Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle

Itzhak Venezia

Optimal sequencing of capacity expansion projects under uncertainty

Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle, tome 16, nº 4 (1982), p. 287-307.

<http://www.numdam.org/item?id=RO_1982__16_4_287_0>

R.A.I.R.O. Recherche opérationnelle/Operations Research
(vol. 16, n° 4, novembre 1982, p. 287 à 307)

OPTIMAL SEQUENCING OF CAPACITY EXPANSION PROJECTS UNDER UNCERTAINTY (*) (**)

by Itzhak VENEZIA ([1])

Abstract. — *In this paper we present an optimal policy for sequencing a finite set of capacity expansion projects when the demand is uncertain and growing with time. By a sequencing policy we mean a rule for determining the order and timing in which the projects should be introduced.*

We analyze the effects of greater uncertainty on the optimal policy and on the results obtained using the optimal policy. These effects depend on the shape of the cost functions. It is shown that with increased uncertainty, if the cost functions are linear one must wait for higher demands before the maximal capacity is reached. It is also shown that increased dispersion of the demand tends to decrease the minimal expected discounted costs. The same conclusions hold true when costs are linear up to some large output and in the case where costs are concave. Contrary to what might be expected, the above results are not necessarily reversed when costs are convex. Intuitive explanations are provided for all these results.

Keywords: Capacity Expansion; sequencing; uncertainty.

Résumé. — *Dans cette étude, nous présentons une politique optimale pour déterminer la séquence d'un ensemble fini de projets d'expansion de capacité quand la demande est incertaine et croissante en fonction du temps. Il s'agit d'établir une règle déterminant dans quel ordre et à quel moment les projets devraient être introduits.*

Nous analysons les effets d'une incertitude accrue sur la politique optimale et ses résultats. Ces effets dépendent de la forme des fonctions coût. On montre que dans le cas d'une incertitude croissante et de fonctions coût linéaires, on doit attendre que les demandes soient plus fortes avant d'atteindre la capacité maximale. On montre également qu'une dispersion accrue de la demande tend à diminuer la valeur de l'espérance du coût total amorti.

Les conclusions restent identiques quand les fonctions coût sont concaves, et dans le cas ou elles sont linéaires jusqu'à une certaine limite supérieure. Contrairement à ce que l'on pourrait penser, les résultats ci-dessus ne sont pas nécessairement inversés quand les fonctions coût sont convexes.

Des explications intuitives accompagnent tous ces résultats.

1. INTRODUCTION

In this paper we investigate the optimal strategy for sequencing capacity expansion projects facing uncertain growing demand. We consider both the optimal determination of the order in which projects must be introduced and

(*) Received June 1980.

(**) The author wishes to thank Professor Charles Tapiero for his helpful comments and suggestions.

([1]) The Hebrew University of Jerusalem and the University of California, Berkeley, California 94720, U.S.A.

the optimal timing of such introductions. The optimality criterion used here is minimization of the expected discounted costs involved in satisfying the demand.

Problems of this type under certainty have been extensively discussed in the literature (*see*, e. g., Butcher *et al.* [1969], Erlenkotter [1973], Erlenkotter and Trippi [1976], Manne [1972]). Less attention has been paid to the case where uncertainty prevails (*see*, e. g., Manne [1961], Tapiero [1973], [1979], Giglio [1970]). These publications have usually treated capacity as continuous and assumed that capacity increments of any size are possible. Most of them considered regenerative or recurrent expansion projects and assumed that the demand is described by a Wiener process (thus allowing for a negative demand). Another feature common to papers concerning capacity expansion is the treatment of capacity as a completely rigid notion.

Here we present a different framework, which may fit a large class of capacity expansion projects. Our system services some uncertain growing demand and is composed of several projects, introduced sequentially into the system. Our analysis is applicable to the following kinds of systems:

(*a*) A water system servicing some region, each project representing a dam;

(*b*) A transportation facility such as a highway or an airport, each project representing additional lanes or runways;

(*c*) A power system generating electricity, each project representing a power plant.

The demand for the services of the system (henceforth, the demand) is a random variable, always depending on economic conditions and on time. At any period, the economic conditions may be in one of a finite number of states (depression, economic boom, and so on), which evolve over time following a Markov chain process. Conditional on the state of the economy, the demand is a random variable depending on a multitude of unpredicted factors such as relative prices, changes in tastes, world events, weather conditions, etc. The demand is growing, in the sense that, for any state of the economy, the sequence of distribution functions of the demand shifts to the right.

We assume that the system can always satisfy the demand. However, the costs involved in so doing depend on the available capacity relative to the demand. We assume that, the larger the capacity, the higher the fixed costs and the lower the marginal costs of supplying the demand. The costs include both pecuniary and nonpecuniary costs incurred by the users of the system, the latter category including poor quality of water, road congestion, low reliability of the power system, etc. Thus, our model is applicable to public projects as well as to profit motivated projects. Since there is no agreement in the literature concerning the appropriate shape (linear, concave, or convex) of a representative cost function (*see* Walters [1963], pp. 48-52) our model has the advantage that it can accommodate any shape of cost function.

In Section II we assume that the order in which projects must be introduced into the system is predetermined by engineering or location constraints. Under this assumption, we concentrate on finding the optimal timing for the introduction of new projects. It is shown that the optimal sequencing policy is characterized by certain sequences of critical numbers and critical intervals (nonoverlapping and exhaust the real line to the right of the critical values) which are functions both of time and of the economic state and capacity at that time. In each period one must compare the current demand with the appropriate critical value. Then, if the demand exceeds the critical value, capacity must be increased. Otherwise, the decision should be postponed to the next period. Given that demand exceeds the appropriate critical value, one can determine according to the critical interval that covers the demand, which project or projects should be added to the system.

In Section III we analyze the effects of increased uncertainty (dispersion, risk) on the optimal policy and on the results obtained using the optimal policy. In the literature only Manne [1961] investigated this problem. In his model, treating the case where backlogs of demand are allowed (the case most similar to ours, and most relevant for practical purposes), Manne (*ibid.*, p. 648) has shown that when the costs of backlogged demand are linear, increased variance has an undetermined effect on the minimum expected discounted costs. No intuitive explanation has been provided for this result. Here we shall analyze the effect of increased uncertainty using a broader (than variance) definition of increased risk, and investigate this effect under various assumptions about the shape of the cost functions. It is shown that with increased uncertainty, if the cost functions are linear, one must wait for higher demands before the maximal capacity is reached. Surprisingly, the expected time that elapses until maximal capacity is reached may not increase with greater uncertainty. The reason for this result is that greater uncertainty implies a higher probability of obtaining high values for the demand and thus of exceeding the critical values. This may offset the effect of the higher critical numbers.

