DISCUSSION PAPER NO. 198

DYNAMIC INVESTMENT STRATEGIES FOR A RISKY R AND D PROJECT

by

S. D. Deshmukh† and S. D. Chikte*

February, 1976

† Graduate School of Management, Northwestern University, Evanston, Illinois.
This research was supported by NSF Grant GS-28287.

* Systems Analysis, Department of Electrical Engineering, Polytechnic Institute of New York, Brooklyn, New York.

DYNAMIC INVESTMENT STRATEGIES FOR A RISKY R AND D PROJECT

1. INTRODUCTION

A significant feature of a research and development (R and D) project as compared to other investment activities is the considerable amount of uncertainty associated with it. The amount of time and effort required to achieve progress as well as the final outcome of the project and its value are all uncertain. The sources of these uncertainties are usually classified into (i) technological factors, that cause the effectiveness of resource allocations to be uncertain due to the intrinsically stochastic nature of the R and D process, and (ii) market conditions, that cause the value of the project outcome to be uncertain due to random changes in consumer tastes, competitors' actions, etc. (see, for example, Mansfield [17], and Marschak, Glennan and Summers [18]). In presence of these uncertainties the R and D manager is faced with the following three interrelated decisions: (1) determine whether it is at all worthwhile undertaking the project in the first place, (2) once it is undertaken, determine an allocation strategy for financing the project through time, and (3) given that the project is on its way, determine the point at which it should be discontinued, whereupon a terminal reward is obtained as a function of the final outcome. In this paper we consider these problems within the context of a firm attempting to develop an improved quality product before marketing it, so as to maximize the profit net of R and D expenditures. These problems of optimally undertaking, financing and terminating an R and D project in presence of the technological and market uncertainties are unified into the single problem of determining an optimal stochastic dynamic resource allocation strategy, which will be shown to have economically meaningful characterization.

In the area of dynamic resource allocation in an R and D project, Kamien and Schwartz [11, 12] and Lucas [15] have obtained forms of optimal expenditures as functions of time using control theory, while Aldrich and Morton [1] have obtained similar results using a dynamic programming model. The corresponding problem of optimally distributing resources among several projects has been considered by Gittins [6, 7, 8] and Laska, Meisner and Siegel [13]. In these and other related models (such as Hess [9]), the state of the project at any time is classified as either being (completely) successful or not, where "success" is precisely defined a priori and "partial success" achieved during the conduct of the project is meaningless and inconsequential. In such models, success may occur instantaneously at any point in time and the conditional probability density of the project completion time (i.e. the hazard rate) is a function of the total effort accumulated till that time, thereby reflecting the (internal technological) uncertainty regarding the total effort required for successful completion of the project. Similarly, project "failure" (premature termination) may occur due to a rival's introduction of a similar competing product, the (external market) uncertainty regarding the time of such an event being represented by the corresponding hazard rate function. Such a model may realistically represent, for example, the development of a specific chemical compound or a cure for a disease or, more frivolously, the activity of participating in a jig-saw puzzle competition, where the binary description of the project may be appropriate.

On the other hand, during the course of an R and D project, it is often meaningful and possible to assess its progress up to any point in time (see, for example, Radner and Rothschild [19], where the progress of an activity is represented by a scalar, denoting its worth). The project resource allocation may then be based on its current progress. Moreover, the level of progress achieved at which the project is to be declared successfully completed is often not pre-specified, so that the project goal itself is a decision variable. Specifically, in this paper we consider a development project aimed at improving the quality of a product relative to that of other competing products of the same generic type in the market, so that the progress of the project is represented by the relative quality achieved so far. Such a relative product quality may be perceived by the manager in terms of, for example, the potential rating of the product by a consumer research organization or the potential profit obtainable if the product developed so far was marketed. The amount of resources to be allocated at any time during the course of such an R and D project then becomes a function of the relative quality of the product developed so far. In addition to determining such a strategy for financing the project through time, the R and D manager must also select the target quality to be achieved before introducing the developed product into the market and collecting the resulting profit.

The progress of the project up to any point in time, as measured by the relative quality of the product developed by then, is a cumulative result of all the past partial successes and failures of possibly differing degrees. A partial success is due to a technological advance by our firm, which occurs at random times that depend on the development expenditures and results in an improvement in the quality of our product relative to the competitors. On the other hand, a partial failure is assumed to be due to a competitor's introduction of an improved product, occurring at random times that are beyond our firm's control and results in making our firm's product relatively inferior. Thus the relative quality of the product developed up to any point in time represents the current project status and summarizes the net effectiveness of the past resource allocations that have resulted in a series of partial successes (improvements) and failures (reductions). The optimal resource allocation strategy for financing the R and D project through time then turns out to be a function of the current quality, so that we have an adaptive feedback control strategy instead of a time expenditure plan which has to be revised every time the project status changes during its conduct. Along with the characterization of such an allocation strategy, our model also yields the optimum goal quality that the R and D manager must strive to develop. As the project progresses through time, it may be optimal to discontinue its further development and market the product, either because the goal quality has been attained, or because the currently attained quality is so inferior in comparison with the goal quality that the expected expenditure of resources required to reach the goal is greater than the associated expected reward incentive. The decision as to whether to undertake the development of the product of a given initial quality at all is then a special case of the above problem.

