+	+	-	-	-	+	+	-1,210.0	-1,084.0
5	+	+	-	+	+	-	+	-	-	-	-1,301.7	-1,339.6
6	+	+	+	-	+	+	-	+	-	-	359.0	-2,639.2
7	-	+	+	+	-	+	+	-	+	-	-997.8	-1,362.2
8	-	-	+	+	+	-	+	+	-	+	454.2	-1,353.7
9	-	-	-	+	+	+	-	+	+	-	-602.6	-2,045.6
10	+	-	-	-	+	+	+	-	+	+	-175.6	-1,324.4
11	-	+	-	-	-	+	+	+	-	+	-1,341.7	-278.1
12	-	-	-	-	-	-	-	-	-	-	-2,985.3	1,750.9

$13 = (2 \times 12) - (1 + 10)$). A popular measure for the fit of the regression model is R_{adj}^2 which is 0.88. Note that because the matrix D is orthogonal, the values of the estimates of the main effects do not change when regressors are deleted or added.

Once the relevant main effects are known, significant two-factor interactions can be determined. There are many possible ways to augment the first-order metamodel in eqn (6) with interactions. We might arbitrarily take the first eleven two-factor interactions $\{\beta_{1,2}, \beta_{1,3}, \dots, \beta_{2,6}\}$ and test the hypothesis $H_0 = \beta_{1,2} = \beta_{1,3} = \dots = \beta_{2,6} = 0$ through an F-test on model reduction (Kleijnen, [37], pp. 156-7). The F-value for this NPV model is insignificant; so H_0 is accepted. However, since the two-factor interactions are confounded, *failure to reject H_0 does not prove that there are no interactions!* It only says that the sum of confounded two-factor interactions is not significant. Testing all possible ways to augment eqn (6) is not a feasible option. What we need is a strategy to choose possible interactions.

A popular assumption is that there are interactions only between factors with *significant* main effects. This is a reasonable approach when no other information is available; the simulation model is then treated as a black box. But in this case (and in many other cases) this approach is not necessary. Some clues can be derived from the simulation model itself and from the intermediate simulation results that lead to the outputs in Table 2. That table gives only NPV, not any more details of the simulation runs. In the simulation model, economic growth plays an important role. However, economic growth (factor 4) has no significant effect on the NPV. This seems odd, and also conflicts with economic theory. After we studied the detailed simulation results of the twenty-four simulation runs, we concluded that economic growth does strengthen the effect of a change in the West Java reserves (factor 3), and it also strengthens the effects of some of the changes in prices (factors 5, 6, 7,

and 8). Therefore we restricted the search for interactions to these six variables (factors 3 through 8). After testing several alternative specifications using a form of forward selection ([38], pp 71-2), the following model gave the best test results:

$$NP\hat{V} = -1051.6 + 142.5x_1 + 461.2x_3 + 659.2x_5 \\ + 447.2x_6 + 242.7x_7 + 236.4x_8 + 226.0x_3x_4 \\ + 112.1x_3x_5 - 107.7x_3x_6 + 128.6x_3x_8. \quad (7)$$

This model has main effects equal to the significant main effects in the first-order model. It has $R_{adj}^2 = 0.98$. The hypothesis $H_0: \beta_{3,4} = \beta_{3,6} = \beta_{3,8} = 0$ yields $F_{13}^4 = 10.81$, and is significant even at the 0.5% level ($F_{13;0.005}^4 = 6.23$).

The validity of approximation eqn (7) can be tested through *cross-validation*. Cross-validation means that factor input combinations are eliminated one by one, the regression model is re-estimated, and the resulting model is used to predict the simulation result for the combination eliminated [39, 40]. To indicate the quality of the predictions obtained through cross-validation, we use a scatter plot (see Fig. 1). If eqn (7) were perfect, the scatter plot would be a straight line. This performance can be quantified by the correlation coefficient between \hat{y}_{-i} and y_i , which is 0.996.

6 CONCLUSIONS

Hertz's criterion $P(NPV \leq 0) \leq \alpha$ for the riskiness of an investment project is equivalent to the criterion for technological risk. However, technological risk and economic risk differ essentially. Contrary to technological risk, economic risk is based on the relation between the project analyzed and related projects and national income. Hertz's risk analysis has the following theoretical and practical flaws. In general no adequate information on the (joint) probability distributions of the input variables for a large

