Discussion Paper No. 756

OPTIMAL ORDERINGS FOR PARALLEL SEARCH*

by

Tara Vishwanath
Department of Economics
Northwestern University
Evanston, IL 60208

August 1987

Abstract

This paper analyzes the optimal sequential strategy in a parallel search problem arising in many economic situations. A decision maker has a finite number n of activities or projects, each yielding an unknown reward at an uncertain time. Several m (< n) projects may be undertaken in parallel (simultaneously), and the projects may be selected sequentially in any order desired. Optimal strategies, which maximize the expected discounted utility of the rewards obtained, are in general complex to determine. We present general conditions in terms of risk and stochastic ordering of the distributions associated with projects, which result in simple optimal rules.

*This research has been partially supported by National Science Foundation Grant SES-8708325.

1.0 Introduction

Many economic problems fall into the broad category described below. A decision maker has a number (n) of activities or projects, all of which must be undertaken. Each project yields an unknown reward at an uncertain time, and is characterized by its own independent joint probability distribution of reward and yield time. Once a project is selected, the reward from it is revealed after a random time lag when it is collected. The projects/activities are selected sequentially in any order desired. Furthermore a number m ($1 \leq m < n$) of activities may be explored simultaneously, i.e., m projects can be carried out in parallel. Once the reward from a project is received, another one from the set of remaining projects is selected. Given that m projects can be undertaken simultaneously, the problem is to determine the sequential strategy (the order of project selection) that maximizes the expected present discounted utility of the stream of rewards obtained from the projects.

In principle, the above problem may be formulated as a dynamic programming problem, and the optimal strategy determined via a straightforward solution to the dynamic programming equations (using backward induction or fixed point techniques). However in most actual cases, this approach, besides shedding little economic insight, would be a combinatorially complex brute force task especially for large values of n and m.

The purpose of this paper is to demonstrate that simple and intuitive rules for the optimal sequential strategies are obtained, when the distributions of the reward and yield times of the various projects satisfy certain stochastic ordering relations. The optimal order for sequential selection of the projects can be predetermined, and may be simply stated as follows. Associated with each project is its expected (discounted) utility which depends on the reward and yield time distribution of that project. Order the projects from the one that gives the highest expected utility to the one which gives the least, and select the projects according to this predetermined order, i.e., at an instant of time, when an opportunity to select a project arises, choose from the set of remaining projects the one which has the highest expected utility. Conditions for the optimality of this rule are given in terms of the risk associated with each distribution and stochastic dominance.

The sequential strategy based on the expected utility ordering is myopic, and in general, myopic rules are not optimal in a dynamic environment such as that of the problem described above. The principal feature of this paper is the demonstration of the optimality of this ordering when several projects can be undertaken in parallel. If only one project may be undertaken at a time (the case of $m=1$), the model described above falls into the general class of bandit processes (Gittens [1979], Whittle [1980], Weitzman [1980]). The solution to bandit problems can be usually characterized by a reservation rule. Each project is assigned a reservation number or an index (a number analogous to the internal rate of return, which in this model coincides with the expected utility of the project) depending only on the project and independent of all other projects. At every decision instant the project with the highest reservation number is selected. Bandit processes have been extensively studied when only a single project can be selected at a time ($m=1$), and the solution concept of a reservation number is applicable to more general models. However, the case of parallel operation of projects has qualitatively different features and there have been relatively few studies in this direction.[1]

The problem addressed in this paper has many natural economic applications. For example, the projects can be various products that require an uncertain amount of time for their development, and the rewards may be interpreted as the estimated profits from marketing them. Depending on the context at hand, the various projects may be interpreted as mines, oil wells, resource-based technologies or investment opportunities.[2] For a further discussion of the applications and the wide scope of such problems, see Weitzman [1976] and Roberts and Weitzman [1980].

The paper is organized as follows. The assumptions and the formal description of the model for the general case of parallel operation of projects are given in section 2. The main theorem concerning the optimal policy is also stated in section 2. The proof of this theorem is presented in a later section viz., section 5. The simple case of m=1 (single project) is discussed in section 3. The parallel project case has qualitatively different features from the single project case, and this is discussed in section 3. For this case it is shown that the expected utility ordering is optimal in a more general situation when new projects become available in the future. Limitations of the optimal ordering given in the main theorem, and the role of the various assumptions for the case $m > 1$ are discussed through examples in section 4. Conclusions are given in section 6.

2.0 The Model and the Main Result

There are n projects at the beginning (i.e., time t=0), all of which must be undertaken. Project i, $1 \leq i \leq n$, yields a reward Z_i and the time taken to collect the reward (from the time the project is started) is X_i (i.e., if project begins at any time t, the reward is received at time $t + X_i$). Z_i and X_i are nonnegative real valued random variables with finite expectations and having the joint probability distribution denoted by $F_i(x,z) \equiv Pr(X_i \leq x, Z_i \leq z)$. It is assumed that F_i and F_j are independent for all $i \neq j$. Let $H_i(x) \equiv Pr(X_i \leq x)$ and $G_i(z) \equiv Pr(Z_i \leq z)$ respectively denote the (marginal) distributions of X_i and Z_i, for $1 \leq i \leq n$.

Let $m(\tau)$ denote the number of projects that can be simultaneously carried out (in parallel) at time $\tau \varepsilon [0,\infty)$. We shall assume that $m(\tau)$ is nondecreasing in τ. Specifically it is assumed that $m(\tau)$ has unit upward jumps at times $\tau_1 \leq \tau_2 \leq \tau_3 \leq \ldots \tau_m$, where $m < n$. The first project is started at time τ_1, the second additional project can be started in parallel at time τ_2, and so on. (Notice that if $\tau_1 < \tau_2$, and if the project that was begun at τ_1 is completed at some time $t < \tau_2$, another one from the set of remaining projects is selected for execution at time t, but during the time interval $[\tau_1,\tau_2)$ only one project is carried out).