The mathematical model presented in the next section makes precise the problems qualitatively outlined above and is analyzed, using the semi-Markov decision theory. In section 3, we establish the structures of the optimal strategies for undertaking, financing and terminating decisions and characterize their dependence on the parameters of the problem. The final section concludes with a summary of the salient features of the model and the results.

2. THE SEMI-MARKOV DECISION MODEL

The total revenues from the R and D project completion date onwards depend upon the quality of the end product developed, in terms of its ability to satisfy the existing and latent customer preferences, and the actions of rivals, in terms of their current and future introduction of competing products (see Scherer [20] and Mansfield [17], p. 129). The terminal reward from the product development project is then the expected value of the resulting profit stream, discounted to the project completion date, and is a function of the final product quality developed relative to that of the competing products in the market on that date. This relative quality rating may be looked upon as a shift parameter in the demand function of that product (as in Dorfman and Steiner [4]), and will be denoted by $q \in Q$, where, for convenience, we take the quality space Q to be the set of all integers, a higher value of q corresponding to a better quality product in comparison with other similar products currently in the market. Each unit of improvement in the rating of the product quality relative to the competition will be assumed to be worth $\beta > 0$, as a result of the increased sales due to the introduction of the better product. Thus, greater improvements over existing competing products bring in proportionately higher rewards, reflecting the consumers' ability in discriminating among products on the basis of their relative qualities and in selecting a superior product proportionately more often at a given price. Such an assumption of constant returns to scale in carrying out the activity of improving the product quality and marketing it corresponds to the linearity assumption made in the economic activity analysis and is, in fact, automatically satisfied with $\beta = 1$ if the quality itself is measured in terms of the potential profit obtainable upon marketing the product as in [19]. Let $\underline{q}$ be the minimum quality our product must have relative to competing products available in the market in order to induce sufficient demand necessary to cover the fixed costs of production and distribution; a product of quality worse than $\underline{q}$ is not worth commercial production. Thus, if the product of relative quality q is introduced into the market, the reward collected summarizes its worth and is given by for all i, so that the average time interval between successive advances of magnitude i, which is $\frac{1}{\lambda_i(a)}$, is non-increasing in a, (conforming to the above mentioned time-cost tradeoff relationship) and that a positive effort is necessary for advancing in the competitive environment. We also assume that $\sum_{i=1}^{\infty} \lambda_i(B)$ is finite, to be denoted by $\lambda(B)$, to ensure that in a finite amount of time only an advance of a finite magnitude and only a finite number of such advances are possible for any resource allocation. Furthermore, suppose that $\sum_{i=1}^{\infty} i\lambda_i(B) < \infty$, so that the expected magnitude of an advance during any stage of the development process is also finite.

As the product development proceeds, an arrival of a better competing product in the market will result in lowering the relative quality ranking of, and thereby adversely affecting the demand for, our product. Such a deterioration in the relative quality may be looked upon as a partial failure, whose magnitude depends upon the extent of advance achieved by the rivals. The arrivals of these competing innovations into the market may also be assumed to be generated by Poisson processes as above, which, however, are beyond the control of our firm. Let $\mu_k \geq 0$ denote the Poisson rate at which a partial failure of magnitude k occurs, where $k \in \{1, 2, \ldots.\}$, so that the time intervals between successive partial failures of magnitude k are exponentially distributed with mean $\frac{1}{\mu_k}$. Such an assumption is reasonable if there is a large number of competing innovators, whose behavior pattern is stationary in time and, independent of one another. As before, we will assume that $\mu = \sum_{k=1}^{\infty} \mu_k < \infty$ and that $\sum_{k=1}^{\infty} k\ \mu_k < \infty$ implying that only a finite number and magnitude of rival innovations are possible in a finite amount of time and that the expected amount of reduction caused by each such innovation is finite.

$$(2.1) \qquad R(q) = \begin{cases} \beta(q - \underline{q}) & \text{if} \quad q \geq \underline{q} \\ 0 & \text{if} \quad q \leq \underline{q} , \end{cases}$$

This terminal reward to be received upon the completion of the development project serves as the incentive for investing capital and labor resources during its conduct. An allocation of resources may be aggregated into the monetary expenditure $a \in [0, B]$, where the upper bound B corresponds to the maximum amount of resources at the firm's disposal. The interest rate will be denoted by $r > 0$, so that one dollar at time t has the present value e^{-rt}.