Another surprising result is that the expected discounted costs decrease with increasing uncertainty. An intuitive reason for this result is that greater uncertainty implies a higher probability of obtaining extreme values. An increase in the probability of high values tends to increase the expected costs, while an increase in the probability of low values tends to decrease these costs. At first glance, it seems that these two effects should cancel out. The first effect, however, is somewhat weaker than the second, because when a high demand is observed its influence on the costs may be partially offset by expanding capacity. This net effect on expected costs will be called in the sequel the "sequential decision factor".

The same conclusions are reached also when costs are concave, or linear up to some large output. Contrary to what might be expected, the above results are not necessarily reversed when costs are convex.

In Section IV we drop the assumption that the order of introducing new projects is predetermined, and present an optimal strategy for simultaneously determining the order in which projects must be introduced and timing such introductions.

2. THE MODEL

Suppose a system services the demand for some product (henceforth, the demand). In order to satisfy this demand, the capacity of the system may be increased by introducing projects $1, 2, \ldots, K$. The problem considered here is to determine the optimal timing for introducing the projects into the system. It is assumed in this section that the order in which the projects are introduced is predetermined, given by $1, 2, \ldots, K$. Several projects however can be added at the same time. The optimality criterion is minimization of the expected discounted costs involved in servicing the demand, where the discount factor is some known $\beta < 1$.

In any period the demand depends on the economic state at that time. There are I possible states of the economy, $s_1, s_2, \ldots, s_I$. These states evolve over time, following a stationary Markov chain process. That is, the probability π_{ij} that the economy is in state s_j at time $t+1$, given that it was in state s_i at time t, is the same for all t. Conditional on the state s_i of the economy, the demand X_t^i is a random variable with cumulative distribution function (CDF) $F_{it}(x)$ and expectation μ_{it}. The growth trend in the demand is represented by the assumption:

$$F_{it}(x) \geqq F_{i,t+1}(x) \quad \text{for all } t. \qquad (2.1)$$

It is further assumed that for $i = 1, 2, \ldots, I$, the sequences $\{ F_{it}(x) \}$ converage to $F_i(x)$ (1).

We denote by $c_k(x)$ the costs of satisfying the demand given that the system includes projects $1, 2, \ldots, k$ (in this case we say, elliptically, that capacity

(1) It has been shown in some empirical studies, for situations fitting our model, that the demand indeed converges. *See*, e. g., Rausser [1976], p. 326.

is k). The costs of installing projects $k+1, \ldots, m$, given that existing capacity is k, are denoted by M_{km} and we define $M_{kk}=0$. The effect of increasing capacity on costs is represented by assuming that for any demand, x, the higher the fixed costs and the lower the marginal costs of satisfying this demand. That is:

$$\left.\begin{array}{ll} c_k(0) < c_{k+1}(0), & k=1, \ldots, K-1 \\ dc_k(x)/dx > dc_{k+1}(x)/dx, & x \geqq 0. \end{array}\right\} \qquad (2.2)$$

We also have to assume that $E[c_K(x)]$ is finite [where the expectation is taken with respect to $F_i(x)$], since otherwise no meaningful evaluations of expected costs can be made. It is also assumed that for all $k < m \leqq K$, there exists an x^*_{mk} such that $c_k(x) > c_m(x)$ whenever $x \geqq x^*_{mk}$ and $F_m(x^*_{mk}) < 1$ for some i. This assumption simply states that for some relevant demands, the costs when capacity is m, are lower than the costs when capacity is k. Unless this assumption is made, there may be no justification for considering capacity m in our analysis. The above two assumptions imply that two cost functions corresponding to different capacities cross exactly once. A possible description of these functions is provided in figure 1. The way cost functions of various

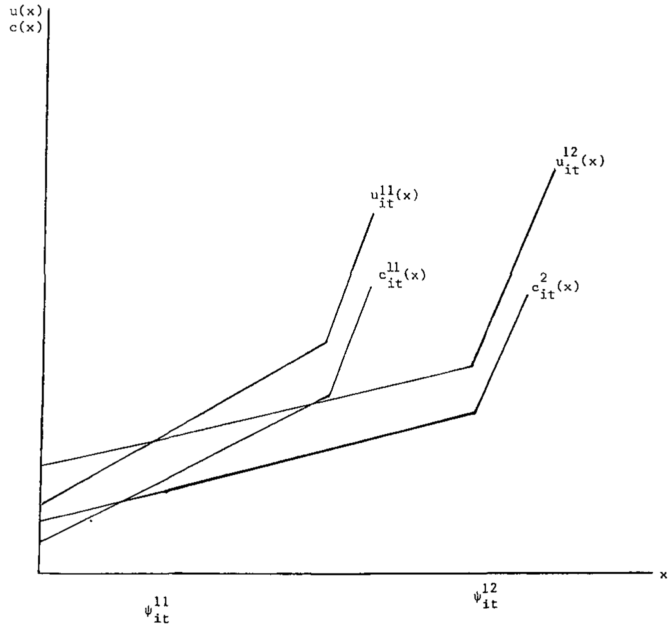

Fig. 1

projects can be aggregated depends on the nature of the projects at hand. When the projects are plants producing some product, the aggregation can be achieved by optimally allocating output among plants. In service systems when the main advantage of increasing capacity is the reduction of delays to customers and greater reliability of the system (e. g., transportation systems, computer facilities, etc.), one must compute the delays and waiting times under various capacities, and compute the costs accordingly.

The optimal expansion strategy is described in:

PROPOSITION 2.1: *For any state of the economy s_i, any period t, and any capacity k, there exist integers m_j, $j=1, 2, \ldots, S$, satisfying* [2]:

$$k<m_1<m_2<\ldots<m_S \leqq K-1,$$

and parameters $\psi_{it}^{km_j}$ satisfying:

$$0<\psi_{it}^{kk}<\psi_{it}^{km_1}<\psi_{it}^{km_2}<\ldots<\psi_{it}^{km_S},$$

such that if $x \leqq \psi_{it}^{kk}$ capacity should not be increased, if $\psi_{it}^{kk} \leqq x<\psi_{it}^{km_1}$ capacity should be increased to m_1, if $\psi_{it}^{km_1} \leqq x \leqq \psi_{it}^{km_2}$ capacity should be increased to m_2, etc.