The present discounted utility of receiving a reward z at time t, in its usual form, is given by $e^{-\rho t}u(z)$, where $\rho > 0$ is the discount rate and $u(\cdot)$ is a concave increasing function. However, for notational convenience we shall use an alternative form which also facilitates consideration of time varying discount rates. It is assumed that the present discounted utility of a reward z at time t is given by the function $u(t,z)$ which satisfies the properties mentioned below. (Again for convenience, we shall use the terms "increasing" and "decreasing" respectively in places of 'nonincreasing' and 'nondecreasing').

The function $U: R^+ \times R^+ \to R^+$ possesses the following properties.

(A.1) $u(t,z)$ is twice differentiable in both arguments, and is convex decreasing in t, for each z.

(A.2) $u(t,z)$ is concave increasing in z, for each t.

(A.3) $u_1(t,z) \equiv \partial u(t,z)/\partial t$ is convex decreasing in z, for each t; and concave increasing in t, for each z.

(A.4) $u_2(t,z) \equiv \partial u(t,z)/\partial z$ is decreasing in z and t;
$u_{22}(t,z) = \partial^2 u(t,z)/\partial z^2$ is increasing in t, for each z; and
$u_{11}(t,z) = \partial^2 u(t,z)/\partial t^2$ is decreasing in t, for each z.

The above assumptions conform to the usual notions of time discounting the utility, and it is easily verified that the special case $u(t,z) = e^{-\rho t}\,\bar{u}(z)$, where $\rho > 0$ and $\bar{u}(\cdot)$ is concave increasing, satisfies the above conditions.

Two kinds of strategies (for carrying out the projects) may be distinguished depending on whether or not there is an option to pull out of a project before its completion and switch to another project. If there is such an option, we shall call the class of policies preemptive; if not, policies are called nonpreemptive. In the following we shall be concerned only with nonpreemptive strategies.

A strategy, thus, is the order in which all the projects are carried out. Without confusion, we let $\pi = (i_1,i_2,\ldots,i_n)$ denote both a <u>permuted listing</u> of $(1,2,\ldots,n)$, and the <u>order</u> which selects the projects in the order $i_1,i_2,\ldots,i_n$ for execution.

Let τ denote the vector of times $(\tau_1,\ldots,\tau_n)$. Let $U(\tau,\pi)$ denote the total expected (discounted) utility from the strategy π.

$$U(\tau,\pi) = E\Big[\sum_{k=1}^{n} u(t_k(\pi),\ Z_k)\Big],$$

where $t_k(\pi)$ is the time at which the reward Z_k from project k is obtained, adopting strategy π.

Our primary interest is in characterizing the (nonpreemptive) optimal policy π^*, the order which achieves the maximum total expected utility i.e.,

$$U(\tau,\pi^*) = \max_{\pi} U(\tau,\pi). \tag{1}$$

Let $\pi = (i_1,i_2,\ldots,i_n)$. Defining $\pi_1 = (i_3,i_4,\ldots,i_n)$, the dynamic programming equation for the above problem may be written as follows.

$$\begin{aligned} U(\tau,\pi^*) = \max_{\pi} E\Big[& I(X_1 < \tau_2 - \tau_1)\{u(\tau_1 + X_{i_1}, Z_{i_1}) + u(\tau_1 + X_{i_1} + X_{i_2}, Z_{i_2}) \\ & + U(\tau_1 + X_{i_1} + X_{i_2},\ \tau_2 \ldots \tau_m;\ \pi_1)\} \\ & + I(X_{i_1} > \tau_2 - \tau_1)\ \{u(\tau_1 + X_{i_1},\ Z_{i_1}) + u(\tau_2 + X_{i_2}, Z_{i_2}) \\ & + U(\tau_1 + X_{i_1}, \tau_2 + X_{i_2},\ \tau_3 \ldots \tau_m; \pi_1)\}\Big]. \end{aligned} \tag{2}$$

In the above equation $I(\cdot)$ denotes the indicator function. If the first project (which is i_1 in the order π) is completed prior to time τ_2, then the second project (i_2) is begun at time $\tau_1 + X_{i_1}$; if not, it is started at time τ_2. In the former event, for the list of projects in π_1, the times at which different projects are started is specified by the vector $(\tau_1 + X_{i_1} + X_{i_2}, \tau_2,\ldots,\tau_n)$. A similar observation holds for the other event also.

As mentioned earlier, equation (2) may be solved, in principle, by backward induction (note that $U(\tau,\phi) \equiv 0$, where ϕ is the null list containing no projects). But, this is likely to be a combinatorial task of unwielding proportions unless n is small.

In what follows, we restrict our attention to a class of problems by imposing conditions on the distributions associated with the projects, and show that, under these conditions, simple rules for ordering the projects are optimal. In stating the assumptions the notions of mean preserving spread (see Rothschild and Stiglitz [1970]) and stochastic dominance are used.

The distribution of rewards and yield times of the projects are assumed to satisfy the following conditions.

(B.1) For each project k, $1 \leq k \leq n$, given $X_k = x$, the conditional distribution of rewards, denoted by $F_k(z|x)$, is either

(a) stochastically decreasing in x,

or has

(b) a spread which is increasing in x (with the mean preserved).[3]

(B.2) For each project k, $2 \leq k \leq n$, the conditional distribution $F_k(z|x)$ is either

(a) stochastically smaller than $F_{k-1}(z|x)$,

or is

(b) a mean preserving spread of $F_{k-1}(z|x)$.

(B.3) For each project k, $1 \leq k \leq n-1$, yield time X_k is stochastically smaller than X_{k+1}.[4]

The projects can have different spreads (with the same mean) in the yield times. However, in this case, an additional stronger assumption is needed on the conditional distribution of rewards.

(B.3)' For each project k, $1 \leq k \leq n-1$ the yield time X_k is a mean preserving spread of the distribution $F_{k+1}(x|z)$. In addition

$E[u(x,Z_k(x)]$ is convex decreasing in x, for all k, where $Z_k(x)$ is the random variable with distribution $F_k(z|x)$, and $u(\cdot)$ is as defined earlier (in assumption A).

