JOURNAL OF ECONOMIC THEORY 10, 358–376 (1975)

On the Allocation of Effort*

ROY RADNER

Department of Economics and Statistics,
University of California, Berkeley, California 94720

AND

MICHAEL ROTHSCHILD

Department of Economics,
Princeton University, Princeton, New Jersey 08540

Received October 16, 1974

1. INTRODUCTION

This paper sets forth a fairly simple model that, we hope, captures the salient points of many managerial dilemmas. Consider a manager who is in charge of several activities. At each point in time, he either devotes his effort to an activity or disregards it. Unattended activities tend to deteriorate (stochastically) while attended activities tend to improve (also stochastically). The manager's problem is to decide how to allocate his effort—which is available only in limited amounts—among the various activities.

There are two distinct approaches that economists could use to analyze such a model. The first would be to describe or derive the manager's preference ordering over all possible outcomes of the process, to find the strategy that maximizes this preference (or its expectation), and then finally to analyze the characteristics of this optimal strategy. The second approach, and that which we follow here, is simply to derive the properties of certain plausible, but not necessarily optimal, behavioral rules. Our preference for this latter approach rests on the conviction that it is simply not tenable to maintain that managers in complex situations, like those which our model describes, formulate complete preference orderings, find

* This paper is based on research supported by the National Science Foundation, the Mathematical Social Science Board, and the Industrial Relations Section of Princeton University. The authors are grateful to David Freedman, Wilhelm Neuefeind and Sidney Winter for helpful suggestions.

optimal strategies, and pursue them. Competing and imperfectly articulated goals, limits on computational ability, organizational slack, and possibly pure sloth combine to make it reasonable that real decision-makers follow attractive, simple rules rather than optimal strategies. These matters are, of course, more complicated than these brief remarks imply. Here it should suffice to say that these notes are an example of the analysis of economic systems in terms of "bounded rationality"—a viewpoint which one of us has described more fully elsewhere [5].

Our model is as follows: There are finitely many activities. Each activity has two modes of evolution. When attended, an activity is a random walk with positive drift. When neglected, it is a random walk with negative drift. We assume that in each period the manager has but one unit of effort to allocate to the various activities and that it makes sense to speak of fractional allocations of effort. A behavior is a rule that determines the current allocation of effort among the activities as a function of the past history of performance.

For example, consider the task of a purchasing agent. The agent must buy a number of items each period. We can assume that, unless renegotiated, contracts of the preceding period tend to stay in effect. In each period the purchasing agent may search for new and better suppliers or attempt to negotiate for improved terms with old suppliers. Although in a stochastic world there is no guarantee that attempts to improve will not lead to temporary setbacks—new materials may work less well than hoped and new suppliers may prove unreliable—it is reasonable that these activities should be productive on the average. On the other hand, it is plausible that unattended contracts tend to deteriorate. For example, general price decreases may not be passed along, or quality may decline when a customer does not evince a continuing interest in improving the terms on which he buys.

Another example is provided by a manager in charge of several independent productive processes. In each period he may devote effort to maintenance of one part of this system. Parts that are maintained tend to improve, while those that are neglected deteriorate. As a third example, consider the situation of a researcher who tries to keep up with the several branches of his field. He is limited to reading one article at a time and finds the reward from time spent examining recent work to be quite variable. However, if he neglects an area for too long, he loses touch with it.

The model obviously has many other applications. We would stress that it can be understood to apply at any of several levels of abstraction. Thus, the activities themselves may be strategies of rules of thumb.

In many of its interpretations, it may not be reasonable to suppose that the manager knows the parameters that govern the process or even completely understands the nature of the process itself. Accordingly, we have emphasized in our analysis (particularly in Sections 5 and 6, below) behavioral rules that can be pursued when such knowledge is absent.

In such a model, performance of the several activities will not typically approach a steady state, even in a stochastic sense, except for very special values of the parameters. We examine the effects of different behaviors on asymptotic performance with respect to two criteria:

(i) the probability of survival, i.e., the probability that performance on one or more activities never falls below certain prescribed levels;

(ii) the long-run average rate of growth per unit time.

The analysis will be concentrated on three types of behavior:

(i) "constant proportions," in which the allocation of effort is constant over time;

(ii) "putting out fires," in which all effort at any date is allocated to those activities that have the worst performance at that date;

(iii) "staying with a winner," in which all effort at any date is allocated to those activities that have the best performance at that date.