Proof: Denote by $V_{it}(x)$ the expected discounted costs of operating the system starting at time t, given that the capacity is k, the economic state is s_i, the observed demand is x, and the optimal policy is always pursued. The expected discounted costs of operating the system, evaluated before the demand at t has been observed, are hence given by $W_{it}^k=E[V_{it}^k(x)]$, where E is the expectation operator with respect to the CDF $F_{it}(x)$.

It then follows from Bellman's [1957] principle of optimality that:

$$V_{it}^k(x)=\min_{m \geqq k}\{c_m(x)+M_{km}+\beta \overline{W_{i,t+1}^m}\} \equiv \min_{m \geqq k}\{u_{it}^{km}(x)\} \quad (2.3)$$

where:

$$\overline{W_{i,t+1}^m}=\sum_{j=1}^{I} \pi_{ij} W_{j,t+1}^m. \quad (2.4)$$

Since for any $m>k$, $c_k(x)$ intersects $c_m(x)$ at most once it follows that also the $u_{it}^{km}(x)$'s intersect at most once (*see* figure 1 for a typical description of these functions). Denoting by $\psi_{it}^{km_j}$ the point where $u_{it}^{km_j}$ and $u_{it}^{km_{j-1}}$ intersect, and noting the properties (2.2) of the cost functions, one can verify that the $V_{it}^k(x)$'s can be written as:

---

[2] The indices $m_1, \ldots, m_S$, and S may vary across different (i, t) combinations.

$$\left.\begin{aligned} V_{it}^{k}(x) &= u_{it}^{kk}(x) && \text{if} \quad x \leqq \psi_{it}^{kk} \\ &= u_{it}^{km_1}(x) && \text{if} \quad \psi_{it}^{kk} < x \leqslant \psi_{it}^{km_1} \\ &\;\vdots \\ &= u_{it}^{km_S}(x) && \text{if} \quad \psi_{it}^{km_S} < x < \infty. \end{aligned}\right\} \quad (2.5)$$

Our result follows from the definition of $V_{it}^{k}(x)$.

Q.E.D.

We next show how the minimal expected discounted costs W_{it}^{k}, and the critical values ψ_{it}^{km} can be computed. If $k=K$, no further expansions of capacity are possible, hence W_{it}^{K} denotes the maximal expected discounted costs assuming that capacity remains fixed at K, i. e.:

$$W_{it}^{K} = E\left[\sum_{\tau=0}^{\infty} \beta^{\tau} \sum_{j=1}^{I} \pi_{it}^{\tau} c_K(X_{t+\tau}^{j})\right] = \sum_{\tau=0}^{\infty} \beta^{\tau} \sum_{j=1}^{I} \pi_{ij}^{\tau} E[c_K(X_{t+\tau}^{j})], \quad (2.6)$$

where the π_{ij}^{τ}'s denote the τ-period lag transition probabilities ([3]). Convergence of the infinite sum in (2.6) is guaranteed since $E[c_K(x)]$ is finite.

When $k<K$, one observes from the definition of the W_{it}^{k}'s that:

$$W_{it}^{k} = E[V_{it}^{k}(x)] = \int_{0}^{\infty} V_{it}^{k}(x)\, dF_{it}(x) = \int_{0}^{\psi_{it}^{kk}} [c_k(x) + \beta\, \overline{W_{i,t+1}^{k}}]\, dF_{it}(x)$$

$$+ \int_{\psi_{it}^{kk}}^{\psi_{it}^{km_1}} [c_{m_1}(x) + \beta\, \overline{W_{i,t+1}^{m_1}} + M_{km_1}]\, dF_{it}(x)$$

$$+ \ldots + \int_{\psi_{it}^{km_S}}^{\infty} [c_{m_S}(x) + \beta\, \overline{W_{i,t+1}^{m_S}} + M_{km_S}]\, dF_{it}(x) \quad (2.7)$$

where the ψ's satisfy:

$$c_k(\psi_{it}^{kk}) + \beta\, \overline{W_{i,t+1}^{k}} = c_{m_1}(\psi_{it}^{kk}) + \beta\, \overline{W_{i,t+1}^{m_1}} + M_{km_1}, \quad (2.8)$$

$$c_{m_{S-1}}(\psi_{it}^{km_S}) + \beta\, \overline{W_{it}^{S-1}} + M_{km_{S-1}} = c_{m_S}(\psi_{it}^{km_S}) + \beta\, \overline{W_{it}^{m_S}} + M_{km_S}.$$

In what follows we show how these recursion relations can be used to compute the W_{it}^{k}'s for all $k<K$.

([3]) If π is the stationary $(I \times I)$ transition matrix, then π_{ij}^{τ} is the (i,j)th element of π^{τ}.

Since the W_{it}'s represent the costs of supplying the demand and since the demand is growing in time, the W_{it}^k's are nondecreasing in t. From the assumption that $E[c_K(x)]$ is bounded and since $F_{it}(x) \geqq F_i(x)$ for all t, it follows that the W_{it}^k's are bounded. Since these series of expected discounted costs are monotone and bounded they converge and we denote by W_i^k the limit of W_{it}^k, for all i and k. These limits satisfy recursion formulas similar to (2.7) and (2.8) except that one has to replace W_{it}^k and ψ_{it}^{km} by W_i^k and ψ_i^{km}, respectively. We shall next show how the W_i^k's can be computed, and that based on the W_i^k's, the W_{it}^k's can easily be obtained.