Condition (B.1) implies that, for all projects, the longer it takes to receive the reward, the less likely it will be large, or that the risk involved in the reward is greater. (Note that B.1 is satisfied if X_k and Z_k are independent i.e., $F_k(z|x)$ is the same for all x). Conditions (B.2), (B.3), and (B.3)' order the projects according to their distributions. Considering projects i and j, where $i < j$, it is assumed that the former yields a stochastically greater or a less risky reward (conditional on the same yield time), and also that its yield time distribution is stochastically smaller or has a greater spread. It may be noted that if Z_k and X_k are independent, then the additional assumption in (B.3)' is always satisfied.

Before we state the main theorem, the following observation is in order. We shall refer to the assumptions (A.1) - (A.4) as the set A of assumptions, and the assumptions (B.1), (B.2), and (B.3) or (B.3)' as the set of B of assumptions.

Proposition 1: Given sets A and B of assumptions,

$$E[u(X_k,Z_k)] \geq E[u(X_{k+1},Z_{k+1})], \qquad \text{for } 1 \leq k \leq n-1.$$

Proof: Let $Z_k(x)$ denote the random reward from project k conditional on $X_k=x$. Define,

$$g_k(x) \equiv E[u(x,Z_k(x))], \qquad 1 \leq k \leq n.$$

By virtue of assumptions (A.1), (A.2) and (B.1), it follows that $g_k(x)$ is decreasing in x. Moreover,

$$g_k(x) \geq g_{k+1}(x), \quad \text{for all } x, \tag{3}$$

which follows from conditions (B.2) and (A.2). Hence,

$$E[u(X_k,Z_k)] = E[g_k(X_k)]$$

$$\geq E\ [g_{k+1}(X_k)] \tag{4}$$

$$\geq E[g_{k+1}(X_{k+1})] = E[u(X_{k+1},Z_{k+1})]. \tag{5}$$

The first inequality follows from (3). The second inequality is due to assumption (B.3). To prove the proposition under assumption (B.3)', note that $g_k(x)$ is convex decreasing in x and that (5) follows from (4) due to this convexity property.

As is evident from the proposition, the sets of assumptions A and B order the projects in terms of their expected utilities with project 1 having the highest expected utility and project n having the least. Under the stated conditions, we shall demonstrate that the optimal policy π^*, which achieves the maximum total discounted expected utility in (1), is the predetermined order $\pi^* = (1,2,3,\ldots,n)$. That is, the projects are ordered according to their expected utilities, and whenever an opportunity to start a project arises, the project with the highest expected utility is chosen from the set of remaining projects.

<u>Theorem</u>: For the model described above, with the set of assumptions A and B, the optimal sequential order is $\pi^* = (1,2,\ldots,n)$.

The proof of the theorem is deferred to section 5. In the next section, the simpler case of m=1 is discussed.

3.0 Case m=1: Single Project Selection

In this section, we discuss the special case of the problem discussed in the previous section, when only one project can be carried out at a time. This case has qualitatively different features from the case of parallel operation of projects, and the purpose of this section is to point out these differences. (Note that this special case is obtained from the model described in section 2, by letting $\tau_j = \infty$, for $j \geq 2$).

Firstly, the proof of the theorem for this case can be obtained by a simple interchange argument. This is demonstrated below.

To keep the exposition simple, we shall assume, in what follows, that Z_k and X_k are independent for all k, $1 \leq k \leq n$, and also that $u(t,z)$ is given by $e^{-\rho t}\, u(z)$. To prove the theorem for m=1, let $\pi \neq \pi^*$ be an optimal policy.

Then the prescription of π implies that there exists a time $T \geq \tau_1$ when a project j is selected, and immediately upon collecting the reward from this project, another project i is started, with $i < j$, i.e., project i has a higher expected utility than j. Consider another policy $\tilde{\pi}$ which is identically equal to π^* up to time T, but selects i at time T, and upon completion of i, selects project j, and thereafter $\tilde{\pi}$ is the same as π. The theorem is proved by contradiction if the total expected utility (denoted $U_{\tilde{\pi}}$) following policy $\tilde{\pi}$ is at least as large as the total expected utility (U_π) following policy π.

It is easily seen that (letting $e^{-\rho} = \beta$, $0 < \beta < 1$) U_π can be written as,

$$U_\pi = V_1 + E[\beta^{T+X_j} u(Z_j) + \beta^{T+X_j+X_i} u(Z_i)] + V_2,$$

where V_1 is the expected discounted utility obtained up to time T_1, and V_2 is the expected discounted utility obtained after completing projects j and i. Now, because of the construction of $\tilde{\pi}$,

$$U_{\tilde{\pi}} = V_1 + E[\beta^{T+X_i} u(Z_i) + \beta^{T+X_i+X_j} u(Z_j)] + V_2.$$

It follows that

$$U_{\tilde{\pi}} - U_\pi = E[\beta^{T+X_i} - \beta^{T+X_i+X_j}]\, E[u(Z_i)] - E[\beta^{T+X_j} - \beta^{T+X_j+X_i}]\, E[u(Z_j)]$$

$$\geq \beta^T(E[\beta^{X_i}] - E[\beta^{X_j}])\, E[u(Z_j)]$$

$$\geq 0.$$

The first inequality is a consequence of assumptions (B.2) and (A.2). The second inequality follows under both assumptions (B.3) and (B.3)'. The theorem is thus proved for the special case.

As is evident from the above discussion, the feature which makes the interchange argument simple, in the case of m=1, is that from a time the project is selected until its completion, the states of the other projects are unaltered. This is not the situation in the case of parallel project operation, as the states of the other projects being simultaneously carried out, are indeed changing. This aspect of the case of parallel projects is precisely what makes the proof of optimality more involved.

A second distinct feature of the one-project-at-a-time case is that the highest expected utility ordering is optimal even when new projects become available as time progresses, i.e., at any time when a project must be selected, the project with the highest expected utility among the set of currently available projects is chosen. It should be clear that the same interchange argument (presented above) is also valid (interpret U_π as the total expected discounted utility following policy π given that new projects arrive in the future) even when new projects come in at future dates. For the case of parallel selection, however, the myopic ordering implied by the theorem may not be optimal when there is arrival of new projects. This will be discussed in the next section.

4.0 Examples

The main theorem provides an extremely simple rule for ordering the projects, in situations where the assumptions of the model are satisfied. In this section, some examples are presented to show that the highest expected utility ordering of the projects is not, in general, optimal when some of the assumptions are violated.