2. A Formal Model of Controlled Random Walk

Consider an agent who supervises several activities, indexed by $i = 1,..., I$. At each date $t = 0, 1, 2,...$, and for each activity i, the individual perceives a measure of performance, $U_i(t)$. As a function of the history of the vector $U(n) = (U_i(n))$ for $n = 0,..., t$, the agent allocates to each activity i some fraction a_i of his effort during the coming period of time. The vector $a = (a_i)$ of fractions is nonnegative, and the sum of its coordinates is unity (the individual devotes all of his effort to the I activities).

Given past history up through date t, the conditional distribution of the next vector, $U(t + 1)$, of performance levels depends upon the vector $a(t)$ of allocations at date t. Define

$$Z(t + 1) = U(t + 1) - U(t). \tag{2.1}$$

The sequence of vectors $Z(t)$ is the sequence of successive increments in the vectors of performance levels, $U(t)$. Roughly speaking, we shall assume that, given the allocation vectors, the successive increments are independent random vectors. The larger the allocation of effort to any activity, the larger will be the expected increment in its performance. If the agent allocates all of his effort to a particular activity, then its expected increment will be positive; if he allocates none of his effort to the activity, then its expected increment will be negative. In other words, the sequence of performance levels is a *controlled random walk*, with the expected values of the successive increments depending on the corresponding allocations of effort.

To be precise, we shall make the following assumptions about the *conditional* distribution of $Z(t+1)$, given the sequence $U(0),..., U(t)$ and the allocation $a(t)$:

The distribution of $Z(t+1)$ depends only on $a(t)$. (2.2a)

$EZ_i(t+1) = a_i(t)\,\eta_i - [1 - a_i(t)]\,\xi_i$, where ξ_i and η_i are given positive parameters. (2.2b)

$\operatorname{Var} Z_i(t+1) = s_i^2[a_i(t)]$, where s_i is a given strictly positive continuous function. (2.2c)

The coordinates of $Z(t+1)$ are mutually independent. (2.2d)

To minimize technical complications, we also assume the following:

The coordinates of $Z(t+1)$ are integer-valued and uniformly bounded. (2.2e)

It will be useful to have a compact notation to denote partial histories of the process of performance levels and allocations. Following standard practice in probability theory, we regard all random variables as measurable functions on a common probability space, which is endowed with a sigma field $\mathscr{F}$ of (measurable) subsets. Corresponding to each date t is a sigma field $\mathscr{F}_t$ of subsets representing all events that can be observed up to and including date t. Each $\mathscr{F}_t$ contains the preceding one, $\mathscr{F}_{t-1}$, and all $\mathscr{F}_t$ are contained in $\mathscr{F}$. For any $s \leqslant t$, the random vectors $U(s)$, $Z(s)$, and $a(s)$ are measurable with respect to $\mathscr{F}_t$.

A *behavior* is a sequence of random allocation vectors $a(t)$ such that each $a(t)$ is measurable with respect to $\mathscr{F}_t$. A *constant proportions behavior* is one in which the allocations $a(t)$ are the same for all t, i.e., a behavior such that the allocation of effort is the same at all dates and for all histories of the process.

In subsequent sections we shall discuss the consequences of different behaviors for the (random) sequence of performance vectors. However, this is a good place to point out that there is a particular way of measuring "average performance" such that expected increments in average performance do not depend on the agent's behavior.

Define a vector $w = (w_i)$ of "weights" by

$$w_i = \left(\frac{1}{\eta_i + \xi_i}\right)\left(\sum_j \frac{1}{\eta_j + \xi_j}\right)^{-1}. \tag{2.3}$$

These weights are positive, and their sum is unity. Define the corresponding weighted averages of performance and performance increments by

$$\bar{U}(t) = \sum_i w_i U_i(t), \qquad \bar{Z}(t) = \sum_i w_i Z_i(t). \tag{2.4}$$

Notice that

$$\bar{Z}(t+1) = \bar{U}(t+1) - \bar{U}(t). \tag{2.5}$$

It is easy to verify from (2.2b) that for *any* behavior the conditional expected value of $\bar{Z}(t)$, given the past history of the system up through date $(t-1)$, is

$$\bar{\zeta} \equiv \left(1 - \sum_i \frac{\xi_i}{\eta_i + \xi_i}\right)\left(\sum_i \frac{1}{\eta_i + \xi_i}\right)^{-1}. \tag{2.6}$$

This parameter, $\bar{\zeta}$, of the performance process, which does not depend on the agent's behavior, will have an important role in the subsequent analysis.