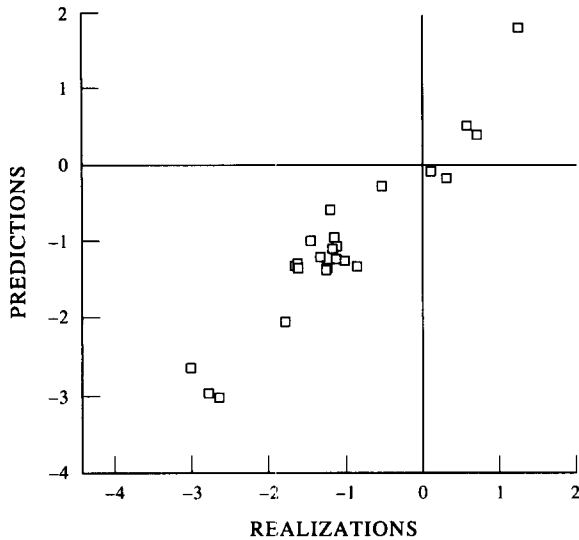


Fig. 1. Scatter plot of NPV regression predictions and simulation realizations.

investment project is available. So it is not an adequate method for assessing this risk. To get information on how the project is related to other projects, we analyzed the effect of changes in input variables and other factors on the NPV. Risk analysis, which results in a probability distribution of the NPV, does not indicate explicitly which factors and factor interactions are important. That information, however, is important for decision makers. Sensitivity analysis seems a better alternative.

Traditional one-factor-at-a-time sensitivity analysis identifies only main effects, not interactions. Such analysis does not meet the information needs of decision makers with respect to a project's risk. They want to know explicitly what the most important factors are and how their effects relate to other factors. These factors deserve special attention during project execution. Decision makers can use this information, when they monitor the progress of the project, and they must design adjustments in case progress is not as expected.

To obtain the required information we advocated design of experiments (DOE) to plan simulation experiments that give BLUE (best linear unbiased) estimators of the main effects, and selected interactions. With respect to risk as defined in economic theory, this is the closest a practitioner can get to indications for the covariances between the project analyzed, and other projects and group income. For our case study (gas transmission in Indonesia) we applied the foldover of the Plackett-Burman design to determine what the most important factors are. So for this investment project with its ten factors, only 24 runs were needed to obtain estimates for the main effects and the most important interactions. One might argue that the interpretation of the interactions is not supported by a complete statistical analysis,

since not *all* two-factor interactions were systematically checked. However, analysts who understand their problem, will normally be able to qualitatively derive which interactions are important. The resulting regression metamodel is indeed statistically sound, and it is supported by knowledge about the problem at hand.

REFERENCES

- Hertz, D. B., Risk analysis in capital investment. *Harvard Business Review*, 1964, 95-106.
- Hillier, F. S., The derivation of probabilistic information for the evaluation of risky investments. *Management Science*, 1963, 9, 443-457.
- Reutlinger S., *Techniques for project appraisal under uncertainty*. World Bank Staff Occasional Paper Nr. 10, International Bank for Reconstruction and Development, John Hopkins Press, Baltimore, USA, 1970.
- Pouliquen, L. Y., *Risk analysis in project appraisal*. World Bank Staff Occasional Paper Nr. 11, International Bank for Reconstruction and Development, John Hopkins Press, Baltimore, USA, 1970.
- Box, G. E. P., Hunter, W. G. and Hunter, J. S., *Statistics for experimenters: An introduction to design, data analysis, and model building*. John Wiley and Sons, New York, 1978.
- Sheng, A. and Yoon Je Cho, Risk management and stable financial structures. Policy Research Working Paper, WPS 1109, World Bank, Washington, DC, USA, 1993.
- Kolbe, A. L., Tye, W. B. and Myers, S. C., *Regulatory risk: economic principles and applications to natural gas pipelines and other industries*. Kluwer Academic Publishers, Boston, 1993.
- Taha, H. A., *Operations Research*. MacMillan Publishing Company, New York, 1989.
- Hull, J. C., *The evaluation of risk in business investment*. Pergamon Press, Oxford, Great Britain, 1980.
- Brealey, R. A. and Myers, S. C., *Principles of corporate finance*, Fourth edition. McGraw-Hill, Inc., New York, USA, 1991.
- Eschenbach, T. G., Spiderplots versus tornado diagrams for sensitivity analysis. *Interfaces*, 1992, 226, 40-46.
- Hertz, D. B. and Thomas, H., *Risk analysis and its applications*. John Wiley and Sons, Chichester, Great Britain, 1983.
- Ho, S. S. M. and Pikes, R. H., The use of risk analysis techniques in capital investment appraisal. In *Risk: analysis, assessment and management*, ed. J. Ansell and F. Wharton. John Wiley and Sons Ltd., Chichester, England, 1992.
- Kunreuther, H. C., Linnerooth, J., Lathrop, J., Atz, H., Macgill, S., Mandl, C., Schwarz, M. and Thompson, M., *Risk analysis and decision processing: the siting of liquefied energy gas facilities in four countries*. Springer-Verlag, Berlin, Germany, 1983.
- Quiggin, J., *Generalized expected utility theory*. Kluwer Academic Publishers, Boston, USA, 1993.
- Arrow, K. J. and Lind, R. C., Uncertainty and the evaluation of public investment decisions. *American Economic Review*, 1970, 60, 364-378.
- Copeland, T. E. and Weston, J. F., *Financial theory and corporate policy* (third edition). Addison-Wesley Publishing Company, Reading, Massachusetts, USA, 1992.