For example, the policy π^* may not be optimal when the number of projects that can be carried out is decreasing over time (i.e., $m(\tau)$ is decreasing in τ) or there are constraints such as due dates for the completion of projects, as is illustrated by the following example.

Example 1: Consider three projects all of which have deterministic yield times, and all yielding the same unit reward, i.e., $u(Z_i) = 1$ for $i=1,2,3$.

Suppose $X_1 = 1$, $X_2 = 2$ and $X_3 = 3$. Further, assume that there are two machines (each machine can undertake one project) available at time t=0. Say, there is a constraint that both machines must be returned at t=4 (the machines are available for four time units), and in addition, one of the machines must be returned after completing the first project.

It is clear that if the highest utility rule π^* is followed, only the projects 1 and 2 can be completed (as one of the machines is returned at t=1, and there is not enough time for the other machine to start project 3 at time t=2 and complete it). If U_{π^*} denotes the total discounted utility following policy π^*, then (letting β denote the one period discount factor)

$$U_{\pi^*} = \beta^1 + \beta^2 .$$

Consider another policy $\pi = (3,2,1)$ which is the reverse order of π^*, which schedules projects 3 and 2 at t=0. Then

$$U_{\pi} = \beta^2 + \beta^3 + \beta^4 .$$

It is clear that for values of β close to unity (for example $\beta \geqslant .8$), $U_{\pi} > U_{\pi^*}$. This demonstrates that, in general, π^* is not optimal if the number of projects that can be carried out in parallel is decreasing over time.

Another assumption that has been made in the model described in section 2 is that once a project is selected, it is carried out until its completion. There is no option of pulling out or backing out from a project in the middle and starting another project. π^* is optimal when there is no option to back out from a project. If there is an option to stop a project before its completion and switch to another project, then, in general, π^* is not optimal, as illustrated by the following example.

Example 2: Suppose there are two projects X_1 and X_2. X_1 equals 1 with probability 2/3, and equals 4 with probability 1/3, and $X_2 = 2$. Suppose that $u(Z_i) = 1$ for i=1,2 and m=1. Since X_1 is a mean preserving spread of X_2, by virtue of assumption (B.3) it follows that $\pi^* = (1,2)$ and

$$U_{\pi*} = \frac{2}{3} (\beta^1 + \beta^3) + \frac{1}{3} (\beta^4 + \beta^6).$$

Assuming that the option to pull out exists, consider the following policy π: Select project 1 first. If it is learned that $X_1 \neq 1$ (at the end of period 2) switch to project 2 and then come back to project 1 at t=3. Then

$$U_\pi = \frac{2}{3} (\beta^1 + \beta^3) + \frac{1}{3} (\beta^3 + \beta^7).$$

It is easily verified that $U_\pi > U_{\pi*}$ for all β, $0 < \beta < 1$, thus demonstrating that among the class of policies which permit preemption (or backing out), π^* may not be optimal in general. Stronger assumptions on the distributions associated with the project are needed for π^* to be optimal within the class of preemptive policies.

In the case of parallel selection of projects, the highest expected utility ordering may not be optimal if new projects arrive in the future, as illustrated by the example below.

Example 3: Suppose at t=0, there are three projects: $X_1 = 1$, $X_2 = 2$, $X_3 = 3$. Let $u(Z_i) = 1$ for i=1,2,3. Further, suppose that two projects arrive at t=3, each carrying a reward z and having a yield time equal to one. Let $m(\tau) = 2$ for all $\tau \geq 0$.

If policy π^* is followed (which schedules projects 1 and 2 at t=0, and project 3 at t=3, and the other at t=4), we get

$$U_{\pi^*} = \beta^1 + \beta^2 + \beta^4 + z\beta^4 + z\beta^5.$$

Consider another policy π which chooses projects 2 and 3 at t=0, starts project 1 at t=2 and completes both new projects at t=4. We get

$$U_\pi = \beta^2 + 2\beta^3 + 2z\beta^4.$$

It is easily verified that $U_\pi > U_{\pi^*}$ for all $z > (1 + \beta^3 - 2\beta^2)/\beta^3(1-\beta)$, thus showing non-optimality of π^* when new projects may arrive, in the parallel operation case.

<u>Example 4</u>: Consider 3 projects. Let $X_1 = 2$, $X_2 = 3$, and $X_3 = 1$ or 6 with equal probability. Let $u(Z_i) = 1$ for i=1,2,3. Notice that project 3 has a higher mean and a higher spread than both projects 1 and 2. Neither assumption (B.3) or (B.3)' is satisfied. Suppose two projects can be chosen at a time, i.e., $m(\tau) = 2$ for all τ. If U_i denotes the expected discounted utility of project i=1,2,3, the for $\beta = 0.9$

$$U_3 = \frac{1}{2}(\beta^1 + \beta^6) = 0.715,$$

$$U_1 = \beta^2 = 0.81 \quad \text{and}$$

$$U_2 = \beta^3 = 0.729.$$

The total discounted utility from the highest expected utility ordering $\pi^* = (1,2,3)$ is,

$$U_{\pi^*} = \beta^2 + \beta^3 + \frac{1}{2}\beta^2(\beta^1 + \beta^6) = 2.114$$

However, consider the order $\pi = (3,1,2)$. Then (for $\beta = 0.9$)

$$U_\pi = \frac{1}{2}(\beta^1 + \beta^4) + \beta^2 + \frac{1}{2}\beta^5 + \frac{1}{2}\beta^6 = 2.149$$

$$> U_{\pi^*}.$$

The last example shows that if the stochastic ordering assumptions on yield times are not satisfied, then π^* prescribed in the theorem may not be optimal. Similar examples may be easily constructed to show the nonoptimality of π^*, in general, if assumption (B.1) or (B.2) is not satisfied.[5]

5.0 Proof of Optimality

In this section the proof of the theorem is presented. The proof is approached through several lemmas.[6]

The first lemma is similar to proposition 1.