At any point in time during the conduct of the project, the relative product quality achieved summarizes the cumulative result of past successes and failures due to past allocations. As the project proceeds, the product quality changes stochastically through time. A partial success corresponds to an improvement in the product quality as a result of a research finding or a scientific discovery or a technological breakthrough by our firm, where such an advance is assumed to take place according to a Poisson process (as in Gaver and Srinivasan [5]). The Poisson rate at which an improvement in the product quality takes place is controlled by the current resource allocation, where a greater allocation will be assumed to stochastically reduce the time required for an improvement, (a stochastic analog of the time-cost tradeoff relationship in R and D studied by Scherer [20]). Furthermore, such an improvement may be of differing magnitudes depending upon the degree of technical advance achieved. Given an allocation $a \in [0, B]$, the uncertainty regarding the time of the next partial success may then be represented by the independent Poisson processes of rates $\lambda_i(a) \geq 0$ where i is the magnitude of the improvement taking values in $\{1, 2, ...\}$. Suppose that for each i, $\lambda_i(a)$ is continuous and increasing in a with $\lambda_i(0) = 0$

Let $X_i(a)(Y_k)$ be the exponential random variable denoting the time interval between successive quality improvements (reductions) of magnitude $i = 1,2,\ldots.$ $(k = 1,2,\ldots.)$. Then $X(a) = \text{Inf}\ \{X_1(a), X_2(a),\ldots\}$ denotes the time interval between successive improvements and is exponentially distributed with parameter

$$(2.2) \quad \lambda(a) = \sum_{i=1}^{\infty} \lambda_i(a).$$

Similarly, $Y = \text{Inf}\ \{Y_1, Y_2\ldots\}$ is the time interval between successive reductions and has the exponential distribution with parameter

$$(2.3) \quad \mu = \sum_{k=1}^{\infty} \mu_k$$

Consequently, the time interval between successive changes in the product quality q is a random variable $Z(a) = \text{Min}\ \{X(a), Y\}$ with exponential distribution with parameter

$$(2.4) \quad \Lambda(a) = \lambda(a) + \mu$$

Now the evolution of the project through time, in terms of the product quality developed, may be described as follows. Suppose that the project starts at time $t_0 = 0$ with the initial quality $q_0 = q$ of the product on hand. The immediate decision to be made is whether to develop this product further before marketing it or to market it as it is, to be denoted by actions c (continue) and s (stop) respectively. If the stopping action s is taken, the development process ends (i.e. the project is not undertaken) and the product is marketed as it is, yielding $R(q_0)$. If the continuation action c is taken, the manager must, in addition, determine the resource allocation $a_0 \in [0,B]$ at time 0. The product quality changes at a random time as a combined result of such an allocation, the internal uncertainties regarding the time of the next improvement (as summarized in $\lambda(a_0)$), and the external uncertainties regarding the time of the next deterioration (as summarized in μ). Thus, the time $Z(a_0) = t_1$ at which such a change in the project status occurs is selected according to the exponential distribution with parameter $\Lambda(a_0)$. Given that such a change occurs at time t_1, the new product quality is therefore

$$q_1 = \begin{cases} q_0 + i & \text{with probability } \dfrac{\lambda_i(a_0)}{\Lambda(a_0)}, \ i = 1, 2, \ldots. \\ q_0 - k & \text{with probability } \dfrac{\mu_k}{\Lambda(a_0)}, \ k = 1, 2, \ldots. \end{cases}$$

In general, at the n^{th} decision epoch at time t_n $(n = 0, 1, 2, \ldots.)$ the state of the project in terms of the currently developed quality q_n is observed and a stop (s) or continue (c) decision is made. The stop decision terminates the process, taken to mean technically that the process instantaneously goes to $-\infty$ so that $t_{n+1} = t_n$ and $q_{n+1} = -\infty$ with a final reward $R(q_n)$, thereby rendering any further resource allocation fruitless and unnecessary. The continue decision on the other hand, necessitates the choice of a further resource allocation. If the allocation $a_n \in [0, B]$ is chosen then the time interval until the next transition, $t_{n+1} - t_n$, has the exponential distribution with parameter $\Lambda(a_n)$, at the end of which the new product quality is

$$(2.5) \qquad q_{n+1} = \begin{cases} q_n + i & \text{with probability } \dfrac{\lambda_i(a_n)}{\Lambda(a_n)}, \ i = 1, 2, \ldots. \\ q_n - k & \text{with probability } \dfrac{\mu_k}{\Lambda(a_n)}, \ k = 1, 2, \ldots \end{cases}$$

The above birth and death process dynamics resemble the discrete time random walk model proposed by Radner and Rothschild [19] to represent the progress of any activity. They, however, consider the implications of following certain reasonable behavioral rules, rather than determining optimal strategies for controlling such a process.