Remark. The model just formulated falls within the framework of Markovian dynamic programming. One may take $U(t)$ to be the state of the system at date t and $a(t)$ to be the action. What we have called here a behavior is what is usually called a policy; we have adopted the former term in order to avoid any connotation of optimization. Nothing has been specified to correspond to the usual "immediate return function" of dynamic programming.

3. SURVIVAL

By *survival*, we shall mean staying away from unacceptably low levels of performance. This statement is clearly imprecise, and there are many alternative ways to make it precise. In this section we shall concentrate on one particular alternative, in which *all* of the performance indices are required to be kept above prescribed levels.

"Survival" is perhaps too dramatic a word for what we have in mind here. For the situations that our model describes, the consequence of failure to survive is often not death but reorganization. That is, if a process fails to maintain a certain standard of performance, then the way in which it is managed will be changed, but if its performance is acceptable its management will be allowed to continue unchanged.

By convention, we shall take zero to be the unacceptably low level of performance for each index. Define

$$M(t) = \min_i U_i(t). \tag{3.1}$$

We shall define *survival* by the condition,

$$M(t) > 0 \quad \text{for all } t. \tag{3.2}$$

In other words, the agent fails to survive as soon as any one index of performance falls to zero (or lower).

In the model of the previous section it will not be possible to guarantee survival with probability one. One may at least ask whether there is any behavior for which the probability of survival is positive. The next proposition states that this is possible if and only if $\bar{\zeta} > 0$, and that in this case some fixed proportions behavior (with suitably chosen proportions) is adequate for the purpose.

For the remainder of this section, all of the assumptions of the model of the previous section are maintained, unless notice is given to the contrary. Furthermore, it will be assumed that, for every activity i, the initial performance level, $U_i(0)$, is strictly positive.

THEOREM 1. *The following three statements are equivalent*:

(i) $\bar{\zeta} > 0$.

(ii) *There exists a behavior for which the probability of survival is positive.*

(iii) *There exists a constant proportions behavior for which the probability of survival is positive.*

Proof. We first prove that (i) implies (iii). With a constant proportions behavior, the sequences $(U_i(t))$ are mutually independent, and each sequence is a random walk. Define the (constant) allocation vector $\hat{a} = (\hat{a}_i)$ by

$$\hat{a}_i = \frac{\bar{\zeta} + \xi_i}{\eta_i + \xi_i}. \tag{3.3}$$

Since $\bar{\zeta} > 0$, $\hat{a}_i > 0$. Furthermore, $\sum_i \hat{a}_i = 1$; hence, $\hat{a}$ is an allocation vector. It is easy to verify that, for every i, $EZ_i(t) = \bar{\zeta}$. Hence, for every i, the probability that $U_i(t)$ remains positive for all t is positive [7, p. 189]. Since the performance indices are mutually independent, the probability of survival is the product of these probabilities, and hence is also positive.

Since (iii) implies (ii), it remains only to show that (ii) implies (i) or, alternatively, that if $\bar{\zeta} \leqslant 0$ then there is no behavior for which the probability of survival is positive. Note first that $M(t) \leqslant \bar{U}(t)$, and therefore it suffices to show that, with probability one, $\bar{U}(t)$ eventually becomes nonpositive. This will be done using a martingale argument. Observe that, in general, $\bar{U}(t)$ will not be a random walk for arbitrary behaviors. However, it was pointed out at the end of the previous section that, for any behavior, the conditional expected value of $\bar{U}(t+1)$, given the history of the system up through date t, is equal to $\bar{U}(t) + \bar{\zeta}$; hence, if $\bar{\zeta} \leqslant 0$, then $\bar{U}(t)$ is a supermartingale.