Lemma 1: Let $g(x,z)$ be a real valued function, concave increasing in x, and convex decreasing in z. If the random variables (X_i, Z_i), for $1 \leq i \leq n$ satisfy the set (B) of assumptions, then

$$E[g(X_k, Z_k)] \geq E[g(X_{k-1}, Z_{k-1})], \quad 2 \leq k \leq n.$$

Proof: Let $Z_k(x)$ denote the random reward from project k, conditional on $X_k = x$. Define

$$h_k(x) = E[g(x, Z_k(x))].$$

By virtue of assumption (B.1), and due to the assumed properties of $g(\cdot)$, it follows that $h_k(x)$ is increasing in x. Moreover, due to assumption (B.2), $h_k(x) \geq h_{k-1}(x)$, for $2 \leq k \leq n$. Hence we have,

$$\begin{aligned}
E[g(X_k, Z_k)] &= E[h_k(X_k)] \\
&\geq E[h_{k-1}(X_k)] \\
&\geq E[h_{k-1}(X_{k-1})] = E[g(X_{k-1}, Z_{k-1})]
\end{aligned}$$

The second inequality is obtained from (B.3) or (B.3)' due to the assumed dependence of $g(x,z)$ on x. This proves lemma 1.

Lemma 2: Suppose Y_1 and Y_2 are independent real valued random variables.

a. If $h(x_1, x_2)$ is real valued function increasing in x_1 and decreasing in x_2, and if Y_1 is stochastically smaller than Y_2 then

$$E[h(Y_1, Y_2)] \leq E[h(Y_2, Y_1)] \qquad (6)$$

b. If $h(x_1, x_2)$ is a real valued function which is concave in x_1 and convex in x_2, and if Y_1 is a mean preserving spread of Y_2, then inequality (6) holds.

Proof: Define $g_i(x_2) = E[h(Y_i,x_2)]$ for $i=1,2$. Under the hypothesis of part a), it is obvious that $g_i(x_2)$ is decreasing in x_2 for $i=1,2$. Moreover, $g_1(x_2) \leq g_2(x_2)$ for all x_2. Therefore,

$$E[h(Y_1,Y_2)] = E[g_1(Y_2)] \leq E[g_2(Y_2)]$$

$$\leq E[g_2(Y_1)] = E[h(Y_2,Y_1)],$$

which proves part (a) of the lemma. The argument for (b) is similar, after observing that $g_i(x_2)$ is convex in x_2 for $i=1,2$.

The next lemma derives some properties of the total discounted expected utility when the optimal policy π^* is followed. Specifically, the lemma characterizes the rate of change of the total expected discounted utility with respect to the times τ_i, $1 \leq i \leq n$. Note that τ_i is the time at which the i^{th} parallel project can be started. For ease of exposition only, we shall refer to τ_i as the instant at which the i^{th} machine is available. Each machine can undertake one project at a time. The first machine arrives at time τ_1, when the first project is begun; the second machine arrives at time τ_2, when the second project is started and so on. Define

$$U_i(\tau,\pi) \equiv \frac{dU(\tau,\pi)}{d\tau_i}, \qquad 1 \leq i \leq m. \tag{7.a}$$

To avoid any ambiguity in the above definition, when more than one machine is available at τ_i, we shall define $U_i(\tau,\pi)$ to be the right hand derivative,[7] i.e., for $\varepsilon > 0$

$$U_i(\tau,\pi) = \lim_{\varepsilon \to 0} [U(\tau_1,\ldots,\tau_{i-1},\tau_i+\varepsilon,\ \tau_{i+1},\ldots,\tau_m;\pi) - U(\tau,\pi)]/\varepsilon. \quad (7.b)$$

Similarly let $U_{ii}(\tau,\pi)$ denote the right hand second derivative,

$$U_{ii}(\tau,\pi) \equiv d^2U(\tau,\pi)/d\tau_i^2; \quad (8)$$

Lemma 3: Suppose there are n projects which satisfy the set B of assumptions, and the strategy $\pi^* = (1,2,\ldots,n)$ is followed. Let τ be the vector of instants $(\tau_1,\tau_2,\ldots,\tau_m)$ where τ_i denotes the instant at which the i^{th} parallel project is started. Then the following properties hold.

a. For $j \neq i$, and $n > 2$,

$U_i(\tau,\pi^*)$ is increasing in τ_i and decreasing in τ_j.

b. For $j \neq i$, and $n > 1$,

$U_{ii}(\tau,\pi^*)$ is decreasing in τ_i and increasing in τ_j.

c. Suppose $\pi_2^* = (2,3,\ldots,n)$, omitting some $k > 2$. Then for $n > 2$

$E[U_1(\tau_1+X_1,\tau_2+X_k,\ldots,\tau_m;\pi_2^*) - U_1(\tau_1+X_k,\tau_2+X_1,\ldots,\tau_m;\pi_2^*)] < 0.$

Proof: The proof is by induction on n. Parts (a) and (b) are trivial for n=1. For this case $U(\tau,\pi^*) = E[u(\tau_1+X_1,Z_1)]$. Therefore $U_1(\tau,\pi) = E[u_1(\tau_1+X_1,Z_1)]$ is increasing in τ_1 (due to assumption A.1) and non-decreasing in τ_j, $j \neq 1$. Part (b) is proved for n=1, similarly by virtue of assumption A.3. Part (c) is obvious for n=2.

Suppose the lemma is true when there are fewer than n projects. We shall show it holds for n projects. To show that $U_i(\tau,\pi^*)$ is increasing in τ_i (in part (a)) let $\tau_2 < \tau_3 < \ldots < \tau_m$. Let i=1, without loss of generality.

We shall outline the behavior of $U_1(\tau,\pi^*)$ as τ_1 increases, considering three separate cases viz: (i) $\tau_1 < \tau_2$, (ii) $\tau_1 > \tau_2$, and (iii) $\tau_1 = \tau_2$. For $\tau_1 < \tau_2$, (defining $\pi_1^* = (2,3, \ldots n)$)

$$U(\tau,\pi^*) = E[u(\tau_1 + X_1, Z_1) + U(\tau_1 + X_1, \tau_2, \ldots, \tau_m; \pi_1^*)].$$

Hence, we have

$$U_1(\tau,\pi^*) = E[u_1(\tau_1 + X_1, Z_1) + U_1(\tau_1 + X_1, \tau_2, \ldots, \tau_m; \pi_1^*)]. \tag{9}$$

It follows from assumption A.1 and the induction hypothesis (note π_1^* has fewer than n projects) that the expression on the right side of (9) over which the expectation is taken, is increasing in τ_1. Thus we have shown that $U_1(\tau,\pi^*)$ is increasing in τ_1, within the region $\tau_1 < \tau_2$.