In order to control the evolution of the project through time, the R and D manager uses a strategy for making the stop or continue decisions together with resource allocations at the decision epochs. A development strategy π is a sequence of (possibly randomized) decision rules $\{\pi_n : n = 0, 1, 2, \ldots\}$. Here π_n specifies at the decision epoch n whether to stop the development process together with an expenditure $a_n \in [0, B]$ in case of continuation, as a function of the past history, i.e. $\pi_n(\cdot|h_n)$ is a probability measure on the decision space $D = \{s, c\} \times [0, B]$, conditional on the past history $h_n = (q_0, d_0, \ldots q_n)$ of states and decisions. A strategy is said to be (non randomized) Markov if $\pi = \{\delta_n : n = 0, 1, 2, \ldots\}$, where $\delta_n : Q \to D$, specifies a decision $\delta_n(q_n)$ if the product quality at the n^{th} decision epoch is $q_n \in Q$. A (non randomized) stationary strategy is given by a single function $\delta : Q \to D$, so that at any decision epoch, if the product quality is q, the decision $\delta(q)$ is specified, thereby eliminating the need for remembering the past progress and allocations.

Starting with an initial quality $q_0 = q$ and following a development strategy π, let

(2.6) $$N = N(q, \pi) = \text{Inf } \{n : d_n \in \{s\} \times [0,B]\}$$

be the random stopping time (possibly infinite) at which the development process is terminated, thereby yielding the reward $R(q_N)$. The net expected discounted return, starting with quality q and following strategy π may then be written as

(2.7) $$W(q, \pi) = E_\pi [e^{-rt_N} R(q_N) - \sum_{n=0}^{N} e^{-rt_n} a_n | q_0 = q]$$

The optimal return function is then denoted by

(2.8) $$V(q) = \sup_\pi W(q, \pi) \quad , q \in Q$$

which may be called the project value function. It is the worth of the project if the product of initial relative quality q is developed optimally. A strategy π^* is said to be optimal at q if $W(q,\pi^*) = V(q)$ and is said to be optimal if it is optimal at every $q \in Q$.

Our objective is to characterize the project value function and the optimal development strategy, so as to determine whether it is worthwhile undertaking the project; if so, how best to finance it through time; and finally, the point at which it is best to terminate the project. In the next section we establish the existence and the structure of the project value function and the existence of an optimal (stationary) strategy, using the methods developed by Lippman [14] and Ross [20], based on the fundamental works by Blackwell [2], Strauch [22] and Maitra [16].

3. THE PROJECT VALUE FUNCTION

In order to establish the existence of an optimal stationary strategy and the functional equation satisfied by the project value function, we invoke the results of Lippman [14], upon verifying that his conditions are met in our problem. An upper bound on the expected discount factor between successive decision epochs is given by

$$\sup_{(c,a)\in D} \int_0^\infty e^{-rt}\, \Lambda(a)\, e^{-\Lambda(a)t}\, dt = \frac{\Lambda(B)}{r+\Lambda(B)} < 1$$

so that Lippman's assumption 1 is satisfied. Next, in our problem, the immediate return function is

$$r(q,d) = \begin{cases} R(q) - a & \text{if } d = (s,a) \\ -a & \text{if } d = (c,a) \end{cases} \tag{3.1}$$

By taking m = 1, and defining

$$w(q) = \left(1 + \frac{\lambda(B)}{\Lambda(B)}\right) \max\{R(q), 1\} \quad \text{and} \tag{3.2}$$

(3.3) $$b = \beta(1 + \frac{\lambda(B)}{\Lambda(B)}) \sum_{i=1}^{\infty} i \frac{\lambda_i(B)}{\Lambda(B)}$$, we find that

$$\sup_{q \in Q} \frac{|\sup_{d \in D} r(q,d)|}{w(q)} \leq \frac{\Lambda(B)}{\lambda(B)+\Lambda(B)} < \infty$$

and that, for each $q \in Q$,

$$\sup_{(c,a)\in D} \left[\sum_{i=1}^{\infty} w(q+i) \frac{\lambda_i(a)}{\Lambda(a)} + \sum_{k=1}^{\infty} w(q-k) \frac{\mu_k}{\Lambda(a)} \right] \leq w(q) + b$$

while $\sup_{(s,a)\in D} [w(-\infty)] \leq w(q) \leq w(q) + b$, so that Lippman's assumptions 2 and 3 also hold in our problem. Furthermore, the quality space Q and the decision space D are clearly complete, separable metric spaces, and D is, in addition, compact. Finally, the expected discount factor $\frac{\Lambda(a)}{r+\Lambda(a)}$, the immediate return function $r(q,d)$ and the transition probabilities $\frac{\lambda_i(a)}{\Lambda(a)}$ and $\frac{\mu_k}{\Lambda(a)}$ are all continuous in d. Therefore, applying Lippman's theorem 1 to our problem yields the following