Thus, assume that $\bar{\zeta} \leqslant 0$, and fix any behavior. Let T be the first date t such that $\bar{U}(t) \leqslant 0$; T may, in principle, be infinite, but our task is to show that, with probability one, T is finite. Let $S(t)$ be the so-called "stopped process" corresponding to $\bar{U}(t)$; i.e., $S(t)$ is equal to $\bar{U}(t)$ for $t < T$, and equal to $\bar{U}(T)$ for $t \geqslant T$. The process $S(t)$ is also a supermartingale and is bounded below. Therefore, $S(t)$ converges almost surely (a.s.).[1]

It remains to show that $S(t)$ cannot converge to any positive number, and for this it suffices to show that $\bar{U}(t)$ cannot converge. For the case $\bar{\xi} < 0$, this is implied by the following law of large numbers, which is, in turn, an immediate consequence of Theorem 40 of Freedman's paper [1].

PROPOSITION 1. *Let $\{Y_t\}$ be random variables, with Y_t $\mathscr{F}_t$-measurable, and let $\mu_t = E\{Y_t \mid \mathscr{F}_{t-1}\}$. If, for some strictly positive b,*

$$| Y_t | \leqslant b,$$

for all t, a.s., and if

$$\bar{\mu} = \lim_{t\to\infty}(1/t)(\mu_1 + \cdots + \mu_t)$$

exists a.s., then

$$\lim_{t\to\infty} (1/t)(Y_1 + \cdots + Y_t) = \bar{\mu}, \quad a.s.$$

To apply Proposition 1, take $Y_t = \bar{U}(t) - \bar{U}(t-1)$; then $\mu_t = \bar{\zeta}$. By

[1] For facts about martingales used in this proof, see, e.g., Neveu [4, Chaps. II and IV]; in particular, for this convergence theorem, see [4, Proposition II-2-9, p. 25].

assumptions (2.2c)–(2.2e) the hypotheses of the proposition are satisfied, and hence

$$\lim_{t\to\infty} \bar{U}(t)/t = \bar{\zeta}. \tag{3.4}$$

If $\bar{\zeta} < 0$, then $\bar{U}(t)$ diverges to $-\infty$.

If $\bar{\zeta} = 0$, then $\bar{U}(t)$ is a martingale with uniformly bounded increments (assumption (2.2e)) and with

$$\sum_{t=1}^{\infty} \mathrm{Var}\,[\bar{U}(t) - \bar{U}(t-1) |\, \mathscr{F}_{t-1}] = \infty$$

(assumptions (2.2c) and (2.2d)). Hence by Corollary 4.5 of [2], $\sup_t \bar{U}(t) = +\infty$ and $\inf_t \bar{U}(t) = -\infty$, a.s. Thus, $\bar{U}(t)$ diverges a.s. also in the case $\bar{\zeta} = 0$. This completes the proof of Theorem 1.

If an agent were concerned *solely* with survival, he would seek a behavior that maximized the probability of survival. We know of no simple characterization of behavior that is "optimal" in this respect, even for the special case of two activities and simple (one-step) random walks.

We close this section with some remarks about alternative definitions of survival. Survival might be defined as keeping positive some given subset of performance indices. The preceding analysis would then be applicable directly to that given subset. In particular, with the assumptions that have been made here, a positive probability of survival is always attainable with respect to any given *single* performance index.

Survival could also be defined as keeping *some* performance index positive, i.e., keeping $\max_i U_i(t)$ positive. But with the present assumptions, this, too, is possible with positive probability by allocating all effort to any one single activity.

4. Rate of Growth

For a given behavior in the controlled random walk model that we have been studying, the vector $U(t)$ of performance indices typically has no steady-state distribution, even in the limit as t increases. However, a measure of performance that may converge, as the horizon increases, is the average rate of change per unit time. For example, in the case of a constant proportions behavior, each performance index, $U_i(t)$, is a random walk, so the strong law of large numbers implies that $(1/t)[U_i(t) - U_i(0)]$ converges almost surely to $EZ_i(t)$, which is the same for all t.

In the case of a random walk, it is natural to consider the average rate of change per unit of time in conjunction with the probability of survival.