Considering the region $\tau_1 > \tau_2$, observe that project 1 is started at time τ_2. It follows that

$$U(\tau,\pi^*) = E[u(\tau_2 + X_1, Z_1) + U(\tau_1, \tau_2 + X_1, \ldots, \tau_m; \pi_1^*)]$$

and, therefore

$$U_1(\tau,\pi^*) = E[U_1(\tau_1, \tau_2 + X_1, \ldots, \tau_m; \pi_1^*)]$$

Again from the induction hypothesis, it is clear that $U_1(\tau,\pi^*)$ is increasing in τ_1, in the region $\tau_1 > \tau_2$.

It remains to consider the change in $U_1(\tau,\pi^*)$ when τ_1 takes the value τ_2. Suppose $\tau_2 = \tau_3 = \ldots \tau_k < \tau_{k+1} \leq \ldots \tau_m$. Let $\pi_2^*=(2,3,\ldots,n)$, omitting project k. First consider case $k < n$.

Let $\tau_1 = \tau_1 + \varepsilon$, where $\varepsilon > 0$ is small. Then, projects 1 through k-1 are started at τ_2, and project k is started at $\tau_2 + \varepsilon$. Letting $\varepsilon \to 0$, the right hand limit

$$U_1(\tau_2,\tau_2,\ldots,\tau_m;\pi^*) = \lim_{\varepsilon\to 0} U_1(\tau_2+\varepsilon,\tau_2,\ldots,\tau_m;\pi^*) \tag{10}$$

$$= E[u_1(\tau_2+X_k,Z_k) + U_1(\tau_2+X_k,\tau_2+X_1,\tau_3,\ldots,\tau_m;\pi_2^*)].$$

Similarly, if $\tau_1 = \tau_2 - \varepsilon$, where $\varepsilon > 0$, then project 1 is started at $\tau_2 - \varepsilon$, whereas projects 2 through k are started at τ_2. Hence, the left hand limit is

$$U_1(\tau_1^-,\tau_2,\tau_3,\ldots,\tau_m;\pi^*) = E[u(\tau_2+X_1,Z_1) + U_1(\tau_2+X_k,\tau_2+X_1,\tau_3,\ldots,\tau_m^*;\pi_2)] \tag{11}$$

Thus, for the case $k < n$, the change in $U_1(\tau,\pi^*)$ when τ_1 takes the value τ_2 is the difference between the values in equations (10) and (11), which equals

$$E[u_1(\tau_2+X_k,Z_k) - u_1(\tau_2+X_1,Z_1)] \geq 0 \tag{12}$$

The nonnegativity of the above expression is a consequence of assumption A.3, and the set B of assumptions. (Note lemma 1 can be directly applied).

For the case of $k \geq n$, the change in $U_1(\tau,\pi^*)$ when τ_1 takes the value τ_2 equals

$$E[-u_1(\tau_2+X_1,Z_1)] \geq 0, \tag{13}$$

since $u_1(\cdot) \leq 0$ by virtue of assumption A.1. The arguments presented above, establish that the change in $U_1(\tau,\pi^*)$ is positive at $\tau_1 = \tau_2$. This completes the induction for part (a). (Arguments are similar to show that $U_1(\tau,\pi^*)$ is decreasing in τ_j, for $j \neq 1$, and are omitted).

It is easy to see that part (b) can also be established along the same lines as the proof for (a) given above. (In this case inequalities (12) and (13) will be reversed).

The inductive step for (c) is directly established by letting $h(x_1,x_2) = U_1(\tau_1 + x_1, \tau_2 + x_2, \tau_3, \ldots, \tau_m; \pi^*)$ and by using part (a) of the lemma, the assumption (B.3) or (B.3)', and lemma 2. This completes the proof of lemma 3.

The next lemma states that the marginal expected utility $U_1(\tau,\pi^*)$, when following optimal policy π^* is greater if project 1 were started later and then the other projects are carried out in the optimal order.

Lemma 4: Suppose the optimal order $\pi^* = (1,2,\ldots,n)$ is followed. Let $\pi_1^* = (2,3,\ldots,n)$. Then for $\tau_1 \leq \tau_2 \leq, \ldots, \leq \tau_m$,

$$\text{a.} \quad U_1(\tau,\pi^*) \leq E[U_1(\tau_1, \tau_2 + X_1, \ldots, \tau_m; \pi_1^*)], \tag{14.a}$$

and

$$\text{b.} \quad U_{11}(\tau,\pi^*) \geq E[U_{11}(\tau_1, \tau_2 + X_1, \ldots, \tau_m; \pi_1^*)]. \tag{14.b}$$

Proof: The proof is by induction on n. The lemma is trivially true for n=1, since the left side of (14.a) is negative and the right side is zero. Suppose the lemma is true when there are fewer than n projects. Let $\pi_2^* = (3,4,\ldots,n)$. For $\tau_1 \leq \tau_2$,

$$U_1(\tau,\pi) = E[u_1(\tau_1 + X_1, Z_1) + U_1(\tau_1 + X_1, \tau_2, \ldots, \tau_m; \pi_1^*)]$$

$$\leq E[u_1(\tau_1 + X_1, Z_1) + U_1(\tau_1 + X_1, \tau_2 + X_2, \ldots, \tau_m; \pi_2^*)]$$

$$\leq E[u_1(\tau_1 + X_2, Z_2) + U_1(\tau_1 + X_2, \tau_2 + X_1, \ldots, \tau_m; \pi_2^*)]$$

$$= E[U_1(\tau_1, \tau_2 + X_2, \ldots, \tau_m; \pi_1^*)]$$

The first inequality follows from the induction hypothesis. The second inequality follows from part (c) of lemma 3, and the assumptions (A.2), (B.2), and (B.3) or (B.3)'. This completes the inductive proof for (a). The proof of (b) is similar, and is omitted.