<u>Proposition 1</u> (Existence and uniqueness)

The project value function $V:Q \to \mathbb{R}$ as defined in (2.8) is the solution of the optimality equation (3.4) and (3.5) below, for all $q \in Q$,

(3.4) $V(q) = \mathrm{Max}\ \{R(q); U(q)\}$ where

(3.5) $$U(q) = \sup_{a\in[0,B]} \left[-a + \sum_{i=1}^{\infty} V(q+i) \frac{\lambda_i(a)}{r+\Lambda(a)} + \sum_{k=1}^{\infty} V(q-k) \frac{\mu_k}{r+\Lambda(a)} \right]$$

which is unique in the Banach space

$$\mathcal{B} = \{u:Q \to \mathbb{R} \mid \sup_{q\in Q} \frac{|u(q)|}{w(q)} < \infty\} ,$$

where $w(q)$ is as in (3.2). Furthermore, there exists an optimal stationary development strategy $\delta^*: Q \to D$, which, whenever the project is in state q, selects a decision $\delta^*(q)$ maximizing the right-hand side of (3.4).

An immediate consequence of the result that there exists an optimal development strategy that is stationary, is that during the course of the project, the optimal investment decision depends only on the current status of the project which summarizes the effectiveness of past investments and progress history.

The project value $V(q)$ is the net expected discounted return from the project, starting with an initial product of quality q and optimally stopping or continuing its further development. Similarly, $U(q)$ is the optimal return if we are forced to continue for one stage and if we follow an optimal policy then on. The next proposition characterizes the structure of the project value function $V(\cdot)$.

Proposition 2 (Properties of $V(\cdot)$)

For all q, q_1, $q_2 \in Q$ with $q_1 \leq q \leq q_2$, we have

(3.6) $\qquad V(q) \geq 0 \qquad$ (nonnegativity),

(3.7) $\qquad V(q_2) \geq V(q_1) \qquad$ (monotonicity),

(3.8) $\qquad V(q) \leq \lambda V(q_1) + (1-\lambda)\, V(q_2)$, where $\lambda = \dfrac{q_2-q}{q_2-q_1} \in [0,1] \qquad$ (convexity) .

Proof:

From (3.4) we have $V(q) \geq R(q) \geq 0$ for all q i.e. (3.6) holds. To prove (3.7) and (3.8) let

$$V_{n+1}(q) = \operatorname{Max} \begin{cases} R(q) \\ U_{n+1}(q) \end{cases} \qquad n = 0,1,2,\ldots, q \in Q \tag{3.9}$$

where

$$U_{n+1}(q) = \sup_{a \in [0,B]} \left[-a + \sum_{i=1}^{\infty} V_n(q+i) \frac{\lambda_i(a)}{r+\Lambda(a)} + \sum_{k=1}^{\infty} V_n(q-k) \frac{\mu_k}{r+\Lambda(a)} \right] \tag{3.10}$$

with $V_0(q) = R(q)$, $q \in Q$ be the usual finite stage versions of (3.4) and (3.5). Clearly, $V_0(q)$ is nondecreasing and convex, from (2.1). Suppose that $V_n(\cdot)$ is nondecreasing and convex. Then, for each $a \in [0,B]$, the bracketed quantity in (3.10) is nondecreasing and convex, so that $U_{n+1}(\cdot)$ also has these properties and hence, from (3.9), so does $V_{n+1}(\cdot)$, completing the induction argument. Thus, each $V_n(\cdot)$ is nondecreasing and convex, and, since for some J, $V_{nJ}(\cdot)$ converges to $V(\cdot)$ (by the usual J-stage contraction property), these properties are preserved in the limit. QED.

By combining this proposition with the definition (3.5) we get

Corollary: $U(\cdot)$ is also nonnegative, nondecreasing and convex.

Thus, optimal product development never results in losses in the long run, while starting the project with a better quality product on hand always yields a higher net return and, furthermore, this marginal advantage from starting with a better product increases with its quality.

In the next section, using the above results, we establish the structure of an optimal stationary investment strategy as a function of the (current) quality attained.

4. THE OPTIMAL INVESTMENT STRATEGY

Define the maps $\bar{V}:Q \times D \to \mathbb{R}$ and $\bar{U}:Q \times [0,B] \to \mathbb{R}$, by

$$(4.1) \quad \left\{ \begin{aligned} \bar{V}(q,d) &= \begin{cases} R(q) - a & \text{if} \quad d = (s,a) \\ \bar{U}(q,a) & \text{if} \quad d = (c,a) \end{cases} \\ \bar{U}(q,a) &= -a + \frac{1}{r+\Lambda(a)} \left[\sum_{i=1}^{\infty} V(q+i)\, \lambda_i(a) + \sum_{k=1}^{\infty} V(q-k)\, \mu_k \right] \end{aligned} \right.$$

so that (3.4) and (3.5) can be written compactly as

$$(4.2) \qquad V(q) = \sup_{d \in D} \bar{V}(q,d) \ , \ q \in Q$$

By Proposition 1, the optimal development strategy $\delta^*:Q \to D$ has the property

$$(4.3) \qquad \bar{V}(q,\delta^*(q)) = \sup_{d \in D} \bar{V}(q,d) = V(q)$$

In case of more than one d's attaining the supremum in (4.3), $\delta^*(q)$ will be understood to specify stopping rather than continuation and less expenditure rather than more, thus implying a risk-averting and thrifty behavior.