Consider for the moment a single activity, so that $U(t)$ is a scalar, and with a fixed effort allocated to that activity. Then $U(t)$ is a random walk with $EZ(t) = \zeta$, a constant. Let $S(t)$ denote the "stopped" process corresponding to $U(t)$, as defined in the previous section, but assume that $\zeta > 0$. After a sufficiently long time has elapsed, the process $S(t)$ is very likely to be either "stopped" ($S(t) \leqq 0$) or quite large. In the latter case, $S(t)/t$ will be close to ζ. If $U(t)$ is aperiodic,[2] then the probability of survival, given $U(0) = u$, goes to 1 like $1 - e^{-u}$ as u gets large.[3] Since $U(t)$ is Markovian, this implies that the probability of survival after date t, given that $U(t) = u$, goes to 1 like $1 - e^{-u}$ as u gets large. Thus, the larger $U(t)$, the larger the probability of survival beyond date t. In other words, once the hurdle of initial survival has been overcome and $U(t)$ has grown large, failure is no longer a serious threat, and the rate of growth could compete more effectively for attention.

In the multiactivity case, with a general behavior, the process $U(t)$ will not typically be a random walk. However, as we have shown in (3.4), the law of large numbers holds for the "natural weighted average performance," $\bar{U}(t)$. We show in Theorem 2(b) that the law of large numbers also holds for an individual index, if the fraction of the time that effort is allocated to that activity converges, a.s.

If, for a given behavior,

$$R_i \equiv \lim_{t\to\infty} (1/t)[U_i(t) - U_i(0)] \tag{4.1}$$

exists a.s., then we shall say that R_i is the *rate of growth* of $U_i(t)$ for activity i. As already noted, for constant proportions behaviors the rates of growth are particularly easy to determine. For a constant proportions behavior with allocation vector $a = (a_i)$, $EZ_i(t) = a_i\eta_i - (1 - a_i)\,\xi_i$, so that, by the strong law of large numbers,

$$R_i = a_i\eta_i - (1 - a_i)\,\xi_i\,, \quad \text{a.s.} \tag{4.2}$$

It is easy to verify that the constant proportions behavior that maximizes $\min_i R_i$ is exactly the one with the allocation $\hat{a}$ in Eq. (3.3) of the proof of Theorem 1 in the previous section. We shall call this the *balanced growth (constant proportions) behavior*. For this behavior, of course, $R_i = \bar{\zeta}$ for every i.

[2] $U(t)$ is called aperiodic if 1 is the greatest common divisor of the support of the distribution of $Z(t)$.

[3] We use here the fact that, for an aperiodic random walk $U(t)$ with positive mean, the conditional probability that $U(t) \leqslant 0$ for some t, given that $U(0) = x$, approaches 0 exponentially in x as x increases without limit (see Spitzer [7, pp. 217–218]).

THEOREM 2. (a) *For any behavior*

$$\lim_{t\to\infty} \frac{\bar{U}(t) - \bar{U}(0)}{t} = \bar{\zeta} \quad a.s.$$

(b) *Let* $A_i(t) = \sum_{s=0}^{t-1} a_i(s)$. *Suppose* $a_i = \lim_{t\to\infty} A_i(t)/t$ *exists almost surely; then* R_i *exists and is equal to* $a_i\eta_i - (1 - a_i)\,\xi_i \quad a.s.$

Remark. R_i and a_i are random variables. We discuss in Section 6 behaviors such that the R_i are nondegenerate random variables. Of course, with constant proportions behavior, a_i and R_i are constants. Another interesting example in which a_i and R_i are constants is given in Theorem 3 of Section 5, below.

Proof of Theorem 2. Part (a) is a restatement of (3.4). To prove part (b), apply Proposition 1 by taking

$$Y_t = U_i(t) - U_i(t-1),$$
$$\mu_t = a_i(t-1)\,\eta_i - [1 - a_i(t-1)]\,\xi_i\,.$$

5. "PUTTING OUT FIRES"

We now examine the consequences of pursuing some common administrative behaviors. In this section we consider "putting out fires," which connotes dealing with crises. In the following section we analyze "staying with a winner." An important feature of both behaviors is that *they can be pursued without any knowledge of the probability laws that govern the performance processes.* This is of considerable significance for these behaviors as examples of bounded rationality. In the context of the present model, we shall define as putting out fires a behavior such that at every date effort is allocated only to the worst performing activity or activities.