The W_i^k's are computed recursively. First one computes the W_i^K's as the limit, as t approaches infinity, of (2.6). Then, in order to compute the W_i^{K-1}'s, one notes that equations (2.7) and (2.8) become, after omitting the subscripts t, the following system of $2I$ equations in $2I$ unknowns $(\psi_i^{K-1}, W_i^{K-1}, i=1, \ldots, I)$.

$$W_i^{K-1} = \int_0^{\psi_i^{K-1}} [c_{K-1}(x) + \beta \overline{W_i^{K-1}}] \, dF_i(x) + \int_{\psi_i^{K-1}}^{\infty} [c_K(x) + \beta \overline{W_i^K} + M_{K-1,K}] \, dF_i(x), \qquad i=1, \ldots, I, \quad (2.9)$$

$$c_{K-1}(\psi_i^{K-1}) + \beta \overline{W_i^{K-1}} = c_K(\psi_i^{K-1}) + \beta \overline{W_i^K} + M_{K-1,K}, \qquad i=1, \ldots, I. \quad (2.10)$$

This system of equations can be solved numerically (*see* lemma A.1 in the appendix for an existence proof). The W_i^k's can then be similarly computed successively for $k=K-2, \ldots, 1$ using at each step (2.7), (2.8), and the W_i^{k+1}'s obtained in the previous stage.

Having computed the W_i^k's, one proceeds to compute the W_{it}^k's. For this we note that since the W_{it}^k's converge to W_i^k, by choosing a large enough $\bar{T}$, W^k can be made as close to W_i^k as one wishes. Substituting W_i^k for W_{iT}^k, (2.7) and (2.8) can be used to compute the $W_{i,\bar{T}-1}^k$'s, which in turn can be used to compute the $W_{i,\bar{T}-2}^k$'s, and similarly all W_{it}^k's, $< \bar{T}-1$ can be computed.

We thus have an optimal strategy for timing the introduction of new projects and a numerical method for computing the critical values ψ_{it}^{km}. Below we present some properties of the solution.

3. EFFECT OF INCREASED UNCERTAINTY ON THE OPTIMAL TIMING POLICY

We now evaluate the effect of increased uncertainty on the critical values ψ_{it}^{kK} and on the expected discounted costs W_{it}^k. We define increased uncertainty (dispersion, risk) as an increase in the risk of each X_t^i for all $i=1, 2, \ldots, I$ and $t=1, 2, \ldots, \infty$ (4). We employ the Rothschild-Stiglitz [1970] definition of increasing risk. We briefly review the definition. Let z_1 and z_2 be two random variables with cumulative distribution functions $H^1(z)$ and $H^2(z)$, respectively. Then z_1 is riskier than z_2 if $H^1(z)$ is obtained from $H^2(z)$ by taking weight from the center of the probability distribution and moving it to the tails, while keeping the mean of the distribution constant. Rothschild and Stiglitz have shown that this definition of increased risk is equivalent to two other definitions: that an increase in risk is the addition of white noise to a random variable (5) and that for all concave nondecreasing functions $g(z)$, $E_1[g(z)] \leq E_2[g(z)]$, where $E_j(.)$ denotes the expectation operator with respect to $H^j(z)$, $j=1, 2$. In the sequel we shall mainly make use of the latter definition.

The effect of increased uncertainty on the minimal expected discounted costs and on the optimal strategy depends, as will be shown, on the shape of the cost functions $c_k(x)$. Walters [1963] reviewed the literature dealing with the empirical measurement of cost functions. The main conclusion from his review is that the shape of cost functions differ considerably from industry to industry, and that it is hard to find agreement even about the shape of cost functions within an industry. In public utilities, however, there is relatively greater agreement. Walters (*ibid.*, p. 50) finds about short run cost functions in these industries that "... over the observed range of output, marginal cost is constant". In profit motivated industries one tends to conjecture that U shaped cost functions are the most prevalent ones. Surprisingly, constant marginal costs have been found in a large number of industries. Some of these findings however can be attributed to the fact that measured outputs were below capacity (*ibid.*, p. 51).

Consequently, we have decided to analyze the effect of increased uncertainty under alternative assumptions concerning the shape of the cost functions. We have chosen to start with the case where costs are linear at the relevant range of demands. This is a convenient starting point since it elucidates the effect of increased uncertainty, and the results of other cases can be obtained as simple corollaries of this case.

---

(4) According to our definition the transition probabilities π_{ij} do not change with increased uncertainty

(5) Increased risk thus implies an increased variance, but not vice versa.

In what follows we let r be a shift parameter dentoing risk, and introduce this parameter explicitly in the CDF of X_t^i which will now be denoted by $F_{it}(x, r)$. We denote by $W_{it}^k(r_0)$, $\psi_{it}^{km}(r_0)$ the minimal expected discounted costs and critical values when risk is r_0, and by $W_{it}^k(r_1)$, $\psi_{it}^{km}(r_1)$ the same variables when risk is $r_1 > r_0$.

We show in proposition 3.1 that if the costs are linear, then the minimal expected discounted costs tend to decrease with increased uncertainty, and the critical limits leading to maximal capacity (i. e., the ψ_{it}^{kK}'s) tend to increase with increased uncertainty. In the proof of this proposition we use the following lemma.

LEMMA 3.1: *If for some $t=T$, $W_{iT}^k(r_1) \leqq W_{iT}^k(r_0)$ for all $k \leqq K$, then $W_{it}^k(r_1) \leqq W_{it}^k(r_0)$ and $\psi_{it}^{kK}(r_1) \geqq \psi_{it}^{kK}(r_0)$ for all $t<T$ and for all $k<K$.*

Proof : By induction. Assuming that $W_{i,t+1}^k(r_1) \leqq W_{i,t+1}^k(r_0)$ for some $t+1 \leqq T$, it follows from the definition of the W's and from (2.3) that:

$$W_{it}^k(r_1) \equiv \int_0^\infty V_{it}^k(x, r_1)\, dF_{it}(x, r_1)$$

$$\equiv \int_0^\infty [\min_{m \geqq k} \{ c_m(x) + \beta\, \overline{W_{i,t+1}^m}(r_1) + M_{km} \}]\, dF_{it}(x, r_1)$$

$$\leqq \int_0^\infty [\min_{m \geqq k} \{ c_m(x) + \beta\, \overline{W_{i,t+1}^m}(r_0) + M_{km} \}]\, dF_{it}(x, r_1)$$

$$\leqq \int_0^\infty [\min_{m \geqq k} \{ c_m(x) + \beta\, \overline{W_{i,t+1}^m}(r_0) + M_{km} \}]\, dF_{it}(x, r_0)$$

$$\equiv W_{it}^k(r_0). \quad (3.1)$$

The first inequality stems from the induction assumption. The second inequality follows from the definition of increased uncertainty since the minimum of several linear functions is a concave function (*see* lemma A.2). Thus it follows that $W^k{}_{it}(r_0) \geqslant W_{it}^k(r_1)$ for all $t \geqslant T$.