The proof of the theorem is now established, using the above lemmas.

<u>Proof of the Theorem</u>: The proof is by induction on n. The theorem is obviously true for n=1. Suppose the theorem is true when there are fewer than n projects. We shall establish the induction step, through an interchange argument and contradiction, and prove that the theorem holds for n projects.

Consider a strategy $\pi \neq \pi^*$ which selects project k(>1) at time τ_1. By the inductive hypothesis it is optimal to start project 1 next (at τ_2) and thereafter follow the rule prescribed by the theorem, i.e., the order $\pi_2^* =$ (2,3,...,n) omitting project k. Thus amongst the strategies which start k first, the best strategy is $\pi_{k,1} \equiv (k,1) + \pi_2^*$ which denotes order $(k,1,\pi_2^*)$. We shall demonstrate in what follows that the policy $\pi_{1,k}$ defined by

$$\pi_{1,k} \equiv (1,k) + \pi_2^* ,$$

is better than $\pi_{k,1}$. The proof of this claim will complete the inductive step.

Let $U(\tau,\pi;c,z)$ denote the expected total utility following strategy π, given (or conditional on) $X_k = c$ and $Z_k = z$. Define

$$\Delta(c,z) \equiv U(\tau,\pi_{1,k};c,z) - U(\tau,\pi_{k,1};c,z).$$

To prove the theorem, it is enough to establish that $\Delta(c,z)$ is concave increasing in c and convex decreasing in z. The reason for this is as follows. Let $(\bar{X}_1,\bar{Z}_1)$ be a pair of independent random variables which are identically distributed as the pair $(X_1 Z_1)$. Given Set B of assumptions, and (X_1,Z_1) assuming the above property of $\Delta(c,z)$, it follows from lemma 1 that

$$E[\Delta(X_k,Z_k)] > E[\Delta(\bar{X}_1,\bar{Z}_1)] = E[\Delta(X_1,Z_1)] = 0.$$

This establishes that $\pi_{1,k}$ is better, thus completing the induction. (Note that, in the above, the last equality to zero follows from the fact that if both projects 1 and k were identical, then the policies $\pi_{1,k}$ and $\pi_{k,1}$ will both yield the same expected reward).

We, now, proceed to establish the desired properties of $\Delta(c,z)$. It is readily seen that

$$\begin{aligned} U(\tau,\pi_{1,k};c,z) = E[&I(X_1 < \tau_2 - \tau_1)\{u(\tau_1 + X_1,Z_1) + u(\tau_1 + X_1 + c,z) \\ &+ U(\tau_1 + X_1 + c,\tau_2,\ldots,\tau_m;\pi_2^*)\} \\ &+ I(X_1 \geq \tau_2 - \tau_1)\{u(\tau_1 + X_1,Z_1) + u(\tau_2 + c,z) \\ &+ U(\tau_1 + X_1,\tau_2 + c,\ldots,\tau_m;\pi_2^*)\}]. \end{aligned} \tag{15}$$

Also, we can write

$$U(\tau,\pi_{k,1};c,z) = u(\tau_1 + c,z) + U(\tau_1 + c,\tau_2,\ldots,\tau_m;(1) + \pi_2^*), \tag{16}$$

where $(1) + \pi_2^*$ denotes the list of projects $(1,2,\ldots,n)$ omitting k, which are also carried out in that order.

From (15) and (16) it follows that

$$d\Delta(c,z)/dz = E[I(X_1 < \tau_2 - \tau_1)\{u(\tau_1 + X_1 + c,z) - u_2(\tau_1 + c,z)\}$$

$$+ I(X_1 \geq \tau_2 - \tau_1)\{u_2(\tau_2 + c,z) - u_2(\tau_1 + c,z)\}. \qquad (17)$$

$$\leq 0. \qquad (18)$$

The last step is a direct consequence of assumption (A.4) and the fact that $\tau_2 \geq \tau_1$. Differentiating (17) again with respect to z, yields (using assumption A.4) that

$$d^2\,\Delta(c,z)/dz^2 \geq 0. \qquad (19)$$

Inequalities (18) and (19) establish that $\Delta(c,z)$ is convex decreasing in z. It remains to show that $\Delta(c,z)$ is concave increasing in c. Differentiating (16) with respect to c and then making use of lemma 4, we have

$$dU(\tau,\pi_{k,1};c,z)/dc \leq E[u_1(\tau_1 + c,z) + U_1(\tau_1 + c,\ \tau_2 + X_1,\ldots,\tau_m;\ \pi_2^*)] \qquad (20)$$

Now differentiating (15) w.r.t. (c), and using (20), it follows that,

$$d\Delta(c,z)/dc \geq E[I(X_1 < \tau_2 - \tau_1)\{u_1(\tau_1 + X_1 + c,z) + U_1(\tau_1 + X_1 + c,\tau_2,\ldots,\tau_m;\pi_2^*)\},$$

$$+ I(X_1 \geq \tau_2 - \tau_1)\{u_1(\tau_2 + c,z) + U_1(\tau_2 + c,\tau_1 + X_1,\ldots,\tau_m;\pi_2^*)\}$$

$$- \{u_1(\tau_1 + c,z) + U_1(\tau_1 + c,\tau_2 + X_1,\ldots,\tau_m;\pi_2^*)\}] \tag{21}$$

$$\geq 0. \tag{22}$$

Using part (a) of lemma (3), assumption A.1 and the fact that $\tau_1 \leq \tau_2$, it is easy to verify that the expression on the right side of (21) over which the expectation is taken is positive. This establishes that $\Delta(c,z)$ is increasing in c. Similarly, differentiating (15) and (16) twice with respect to c, and making use of (part b) lemma 4, we get

$$d\Delta^2(c,z)/dc^2 \leq E[I(X_1 < \tau_2 - \tau_1)\{u_{11}(\tau_1 + X_1 + c,z) + U_{11}(\tau_1 + X_1 + c,\tau_2,\ldots,\tau_m;\pi_2^*)\}$$

$$+ I(X_1 \geq \tau_2 - \tau_1)\{u_{11}(\tau_2 + c,z) + U_{11}(\tau_2 + c,\tau_1 + X_1,\ldots,\tau_m;\pi_2^*)\}$$

$$- \{u_{11}(\tau_1 + c,z) + U_{11}(\tau_1 + c,\tau_2 + X_1,\ldots,\tau_m;\pi_2^*)\}]$$

$$\leq 0. \tag{23}$$

The last step follows from part (b) of lemma 3, the fact that $\tau_1 \leq \tau_2$, and assumption A.4 (i.e., $u_{11}(\cdot,z)$ is decreasing in its first argument). Inequalities (18), (19), (22) and (23) establish that $\Delta(c,z)$ is concave increasing in c and convex decreasing in z, thus completing the proof of the theorem.