Putting out fires would appear to be a "conservative" behavior, motivated more by the desire to survive than the desire to grow rapidly, perhaps even more conservative than the balanced-growth constant proportions behavior of the preceding section. Whether putting out fires has a higher probability of survival than balanced-growth behavior is not known to us. However, putting out fires can survive if and only if balanced growth can survive, and if survival is possible then putting out fires and balanced growth have the same growth rate (see Theorem 3, below). To provide a formal definition, let

$$M(t) = \min_i U_i(t), \tag{5.1}$$

and define *putting out fires* by the following:

(a) If $U_i(t) > M(t)$, then $a_i(t) = 0$.
(b) If $U_i(t) = M(t)$ and $a_i(t-1) = 1$, then $a_i(t) = 1$. (5.2)
(c) If neither (a) nor (b) holds, then $a_i(t) = 1$ for $i =$ the smallest j such that $U_j(t) = M(t)$.

THEOREM 3. *If there are I activities, if $\bar{\zeta} > 0$, and if*

$$\begin{aligned} &\mathbf{P}\{Z_i(t+1) = 0 \mid a_i(t)\} > 0, \\ &\mathbf{P}\{Z_i(t+1) = 1 \mid a_i(t) = 1\} > 0, \\ &\mathbf{P}\{Z_i(t+1) = -1 \mid a_i(t) = 0\} > 0, \end{aligned} \tag{5.3}$$

then with putting out fires (i) each activity has almost surely the same rate of growth as does $\bar{U}(t)$, namely, $\bar{\zeta}$, (ii) the long-run average proportion of effort allocated to each activity is almost surely the same as with balanced-growth (constant proportions) behavior, and (iii) survival is possible with positive probability.

The proof of Theorem 3 is based on the following proposition, which is of independent interest. Consider the Markov chain[4]

$$C_I(t) = [m(t), V_1(t), ..., V_I(t)],$$

where $m(t)$ is the activity to which effort is allocated at time t and

$$V_i(t) = U_i(t) - M(t).$$

PROPOSITION 2. *$\bar{\zeta} > 0$ implies that the Markov chain $C_I(t)$ is positive recurrent.*

Remark. This result shows how putting-out-fires behavior differs from balanced growth even though, if $\bar{\zeta} > 0$, all activities grow at the same asymptotic rate when either behavior is adopted. Under balanced growth the various activities tend to spread out. The difference, $D_{ij}(t) = U_i(t) - U_j(t)$, is a random walk without drift. Thus the distribution of $D_{ij}(t)$ tends to no limit, and the expected time of $D_{ij}(t)$'s first return to zero is always infinite. In contrast, if putting out fires is followed, activities tend to stay close together. Since $D_{ij}(t) = V_i(t) - V_j(t)$, Proposition 2 implies that $D_{ij}(t)$ has an asymptotic distribution and that the expected time to return to zero is finite, whatever the current value of $D_{ij}(t)$. The proof of Proposition 2 is given after that of Theorem 3.

[4] It is to be understood that the state space of any chain considered in the connection with putting out fires is to include only states that are consistent with this behavior.

Proof of Theorem 3. Since $C_I(t)$ is positive recurrent, the long-run relative frequency of the events $m(t) = i$ converges almost surely to the invariant probability of that event. Let these probabilities be a_i. By Theorem 2(b), $\lim_{t\to\infty}(1/t)[U_i(t) - U_i(0)]$ converges almost surely to

$$R_i = a_i\eta_i - (1 - a_i)\,\xi_i\,. \tag{5.4}$$

Suppose $R_i \neq R_j$ for some i and j; consider

$$V_i(t) = U_i(t) - M(t) \geqslant |\, U_i(t) - U_j(t)|\,.$$

If $R_i > R_j$, this last quantity diverges to $+\infty$ a.s., which contradicts Proposition 2. Thus all activities grow at the same rate, which must equal $\bar{\zeta}$, which proves (i). To prove (ii), solve (5.4) for a_i, with $R_i = \bar{\zeta}$.

We now prove (iii). We have shown in (i) that $\lim_{t\to\infty}(U_i(t))/t = \bar{\zeta} > 0$ a.s., so that

$$\lim_{t\to\infty} U_i(t) = +\infty \quad \text{a.s.,} \quad i = 1, 2,\ldots, I.$$

Hence

$$\lim_{t\to\infty} M(t) = +\infty \quad \text{a.s.} \tag{5.5}$$

Denote $(u_1,\ldots, u_I)$ by u, and let $T(u)$ be the first t such that $M(t) \leqq 0$ given that $U(0) = u$ and $M(0) > 0$. We must show that $\mathbf{P}\{T(u) = \infty\} > 0$.