The optimal development strategy $\delta^*:Q \to D$ may be decomposed into the <u>optimal stopping strategy</u> $\delta_1^*:Q \to \{s,c\}$ and the <u>optimal allocation strategy</u> $\delta_2^*:Q \to [0,B]$, so that $\delta^* = (\delta_1^*, \delta_2^*)$. The next two propositions establish the structure of δ_1^* and δ_2^*, respectively.

<u>Proposition 3</u> (Optimal stopping region)

There exist q_s, $q^s \in Q$ such that $q_s \leq \underline{q} \leq q^s < \infty$ and $\delta_1^*(q) = s$ if and only if $q \leq q_s$ or $q \geq q^s$.

<u>Proof</u>:

Define

$$(4.4) \qquad q_s = \sup\{q \in Q : V(q) = 0\}.$$

Then, from (2.1) and (3.4), $q_s \leq \underline{q}$. If $q \leq q_s$, then using nonnegativity and monotonicity of Proposition 2 and its corollary, we have

$$0 = V(q) = R(q) = U(q)$$

so that $\delta_1^*(q) = s$, since in the case of a tie δ^* specifies stopping. Similarly, if $q_s < q \leq \underline{q}$, then $V(q) > 0 = R(q)$, implying $\delta_1^*(q) = c$.

Next, we show that there is a unique point q^s of intersection of $R(q)$ and $U(q)$ with $q \geq \underline{q}$ having the desired properties. Suppose that there does not exist such a point, so that for all $q \geq \underline{q}$ we have $U(q) > R(q)$ and therefore $\delta_1^*(q) = c$, for all $q > q_s$. Then $V(q) = U(q) > R(q) \geq 0$ for all $q \geq \underline{q}$. Now, for each $q > q_s$ define the stopping time

$$N = N(q,\delta^*) = \text{Inf}\ \{n : q_n \leq q_s\},$$

so that,

$$V(q) = W(q,\delta^*) \leq E_{\delta^*}[e^{-rt_N} R(q_N) | q_0 = q]$$

Since on $\{N < \infty\}$ we have $R(q_N) = 0$ and on $\{N = \infty\}$ we have $e^{-rt_N} = 0$ (since $\Lambda(B) < \infty$), and since $V(q) \geq 0$, it follows that $V(q) = 0$ for all $q > q_s$, yielding a contradiction. Hence $U(q) > R(q)$ for all $q \geq \underline{q}$ is impossible and we may define

$$q^s = \min\{q : q \geq \underline{q} \text{ and } U(q) \leq R(q)\}. \tag{4.5}$$

Next we claim that $U(q) \leq R(q)$ for all $q \geq q^s$. Otherwise, let $\tilde{q} = \text{Min}\{q : q \geq q^s \text{ and } U(q) > R(q)\}$.

Then, by convexity of $U(\cdot)$ and linearity of $R(\cdot)$, we have $U(q) > R(q)$ for all $q \geq \tilde{q}$, so that $\delta_1^*(q) = c$ for all $q \geq \tilde{q}$ and $\delta_1^*(q) = s$ for all q such that $q^s \leq q < \tilde{q}$. Again, starting in $q \geq \tilde{q}$, define, under δ^*,

$$N = \text{Inf}\ \{n : q^s \leq q_n < \tilde{q} \text{ or } q_n \leq q_s\}$$

yielding

$$V(q) \leq R(\tilde{q}-1) = V(\tilde{q}-1) \text{ for all } q \geq \tilde{q}$$

which contradicts convexity of $V(\cdot)$. Therefore, $U(q) \leq R(q)$ for all $q \geq q^s$, so that $\delta_1^*(q) = s$ whenever $q \geq q^s$.

Finally, by the definition of q^s we have $\delta_1^*(q) = c$ whenever $\underline{q} \leq q < q^s$, which, together with the fact (shown in the first paragraph of this proof) that $\delta_1^*(q) = c$ whenever $q_s < q \leq \underline{q}$, characterizes the convex continuation region (q_s, q^s). QED.

We may interpret q^s as the "optimum goal" the development manager should strive for. Thus, as soon as the relative product quality exceeds q^s, it is best to stop the development and market the product, because the immediate reward from doing so exceeds the maximum expected return from continuing the development further. Similarly, q_s may be interpreted as the minimum acceptable quality level necessary to justify further development. As soon as the relative product quality drops below q_s, it is best to abandon the project, because our performance has been too unsatisfactory (in relation to that of the competitors) to justify further investments.