The linearity assumption and (2.2) imply that:

$$c_k(x) = p_k + q_k x, \qquad k = 1, ,\ldots, K, \qquad (3.2)$$

where $p_k < p_{k+1}$ and $q_k > q_{k+1}$, $k = 1, \ldots, K-1$.

Since, from (2.8), the critical values ψ_{it}^{kK} satisfy:

$$c_k(\psi_{it}^{kK})+\beta\,\overline{W_{i,t+1}^k}=c_K(\psi_{it}^{kK})+\beta\,\overline{W_{i,t+1}^K}+M_{kK},$$

it follows that:

$$\psi_{it}^{kK}=(q_k-q_K)^{-1}[M_{k,K}+(p_K-p_k)+\beta(\overline{W_{i,t+1}^K}-\overline{W_{i,t+1}^k})]. \qquad (3.3)$$

Since $\overline{W_{i,t+1}^K}$ is independent of risk and since $\overline{W_{i,t+1}^k}$ tends to decrease with increased risk it follows that ψ_{it}^{kK} tends to increase with increased risk, i. e.:

$$\psi_{it}^{kK}(r_0)\leqq\psi_{it}^{kK}(r_1) \quad \text{for all } t\leqq T.$$

Q.E.D.

PROPOSITION 3.1: *For all t, $W_{it}^k(r_0)\geqq W_{it}^k(r_1)$ and $\psi_{it}^{kK}(r_0)\leqq\psi_{it}^{kK}(r_1)$.*

Proof : In lemma A.3 it is shown that the limits $W_i^k(r_0)$ and $W_i^k(r_1)$ of $\{W_{it}^k(r_0)\}$ and $\{W_{it}^k(r_1)\}$ satisfy:

$$W_i^k(r_0)\geqq W_i^k(r_1), \qquad i=1, \ldots, I. \qquad (3.4)$$

As explained above, the W_{it}^k's are obtained by recursion starting from some very large period $t=\bar{T}$. For $W_{i,\bar{T}+1}^k(r_0)$ and $W_{i,\bar{T}+1}^k(r_1)$ we use the steady-state limits $W_i^k(r_0)$ and $W_i^k(r_1)$, respectively. It then follows from lemma 3.1 that $W_{it}^k(r_0)\geqq W_{it}^k(r_1)$ for all $t\leqq\bar{T}$. Since $\bar{T}$ may be chosen arbitrarily large, this inequality holds for all t. The inequality $\psi_{it}^{kK}(r_0)\leqq\psi_{it}^{kK}(r_1)$ now follows immediately from (3.3) since $W_{it}^K(r_0)=W_{it}^K(r_1)$. This completes the proof.

We should note that with greater uncertainty one must wait for higher demands to come along before maximal capacity is reached. This, however, does not necessarily imply that the expected time required to reach the maximal capacity increases with increased uncertainty (*see* Section V for a counter example; an intuitive explanation has been given above).

We now turn to examine how the above results change when the linearity assumption is dropped. For this we note that the crucial part in the proof of lemma 3.1 (and hence of proposition 3.1) is the part where the concavity of $V_{it}^k(x)$ has been used to establish the second inequality in (3.1). Thus, it is the concavity of $V_{it}^k(x)$ rather than the linearity of the cost functions, which is the more relevant source of the above results. The same conclusions therefore hold whenever the functions $V_{it}^k(x)$ are concave. Consider for example the case where there are only two projects ([6]), suppose that with probability 1 the maximal capacity can satisfy the demand ([7]), and denote by A_1 the capacity of Project 1. Suppose also that the cost functions are linear up to capacity and then rise sharply. In this case the functions $V_{it}^k(x)$ will be concave if the functions $u_{it}^1(x)$ and $u_{it}^2(x)$ intersect below A_1, i.e., if capacity is "large enough". (See *fig.* 2.) Such a situation is not uncommon. Walters [1963], p. 51

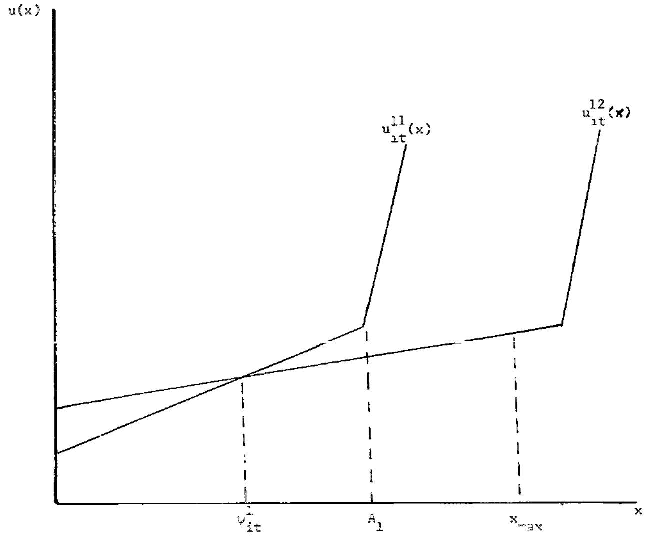

Fig. 2

mentions several time-series studies of cost functions where it has been shown that for the firms in the sample, costs were linear and measured outputs were well below capacity. This implies that over a long period of time, these firms increased capacity before full capacity has been reached. Such a behavior is consistent with the situation depicted in figure 2.

It is evident that the same results (as in proposition 3.1) hold also when the cost functions are concave. Surprisingly however, these results are not necessarily reversed when the cost functions are convex. This is due to the fact that $V_{it}^k(x)$, the minimum of several (in this case) convex functions, is neither necessarily convex nor concave. It thus follows that when costs are convex, the

([6]) This can easily be extended to the case of more than two projects.

([7]) If costs of exceeding capacity are infinite (i.e., if capacity is considered as a completely rigid notion), then the assumption that $E[c_K(x)]$ is finite implies the present assumption.

minimal expected discounted costs can either increase or decrease depending on the parameters of the problem. A heuristic explanation runs as follows: when costs are convex, increased uncertainty has two effects on the minimal expected discounted costs. On the one hand the expected value of any convex function tends to increase with increased uncertainty. However, the "sequential decision factor" tends to decrease the expected costs since the effect of higher demands can be somewhat mitigated by increasing capacity.