Let H be the set of u for which $(\max_i u_i - \min_i u_i) \leqq 1$ and $\min_i u_i > 0$. One can verify that for the Markov chain $[m(t), U(t)]$ the set H' of states for which $U(t)$ is in H can be reached (with positive probability) from any state, and that this can be done while keeping $\min_i u_i$ strictly positive if the point of departure is strictly positive. Furthermore, from any state in H' one can reach any other state in H' without leaving H'. (This shows, incidentally, that the chains $[m(t), U(t)]$ and $C_I(t)$ each have a single class). Therefore, if there is some u^* in H from which the probability of survival is positive, then the same is true for every u for which $\min_i u_i$ is positive. To demonstrate the existence of such a u^* in H, suppose to the contrary that for every u in H the probability of survival from u were zero, i.e., that

$$\mathbf{P}\{T(u) < \infty \mid \mathscr{F}_0 \cdot U(0) = u\} = 1.$$

Since the chain $C_I(t)$ is positive recurrent (by Proposition 2), it would follow that, starting from *any* u, and given any t, there would be a finite $T' \geqq t$ for which $U(T')$ would be in H, and therefore there would be a subsequent finite $T'' \geqq T'$ with $M(T'') \leqq 0$ In other words, starting from any u the probability of survival would be zero, which would imply in turn that $\liminf_t M(t)$ would be $\leqq 0$, contradicting (5.5).

Proof of Proposition 2. The proof is by induction. Clearly the proposition is true for $I = 1$. The induction step—that if it is true for $I = 1, \ldots, J - 1$, then it is true for $I = J$—is accomplished in three lemmas. The first uses the fact, established above, that if Proposition 2 is true for $I = J - 1$ then Theorem 3 is true for $I = J - 1$.

LEMMA 1. *Suppose that $\mathscr{K}$ is any proper subset of $\mathscr{J} = \{1, 2, \ldots, J\}$, that $\mathscr{K}'$ is the complement of $\mathscr{K}$ in $\mathscr{J}$, and that putting out fires is practiced on the activities in $\mathscr{K}$ while no effort is allocated to those in $\mathscr{K}'$. Then there is a (nonrandom) T such that*

$$E \min_{k \in \mathscr{K}} U_k(t) \geqslant \min_{k \in \mathscr{K}} U_k(0) + 1 \quad \textit{for all} \quad t \geqslant T, \tag{5.6a}$$

$$E \max_{k \in \mathscr{K}'} U_k(t) \leqslant \max_{k \in \mathscr{K}'} U_k(0) - 1 \quad \textit{for all} \quad t \geqslant T. \tag{5.6b}$$

T can be chosen so that (5.6) *holds for any $\mathscr{K}$ properly contained in $\mathscr{J}$.*

Proof. If we consider the activities in $\mathscr{K}$ alone, then, in an obvious notation, the rate of growth of $\bar{U}_{\mathscr{K}}(t)$, the natural weighted-average process, is

$$\bar{\zeta}_{\mathscr{K}} = \left(1 - \sum_{k \in \mathscr{K}} \frac{\xi_k}{\eta_k + \xi_k}\right)\left(\sum_{k \in \mathscr{K}} \frac{1}{\eta_k + \xi_k}\right)^{-1}. \tag{5.7}$$

Comparing (5.7) and (2.6), we see that $\bar{\zeta}_{\mathscr{K}} > 0$ whenever $\bar{\zeta} > 0$. Define

$$C_{\mathscr{K}}^{\mathscr{J}}(t) = [m_{\mathscr{K}}(t), \{V_{\mathscr{K}_j}(t)\}_{j \in \mathscr{K}}],$$

where $m_{\mathscr{K}}(t)$ is the activity in $\mathscr{K}$ to which effort is allocated at date t, and

$$V_{\mathscr{K}_j}(t) = U_j(t) - M_{\mathscr{K}}(t), \quad j \text{ in } \mathscr{K},$$

$$M_{\mathscr{K}}(t) = \min_{j \in \mathscr{K}} U_j(t).$$

The process $\{C_{\mathscr{K}}^{\mathscr{J}}(t)\}$ is a Markov chain and, by the induction hypothesis, is positive recurrent. The long-run relative frequency of the events $m_{\mathscr{K}}(t) = k$ converges a.s. to the corresponding invariant probabilities. By the induction hypothesis and (i) of Theorem 3, for every k in $\mathscr{K}$, $U_k(t)/t$ converges to $\bar{\zeta}_{\mathscr{K}}$; hence,

$$\lim_{t \to \infty} \left(\frac{1}{t}\right) \min_{k \in \mathscr{K}} U_k(t) = \bar{\zeta}_{\mathscr{K}}, \text{ a.s.}$$