6.0 Conclusions

This paper addresses a parallel search problem which is common to many economic situations. The problem of determining the optimal sequential strategy is often complex. In general, the optimal policies are hard to characterize and difficult to compute, as the time to arrive at the solution (through standard techniques) grows exponentially in the parameters of the problem. Thus, it is worthwhile investigating conditions under which simple and intuitive rules are optimal. A set of general conditions is presented, in this paper, for the problem of undertaking a finite number of projects (each having uncertain reward and yield time), when several of them can be explored (nonpreeemptively) in parallel, to maximize the expected discounted utility. The optimal policy is shown to be a simple predetermined order for carrying out the projects.

For the problem of parallel projects considered here, conventional thinking (based on static scenario) suggests picking (as many as are allowed) those projects having the highest expected utility, when determined in solution (i.e., computed individually). This paper may be viewed as a study of the conditions for which such a highest expected utility ordering is optimal, when decisions are made sequentially in a dynamic environment. The conditions given are general in the sense that if any of them is violated, this rule is no longer optimal, and the resulting problem becomes complex. This is demonstrated via several examples.

The case of one-project-at-a-time has been dealt with extensively in the literature. In contrast, the studies of the ubiquitous case of parallel projects is relatively few. It appears that the solution to the parallel projects problem is not amenable for characterization via simple reservation rules (as in the single project case), unless some restrictions are imposed.

While investigation of the conditions for which other simple rules are optimal, may prove useful in applications, the study of this paper could be extended in several directions. For example, finding conditions under which the highest expected utility ordering is optimal for the parallel project case even when preemptions (or pull-out options) are allowed is a topic for further investigation. Learning, when new information about a project is continually revealed during its operation, and correlated projects are other topics for further study in this context. Finally, studying this parallel project problem when the decision maker's object is to maximize the expected discounted value of the maximum of the rewards received (stopping problem) has applications in technology choice, and research and development.

References

1. Gittens, J.C. [1979]: "Bandit Processes and Dynamic Allocation Indices," Journal of the Royal Statistical Society, B, 41, No. 2. 148-177.

2. Roberts, K.W.S. and M.L. Weitzman [1981]: "Funding Criteria for Research, Development and Exploration Projects," Econometrica, 49, No. 5, 1261-1288.

3. Rothschild, M. and J. Stiglitz [1970]: "Increasing Risk I: A Definition," Journal of Economic Theory, 9, 185-202.

4. Weber, R.R., P. Varaiya and J. Walrand [1986], "Scheduling Jobs on Parallel Servers to Minimize Expected Flowtime", Journal of Applied Probability, 23.

5. Weitzman, M. [1979]: "Optimal Search for the Best Alternative," Econometrica, 47, No. 3, 641-654.

6. Weitzman, M. and K. Roberts [1980]: "On a General Approach to Search and Information Gathering," Working Paper No. 263, Department of Economics, M.I.T.

7. Whittle, P. [1980]: "Multiarmed Bandits and the Gittens Index," Journal of Royal Statistical Society, B, 42, No. 2.

Footnotes

[1]It is easy to construct examples to demonstrate that picking the projects corresponding to the n largest reservation numbers is not an optimal policy for the parallel projects case, when these reservation numbers are determined from the one-project-at-a-time case.

[2]For an application to scheduling jobs at a service station, see Weber, Varaiya, and Walrand [1980] who consider optimal scheduling on parallel servers to minimize the sum of expected waiting and service times of all jobs, when service times are ordered through first order dominance.

[3]Given the random variables Y_i, i=1,2, with respective distributions $D_i(y)$, the random variable Y_1 is stochastically smaller than Y_2 if $D_1(y) \geq D_2(y)$ for all y; Y_1 is a mean preserving spread of Y_2 if $\int_0^y [D_1(s) - D_3(s)] \geq 0$ for all y. In the former case, $E[f(Y_1)] \geq E[f(Y_2)]$ for every decreasing function $f(\cdot)$, (the inequality is reversed if f is increasing). In the latter case, this inequality holds for every convex function f.

[4]A sufficient condition for B.3 to hold, may be given in terms of the conditional distribution, $F_k(x|z)$, of yield times. Since the marginal $H_k(x) = \int F_k(x|z)\, dG_k(z)$, it follows that $H_k(x) - H_{k+1}(x) =$
$\int [F_k(x|z) - F_{k+1}(x|z)]\, dG_k(z) + \int F_{k+1}(x|z)\, [dG_k(z) - dG_{k+1}(z)]$.
This quantity is nonnegative for each x, if (i) for each z, $F_k(x|z) \geq F_{k+1}(x|z)$, (ii) for each k, $F_k(x|z)$ is increasing in z, and (iii) Z_k is stochastically greater than Z_{k+1}. In otherwords, the conditional yield times must be stochastically increasing in k (for each z), and decreasing in z (for each k).

[5]Such examples will be similar in nature to those that may be constructed to demonstrate, for example, the fact that, given two random variables, the expectation of increasing function of these may not be ordered if they are not stochastically ordered.

[6]Some of our arguments in this section are similar to those in [4].

[7]In otherwords, a project is started on machine i (the one with respect to which the derivative is taken) after assigning a project (according to π^*) on other machines which may become available at the same time. To put in another way, machine i is thought of as becoming available slightly later.