18. Wilson R., Risk measurement of public goods. In *Discounting for time and risk in energy policy*, ed. R.C. Lind. Resources For The Future, Inc, Washington, DC, USA, 1982, pp. 205–49.
19. Kleijnen, J. P. C., Sensitivity analysis and related analyses: a survey of statistical techniques. Paper presented at the International Symposium SAMO95, organized by European Commission DGII/F, 25–27 September, Belgirate, Italy, 1996.
20. Kleijnen, J. P. C., Bettonvil, B. and Van Groenendaal, W. J. H., Validation of simulation models: regression analysis revised. Center Discussion Paper No, Tilburg University, 1995.
21. Kleijnen, J. P. C., Verification and validation of simulation models. *European Journal of Economic Research*, 1995, **82**, 145–162.
22. Duvigneau, J. C. and Prasad, R. N., Guidelines for calculating financial and economic rates of return for DFC projects. World Bank Technical Paper Nr. 33, World Bank, Washington, DC, USA, 1984.
23. UNIDO, *Manual for the preparation of industrial feasibility studies*. United Nations Industrial Development Organization, Vienna, Austria, 1986.
24. Gittinger, J. P., *Economic analysis of agricultural projects*. John Hopkins University Press, Baltimore and London, 1982.
25. Myers, S. C., Postscript: using simulation for risk analysis. In *Modern developments in financial management*, ed. Myers. The Dryden Press, Illinois, 1976, pp. 457–63.
26. Lewellen, W. G. and Long, M. S., Simulation versus single-value estimates in capital expenditure analysis. In *Decision Sciences*, Vol. 3, ed. S.C. Myers. *Modern developments in financial management*. The Dryden Press, Illinois, 1972, pp. 442–56.
27. Tversky, A. and Kahneman, D., Rational choice and the framing of decisions. In *Decision making: Descriptive, normative, and prescriptive interactions*, ed. D.E. Bell, H. Raiffa and A. Tversky. Cambridge University Press, Cambridge, Great Britain, 1988.
28. Cooper, D. F. and Chapman, C. B., *Risk analysis for large projects*. John Wiley and Sons, Chichester, Great Britain, 1987.
29. Bower and Lessard, An operational approach to risk-screening. *Journal of Finance*, 1973, **27**, 321–337.
30. Keeley, R. H. and Westerfield, R., A problem in probability distribution techniques for capital budgeting. *Journal of Finance*, 1973, **27**, 703–709.
31. Lind, R. C., A primer on the major issues relating to the discount rate for evaluating national energy options. In *Discounting for time and risk in energy policy*, ed. R.C. Lind. Resources For The Future, Inc, Washington, DC, USA, 1982, pp. 205–49.
32. Myers, S. C. and Turnbull, S. M., Capital budgeting and the capital asset pricing model. *Journal of Finance*, 1977, **X**, 321–333.
33. Plackett, R. L. and Burman, J. P., The design of optimum multifactorial experiments. *Biometrika*, 1946, **33**, 305–325.
34. Raktoc, B. L., Hedayat, A. and Federer, W. T., *Factorial designs*. John Wiley and Sons, New York, USA, 1981.
35. Box, G. E. P. and Wilson, K. B., On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society, Series B*, 1951, **13**, 1–38.
36. Box, G. E. P. and Hunter, J. S., The 2^{k-p} fractional factorial design. *Technometrics*, 1961, **3**, 311–449.
37. Kleijnen, J. P. C., *Statistical tools for simulation practitioners*. Marcel Dekker, New York, USA, 1987.
38. Brook, R. J. and Arnold, G. C., *Applied regression analysis and experimental design*, Marcel Dekker Inc., New York, 1985.
39. Stone, M., Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society, Series B*, **36**, 1974, 111–147.
40. Kleijnen, J. P. C., Regression metamodels for simulation with common random numbers: comparison of validation tests and confidence intervals. *Management Science*, 1992, **38**, 410–417.