4. THE CASE WHERE THE ORDER OF INSTALLING PROJECTS IS NOT PREDETERMINED

There are no conceptual differences between this case and the former one. Notation however should be somewhat modified and more computations are usually required to arrive at the optimal solution. For this case we need the following definitions and notations:

N, $\{1, 2, \ldots, K\}$;

$\tilde{A}$, set of all subsets of N;

A, element of $\tilde{A}$;

$\alpha(A)$, number of elements in A;

$\bar{A}$, $N-A$, i. e., set of elements of N which are not in A;

φ, empty set;

AB, the set $\{A \vee B\}$;

M_{AB}, costs of increasing capacity from A to AB.

Suppose the costs of supplying the demand are given by $c_A(x)$ if the projects whose subscripts are in A have already been installed (we call these projects, elliptically, the projects in A). Then, in accordance with (2.3), we assume that $A \supset B$ implies $c_A(0) > c_B(0)$, $c'_A(x) < c'_B(x)$. Suppose at time t the projects in A are in operation, and the state of the economy is s_i. Then the minimal expected discounted costs of supplying the current and future demand, $V^A_{it}(x)$, conditional on the current demand x, may be obtained as follows. Under the above conditions one has $2^{\alpha(A)}$ possible actions available at time t. Namely, either install no project, or increase capacity to AB, where $B \in \bar{A}$. In the latter case one should choose the action which minimizes the expected discounted costs. It thus turns out that in this case the number of computations in each stage is $2^{\alpha(A)}$, compared with $\alpha(\bar{A})$ in the former case. In the type of projects that we consider, engineering and location consideration will usually eliminate many expansion possibilities. Thus the amount of computations involved will usually be tolerable even if K is large.

Let:

$$W_{it}^{A}=E\left[V_{it}^{A}(x)\right] \quad \text{and} \quad \overline{W}_{it}^{A}=\sum_{j=1}^{I} \pi_{ij}\, W_{jt}^{A}.$$

Then, if capacity is increased to AB ($B=\varphi$ means that no project is installed), the expected discounted costs of supplying the current and future demand are:

$$u_{it}^{AB}(x)=c_{AB}(x)+\beta\, \overline{W}_{i,t+1}^{AB}+M_{AB}, \tag{4.1}$$

that is, $u_{it}^{AB}(x)$ is the sum of the costs M_{AB} of increasing capacity from A to AB, the costs of supplying the current demand, and the discounted expected costs $\beta\, \overline{W}_{it}^{AB}$ of future operations. Hence:

$$V_{it}^{A}(x)=\min_{B \in A}\left\{u_{it}^{AB}(x)\right\}. \tag{4.2}$$

The same techniques used in section II can also be applied here to determine the W_{it}^{A}'s. Also the optimal expansion policy is of a similar structure as in the case where the order of installing the projects is predetermined. For any i, t and A there is a critical limit ψ_{it}^{A} such that if $x<\psi_{it}^{A}$ capacity is not increased. If $x>\psi_{it}^{A}$, some critical intervals determine which projects should be installed. In the context of a simple example, figure 3 illustrates the structure of the optimal policy. In our example $N=\{1,2,3,4\}$, $A=\{1,2\}$ and the possible AB sets are: A, A_1, A_2, A_3, where $A_1=\{1,2,3\}$, $A_2=\{1,2,4\}$, $A_3=\{1,2,3,4\}$. If $0 \leqq x \leqq \psi_{it}^{A}$ capacity is not increased, if $\psi_{it}^{A}<x \leqq \psi_{it}^{A_1}$ capacity should be increased to A_1, and so on.

5. AN EXAMPLE

In our example we consider a situation in which there are two projects to sequence and we assume that there are only two states of the economy. The notation in this section corresponds to that used in section III. The parameters of the problem are given below.

Transition probabilities:

$$\pi=\begin{bmatrix}\pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22}\end{bmatrix}=\begin{bmatrix}.75 & .25 \\ .25 & .75\end{bmatrix}.$$

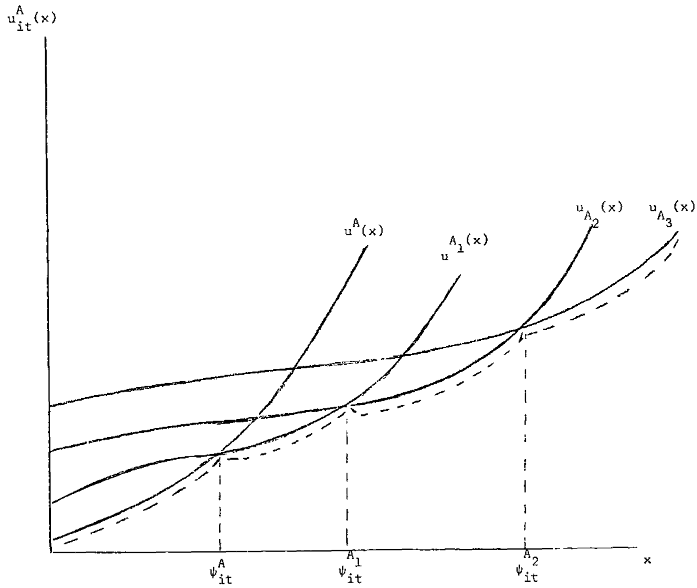

Fig. 3

Means μ_t^i.

Let:

$$\mu_t = 100(1.07)^t, \qquad 1 \leqq t \leqq 10;$$

$$\mu_t = \mu_{t-1}(1.06), \qquad 11 \leqq t \leqq 20;$$

$$\mu_t = \mu_{t-1}(1.05), \qquad 21 \leqq t \leqq 25;$$

$$\mu_t = \mu_{t-1} + 100(0,5)^{t-25}, \ t \geqq 26.$$

The μ_t^i's are given by:

$$\mu_{1t} = 0,96\,\mu_t, \qquad t = 1, 2, \ldots, \infty;$$

$$\mu_{2t} = 1,04\,\mu_t, \qquad t = 1, 2, \ldots, \infty.$$

Cost parameters.

The costs of servicing the demand x when project k is the only one in operation are given by:

$$c_k(x) = p_k + q_k x,$$

where:

$$p_1 = 1\,000, \qquad p_2 = 2\,000;$$
$$q_1 = 100, \qquad q_2 = 50.$$

The costs of servicing the demand x when both projects 1 and 2 are in operation are given by:

$$c(x) = 2\,500 + 30\,x.$$

The costs of installing project k are M_k, $k=1, 2$, where:

$$M_1 = 20\,000,$$
$$M_2 = 70\,000.$$

The discount factor β is assumed to be 0.9.