The problem of determining whether to undertake a project or not becomes a special case of the problem of determining the optimal stopping strategy. Thus, if the product initially on hand is so inferior that $q_0 \leq q_s$, then it is not worthwhile to undertake the project, while, if the product is so superior initially itself that $q_0 \geq q^s$, then its further development is unnecessary.

An intermediate quality q, $q_s < q < q^s$, represents a promising product worth improving through R and D expenditures. The following proposition shows that, in the continuation region (q_s, q^s) specified by the optimal stopping strategy δ_1^*, it is optimal to pursue a vigorous and aggressive allocation strategy δ_2^*, which is strictly positive and increasing in the current progress.

Proposition 4 (Optimal allocation strategy)

With q_s and q^s as in Proposition 3, the optimal allocation strategy δ_2^* satisfies

(4.6) $\quad \delta_2^*(q) = 0 \quad$ if $q \leq q_s$ or $q \geq q^s$,

(4.7) $\quad \delta_2^*(q) > 0 \quad$ if $q_s < q < q^s$, and

(4.8) $\quad \delta_2^*(q_2) \geq \delta_2^*(q_1) \quad$ if $q_s < q_1 \leq q_2 < q^s$

Proof:

Statement (4.6) is immediate from (4.1) and (4.2).

If $q_s = \underline{q}$, then from the proof of Proposition 3 it follows that $q^s = q_s = \underline{q}$, so that (4.7) and (4.8) follow vacuously. Hence it suffices to consider the case in which $q_s < \underline{q}$.

Suppose $\delta_2^*(q) = 0$ for some $q \in (q_s, q^s)$. Then

$$V(q) = U(q) = \overline{U}(q, \delta_2^*(q)) = \sum_{k=1}^{\infty} V(q-k) \frac{\mu_k}{r+\mu},$$

using the definition of (q_s, q^s), (4.1), (4.3) and the assumption $\lambda_i(0) = 0$ for all i. Hence, from monotonicity of $V(\cdot)$, as in (3.7), we have $V(q) \geq V(q-k)$ for all $k \geq 1$, so that the above implies that

$$V(q) \leq \frac{\mu}{r+\mu} V(q),$$

which together with nonnegativity of $V(\cdot)$ implies that $V(q) = 0$. But this contradicts (4.4), since $q > q_s$, completing the proof of (4.7).

To show monotonicity of $\delta_2^*(\cdot)$ in the continuation region, note that, if $q_s < q < q^s$, then

$$V(q) = U(q) = \overline{U}(q, \delta_2^*(q)),$$

so that

$$V(q) = \sup_{a \in [0,B]} \{-a + \sum_{i=1}^{\infty} V(q+i) \frac{\lambda_i(a)}{r+\Lambda(a)} + \sum_{k=1}^{\infty} V(q-k) \frac{\mu_k}{r+\Lambda(a)} \}$$

which can be seen to be equivalent to

$$\text{(4.9)} \qquad rV(q) = \sup_{a \in [0,B]} \{-a[r+\Lambda(a)] + \sum_{i=1}^{\infty} [V(q+i) - V(q)] \lambda_i(a) + \sum_{k=1}^{\infty} [V(q-k) - V(q)] \mu_k\}$$

Now define

$$F(q,a) = -a[r + \Lambda(a)] + \sum_{i=1}^{\infty} [V(q+i) - V(q)]\lambda_i(a) + \sum_{k=1}^{\infty} [V(q-k) - V(q)]\mu_k$$

so that (4.9) becomes

$$\begin{aligned} rV(q) &= \sup_{a\in[0,B]} \{F(q,a)\} \\ &= F(q,\delta_2^*(q)) \quad , \quad q_s < q < q^s \end{aligned}$$

Now, if $q_s < q_1 \leq q_2 < q^s$, then by convexity of $V(\cdot)$ and monotonicity of $\lambda_i(\cdot)$ for each i, it follows that, $[F(q_2,a) - F(q_1,a)]$ is nondecreasing in a. But this implies that $\delta_2^*(q_2) \geq \delta_2^*(q_1)$, because otherwise we get

$$F(q_2,\delta_2^*(q_1)) - F(q_1,\delta_2^*(q_1)) > F(q_2,\delta_2^*(q_2)) - F(q_1,\delta_2^*(q_2))$$

contradicting optimality of δ_2^*. QED.

5. REMARKS

The economically meaningful characterization of the entire stationary optimal development strategy $\delta^*:Q \rightarrow D$, given in Propositions 3 and 4, may now be summarized by

$$(5.1) \qquad \delta^*(q) = \begin{cases} (s,0) & \text{if} \quad q \leq q_s \text{ or } q \geq q^s \\ (c,\delta_2^*(q)) & \text{if} \quad q_s < q < q^s \end{cases}$$

where the expenditure $\delta_2^*(q)$ is positive and nondecreasing over the convex continuation region (q_s,q^s), and the optimum critical levels q_s and q^s are given by (4.4) and (4.5).

This problem of characterizing δ may be viewed as a generalization of the optimal stopping problem (see e.g. Breiman [3]). In our problem, in addition to the usual stop or continue decisions, we have allowed for the choice of intensity with which to continue (searching), which in turn also affects the future "offers". The control limit policy form of the optimal stopping region established here resembles that of Wald's sequential hypothesis testing procedure (e.g. see Ross [20] Ch. 6), the intensity of sampling in the continuation region being the additional decision variable.

The effect of varying the parameters of the problem on the control limits q_s and q^s may be easily investigated. From the definitions (3.5), (4.4) and (4.5) of $U(q)$, q_s and q^s, respectively, the continuation region (q_s, q^s) may be easily seen to be nonincreasing in μ and r and nondecreasing in $\lambda(\cdot)$ and B. Thus a lesser competition, a lower interest rate, and a greater technological or financial capability induce us to pursue a higher goal and not give up too soon.

The structure of the optimal strategy δ^* summarized in (5.1) enables us to write, analogous to (2.6), (2.7) and (2.8),

$$(5.2) \qquad V(q) = \sup_{\substack{\langle q_s, q^s\rangle \in Q^2 \\ q_s \le q^s}} \quad \sup_{\delta_2 \in \Delta(q_s, q^s)} \overline{W}(q, \delta_2)$$

where

$$\Delta(q_s, q^s) = \{\delta_2 : Q \to [0,B] : \delta_2(q) = 0 \text{ if } q \le q_s \text{ or if } q \ge q^s, \text{ and } \delta_2(q) \text{ is positive and nondecreasing in } q \in (q_s, q^s)\}$$

is the set of stationary allocation strategies in a project with prespecified goals q_s and q^s, while

$$(5.3) \qquad \overline{W}(q, \delta_2) = E_{\delta_2}\left[e^{-rt_N} R(q_N) - \sum_{n=0}^{N-1} e^{-rt_n} \delta_2(q_n) \middle| q_0 = q \right] \text{ with}$$

$$(5.4) \qquad N = \operatorname{Inf}\{n : q_n \le q_s \text{ or } q_n \ge q^s\}$$

The entire project planning problem may be thought of as that of determining optimal goals q_s and q^s and an optimal adaptive allocation strategy δ_2^* for attaining these goals. Equation (5.2) suggests an iterative method for arriving at a desired plan, which may be heuristically described as follows, and which has an obvious behavioral interpretation. Given an initial project status, a set of reasonable goal levels is selected and the corresponding optimal strategy for attaining these goals is determined (by using, say, the policy improvement routine of Howard [10]). The goals are then appropriately modified and the process is repeated until a satisfactory combination of the goals and the allocation strategy is found.

References

[1] Aldrich, C. and Morton, T. E., "Optimal Funding Paths for a Class of Risky R&D Projects," Management Science, Vol. 21, No. 5, pp. 491-500 (1975).

[2] Blackwell, D., "Discounted Dynamic Programming," Annals of Mathematical Statistics, Vol. 36, pp. 226-235, (1965).

[3] Breiman, L., "Stopping Rule Problems," Ch. 10 in Applied Combinatorial Mathematics, E. F. Beckenbach (Editor), Wiley (1964).

[4] Dorfman, R. and Steiner, P. O., "Optimal Advertising and Optimal Quality," American Economic Review, 1954 (pp. 826-836)

[5] Gaver, D. P. and Srinivasan, V., "Allocating Resources Between Research and Development: A Macro Analysis," Management Science, Vol. 18, pp. 492-501 (1972).

[6] Gittins, J. C., "Optimal Resource Allocation in Chemical Research," Advances in Applied Probability, Vol. 1, pp. 238-270 (1969).

[7] Gittins, J. C., "Some Problems of Stochastic Resource Allocation," Journal of Applied Probability, Vol. 9, pp. 360-369 (1972).

[8] Gittins, J. C., "Resource Allocation in Speculative Chemical Research," Journal of Applied Probability, Vol. 11, pp. 255-265 (1974).

[9] Hess, S. W., "A Dynamic Programming Approach to R&D Budgeting and Project Selection," IRE Transactions on Engineering Management, Vol. EM 9, pp. 170-178 (1962).

[10] Howard, R. A., Dynamic Programming and Markov Processes, Technology Press and Wiley (1960).

[11] Kamien, M. I., and Schwartz, N. L., "Expenditure Patterns for Risky R&D Projects," Journal of Applied Probability, Vol. 8, pp. 60-73 (1971).