The random variables X_t^i.

For each economic state $i=1, 2$, and each $t=1, 2, \ldots$, the random variable X_t^i may assume one of the seven values $\mu_{it} - v_0\, v_{it}$, $\mu_{it} - v_0\, v_{it}/\sqrt{2}$, $\mu_{it} - v_0\, v_{it}/\sqrt{5}$, μ_{it}, $\mu_{it} + v_0\, v_{it}/\sqrt{5}$, $\mu_{it} + v_0\, v_{it}/\sqrt{2}$, $\mu_{it} + v_0\, v_{it}$, with respective probabilities: 1/32, 2/32, 5/32, 16/32, 5/32, 2/32, 1/32. v_0 is a risk parameter which will be assigned several values in the sequel, and the parameter v_{it} is given by $v_{it} = (\mu_{it})^{1/2}$. It may easily be verified that for $i=1, 2$ and $t=1, 2, \ldots, \infty$, X_t^i has a Gaussian-like shape, with mean μ_{it} and variance $(6/32)\, v_0^2\, \mu_{it}$.

From this form of the variance it follows that, the larger v_0, the larger is the dispersion of X_t^i. Thus the effect on the analysis of changes in risk will be examined by inspecting the effect on the analysis of changes in v_0.

RESULTS

The main results are summarized in tables I and II below. Table I lists the ψ_{it}'s for $v_0 = 0.75$. The optimal policy may be read from this table as follows. Suppose the state of the economy in period 5 is 1, and project 2 is in operation. Then, if the demand exceeds 113, project 1 must be introduced. Otherwise no expansion should take place.

An interesting result revealed in table I is that $\psi_{2t}^k \leq \psi_{1t}^k$. In other words, the critical values corresponding to improved economic conditions (when the demand is on the average higher) are lower than those corresponding to less favorable economic conditions. The reason for this is the following. When

TABLE I (*)

Critical Values ψ^k_{it} *for* $V_0=4.242$

t	ψ^1_{1t}	ψ^1_{2t}	ψ^2_{1t}	ψ^2_{2t}
1.	132.6	128.9	121.8	119.1
2.	125.9	123.7	117.3	116.2
3.	122.2	121.1	115.4	114.3
4.	120.3	119.4	113.7	113.3
5.	118.9	118.4	113.0	112.7
6.	118.3	118.1	112.6	112.5
7.	118.0	117.9	112.4	112.4
8.	117.8	117.8	112.4	112.4
9.	117.8	117.8	112.4	112.4

(*) Since the ψ^k_{it} converge, $\psi^k_{it}=\psi^k_{i9}$ for $t \geq 10$.

economic conditions are favorable, it follows from the Markov process generating the economic conditions that favorable economic conditions are expected to prevail in the near future. Hence it is advantageous to increase capacity in order to meet these high future demands, even if current demand is low.

In table II we present the expected costs, W^k_{it}, $k=1, 2$, $i=1, 2$, for $t=1$ and four values of v_0. We also present the following statistics:

$E(T^{*k}_i)$: the expected time elapsing from the origin until full capacity is reached, if the project introduced at the origin is k and the state of the economy at that time is s_i;

$E(T^*_i)$: the expected time elapsing from the origin until full capacity is reached, if the first project introduced is determined optimally and the state of the economy at the origin is s_i;

ξ_i: the probability that project 1 is the first one introduced, given that the economic state at the origin is s_i.

It follows from table II that $E(T^{*k}_i)$ is not a monotone function of v_0. This confirms our statement in section I that the expected time elapsing until the maximal capacity is reached does not necessarily increase with greater uncertainty. In our example it turns out, owing to the bounded range of the X^i_t's, that it is always optimal to introduce project 1 first, and to introduce project 2 a few years later. The reason for this is that the installation costs of project 2 are high. Thus the benefits derived from postponing the introduction of project 2 are large relative to $(W^1_{i1}-W^2_{i1})$—the benefits derived by introducing it first.

TABLE II

Parameters W_{i1}^{k}, $E(T_i^{*k})$, $E(T_i^{*})$, ξ_i for Selected Values of v_0

v_0	1.414	2.828	4.242	5.656
$W_{1,1}^{1}$.	15341	15316	15285	15250
$W_{2,1}^{1}$.	15474	15444	15409	15372
$W_{1,1}^{2}$.	10471	10464	10456	10446
$W_{2,1}^{2}$.	10538	10530	10522	10513
$E(T_1^{*1})$.	3.14	3.18	3.08	3.09
$E(T_2^{*1})$.	2.41	2.94	2.63	2.64
$E(T_1^{*2})$.	2.68	2.60	2.36	2.56
$E(T_2^{*2})$.	2.06	2.08	2.12	2.19
$E(T_1^{*})$.	3.14	3.18	3.08	3.09
$E(T_2^{*})$.	2.41	2.94	2.63	2.64
ξ_1.	1	1	1	1
ξ_2.	1	1	1	1

REFERENCES

R. BELLMAN, *Dynamic Programming*, Princeton, Princeton University Press, 1957.

R. BELLMAN, *Introduction to Matrix Analysis*, New York, McGraw Hill and Co., 2nd edition, 1970.

W. BUTCHER *et al.*, *Dynamic Programming for the Optimal Sequencing of Water Supply Projects*, Water Resources Research, Vol. 5, 1969, pp. 1196-1204.

D. ERLENKOTTER, *Sequencing Expansion Projects*, Operations Research, Vol. 21, 1973, pp. 542-553.

D. ERLENKOTTER and R. TRIPPI, *Optimal Investment Scheduling with Price-Sensitive Demand*, Management Science, Vol. 23, 1976, pp. 1-11.

R. GIGLIO, *Stochastic Capacity Models*, Management Science, Vol. 17, 1970, pp. 174-184.

A. MANNE, *Capacity, Expansion and Probabilistic Growth*, Econometrica, Vol. 29, 1961, pp. 632-649.

A. MANNE (ed.), *Investments for Capacity Expansion: Size, Location, and Time Phasing*, Cambridge, Massachusetts, The MIT Press, 1967.

G. RAUSSER and C. WILLIS, *Investment Sequencing, Allocation, and Learning in the Design of Water Resources Systems: An Empirical Application*, Water Resources Research, Vol. 12, 1976, pp. 317-330.

S. ROSS, *Applied Probability Models with Optimization Applications*, San Francisco, Holden-Day, 1971.

M. ROTHSCHILD and J. STIGLITZ, *Increasing Risk I, A Definition*, Journal of Economic Theory, Vol. 2, 1970, pp. 225-243.

C. TAPIERO, *Capacity Expansion of a Sequence of Deteriorating Facilities Under Uncertainty-Optimality and Simulation*, Working Paper, Columbia University, 1973.

C. TAPIERO, *Capacity Expansion of a Deteriorating Facility Under Uncertainty*, RAIRO/Operations Research, Vol. 13, 1979, pp. 55-60.

A. WALTERS, *Production and Cost Functions: An Econometric Survey*, Econometrica, Vol. 31, 1963, pp. 1-66.

APPENDIX

LEMMA A.1: *The functions W_i^k, $i=1,\ldots,n$, exist and are unique for all $k<K$.*

Proof : Here we assume that $k=K-1$. A similar proof applies also for all $k<K-1$.

Consider the transformation $h: R_+^I \to R_+^I$ (where R_+^I = the nonnegative part of the I dimensional Euclidean space) defined by:

$$h_i(v)=E[H(x)], \qquad i=1,\ldots,I, \tag{A.1}$$

where:

$$\left.\begin{aligned} H(x)&=\min\{G(x), g(x)\},\\ G(x)&=c_k(x)+\beta\sum_{j=1}^{I}\pi_{ij}v_j,\\ g(x)&=c_K(x)+M_{kK}+\beta\sum_{j=1}^{I}\pi_{ij}W_j^K. \end{aligned}\right\} \tag{A.2}$$

We show that $h(.)$ is a contraction mapping, hence from Ross [1970], it follows that there exists a unique function v^* satisfying:

$$v_i^*=E\left[\min\left\{c_k(x)+\beta\sum_{j=1}^{I}\pi_{ij}v_j^*, g(x)\right\}\right], \qquad i=1,\ldots,I. \tag{A.3}$$

This unique function v^* is by definition, W_i^k. Ross has also shown that:

$$v_i^*=\lim_{n\to\infty} h_i^n(v_0), \qquad i=1,\ldots,I, \tag{A.4}$$

where v_0 is an arbitrary element of R_+^I.

In order to show that h is a contraction mapping we need to demonstrate that:

$$\max_i |h_i(v^1)-h_i(v^2)| \leqq \rho \max_i |v_i^1-v_i^2|,$$

where $\rho<1$ is some parameter. For this we note that:

$$h_i(v^1)-h_i(v^2)\equiv E\left[\min\left\{c_k(x)+\beta\sum_{j=1}^{I}\pi_{ij}v_j^1,\ g(x)\right\}\right]$$

$$-E\left[\min\left\{c_k(x)+\beta\sum_{j=1}^{I}\pi_{ij}v_j^2,\ g(x)\right\}\right]$$

$$=\int_0^{\psi_i^1}[c_k(x)+\beta\,\bar{v}_i^1]\,dF_i(x)+\int_{\psi_i^1}^{\infty}g(x)\,dF_i(x)$$

$$-\int_0^{\psi_i^2}[c_k(x)+\beta\,\bar{v}_i^2]\,dF_i(x)-\int_{\psi_i^2}^{\infty}g(x)\,dF_i(x)$$

$$\leqq\int_0^{\psi_i^1}[c_k(x)+\beta\,\bar{v}_i^1]\,dF_i(x)+\int_{\psi_i^1}^{\infty}g(x)\,dF_i(x)$$

$$-\int_0^{\psi_i^1}[c_k(x)+\beta\,\bar{v}_i^2]\,dF_i(x)-\int_{\psi_i^1}^{\infty}g(x)\,dF_i(x)$$

$$\leqq\beta\,|\bar{v}_i^1-\bar{v}_i^2|\leqq\beta\max_i|v_i^1-v_i^2|,\qquad i=1,\ldots,I,$$

where:

$$\bar{v}=\sum_{j=1}^{I}\pi_{ij}v_j.$$

The first inequality follows from the fact that ψ_i^2 is chosen so as to minimize $h_i(v^2)$. The second inequality is obvious.

Likewise, one can show that:

$$h_i(v^2)-h_i(v^1)\leqq\beta\max_i|v_i^2-v_i^1|.$$

Hence $\max_i|h_i(v^2)-h_i(v^1)|\leqq\beta\max|v_i^2-v_i^1|$.

Q.E.D.

LEMMA A.2: *Let $g_1(x), \ldots, g_n(x)$ be n concave or linear functions, then $g(x)=\min\{g_1(x), \ldots, g_n(x)\}$ is concave.*

Proof : Let $x_\alpha = \alpha x_1 + (1-\alpha) x_2$ for some $0 \leqq \alpha \leqq 1$:

$$
\begin{aligned}
g(x_\alpha) &= \min\{g_1(x_\alpha), \ldots, g_n(x_\alpha)\} \\
&\geqq \min\{\alpha g_1(x_1) + (1-\alpha) g(x_2), \ldots, \alpha g_n(x_1) + (1-\alpha) g_n(x_2)\} \\
&\geqq \alpha \min\{g_1(x_1), \ldots, g_n(x_1)\} + (1-\alpha) \min\{g_1(x_2), \ldots, g_n(x_2)\} \\
&= \alpha g(x_1) + (1-\alpha) g(x_2).
\end{aligned}
$$

Q.E.D.

LEMMA A.3: *The limits W_i^k satisfy $W_i^k(r_1) \geqq W_i^k(r_0)$ for all $k \leqq K$, $i \leqq I$.*

Proof : As shown in lemma A.1:

$$W_i^k = \lim_{n \to \infty} h_i^n(v),$$

where the function h_i is defined in A.1. As shown in lemma A.2, the function $H(x)$ is concave. Hence it follows from the definition of increased uncertainty that for any $v \in R_+^I$ $h_i(v \mid r_1) \leqq h_i(v \mid r_0)$ and thus:

$$W_i^k(r_1) = \lim_{n \to \infty} h_i^n(v \mid r_1) \leqq \lim_{n \to \infty} h_i^n(v \mid r_0) \equiv W_i^k(r_0).$$

Q.E.D.