Recalling that the increments of $U_k(t)$ are uniformly bounded, we conclude from the last limit and the Lebesgue dominated convergence theorem that

$$\lim_{t\to\infty} E\left(\frac{1}{t}\right) \min_{k\in\mathscr{K}} U_k(t) = \bar{\zeta}_{\mathscr{K}} ,$$

so that

$$E \min_{k\in\mathscr{K}} U_k(t) \to +\infty.$$

Thus there exists $T_{\mathscr{K}}$ such that $t > T_{\mathscr{K}}$ implies

$$E \min_{k\in\mathscr{K}} U_k(t) > \min_{k\in\mathscr{K}} U_k(0) + 1.$$

If no effort is allocated to the activities in $\mathscr{K}'$, then the law of large numbers implies that $U_k(t)/t$ converges a.s. to $(-\xi_k)$ for all k in $\mathscr{K}'$, and hence

$$\lim_{t\to\infty} \left(\frac{1}{t}\right) \max_{k\in\mathscr{K}'} U_k(t) = -\min_{k\in\mathscr{K}'} \xi_k .$$

Appealing again to Lebesgue, we conclude that

$$E \max_{k\in\mathscr{K}'} U_k(t) \to -\infty,$$

so that there is a $T_{\mathscr{K}}'$ such that $t \geqslant T_{\mathscr{K}}'$ implies

$$E \max_{k\in\mathscr{K}'} U_k(t) < \max_{k\in\mathscr{K}'} U_k(0) - 1.$$

Letting $T = \max_{\mathscr{K}}(\max(T_{\mathscr{K}} , T_{\mathscr{K}}'))$ as $\mathscr{K}$ ranges over all proper subsets of $\mathscr{J}$ completes the proof.

LEMMA 2. *Let $D(t) = \max_i V_i(t)$, and $G = 2JT(b + 1)$; if $D(0) > G$, then $ED(T) < D(0) - 2$, where T is as in Lemma* 1.

Proof. Recall that b is the uniform bound on the $Z_i(t)$'s. Let ρ be a permutation of $\{1,..., J\}$ such that

$$U_{\rho(1)}(0) \geqslant U_{\rho(2)}(0) \geqslant \cdots \geqslant U_{\rho(J)}(0), \tag{5.8}$$

and let $\Delta_i(t) = U_{\rho(i-1)}(t) - U_{\rho(i)}(t)$. Then (5.8) implies $\Delta_i(0) \geqslant 0$. Since $D(0) = \sum_{i=2}^{J} \Delta_i(0)$, $D(0) > G$ implies $\Delta_j(0) > 2T(b + 1)$ for some j. For this j, $\Delta_j(t) > 0$ for $t \leqslant T$. Thus if $\mathscr{K} = \{k \mid \rho^{-1}(k) \geqslant j\}$, $\mathscr{K}' = \{k \mid \rho^{-1}(k) < j\}$, and effort is allocated among activities 1,..., J according to putting-out-fires behavior for T periods, then all effort is devoted to the activities in $\mathscr{K}$ and none to those in $\mathscr{K}'$. Since

$$D(T) = \max_{k \in \mathscr{K}'} U_k(T) - \min_{k \in \mathscr{K}} U_k(T),$$

the conclusion follows immediately from Lemma 1.

LEMMA 3. *Suppose $D(0) > G$; let T^* be the first t such that $D(T^*) \leqslant G$. Then $ET^* < \infty$.*

Proof. Let $D = D(0)$ and consider the random variables

$$X_n \equiv D[nT] - D[(n-1)\,T].$$

If we let $\mathscr{G}_n = \mathscr{F}_{nT}$, then $\mathscr{G}_n$ is an increasing sequence of sigma fields and X_n is $\mathscr{G}_n$-measurable. Furthermore, if $Y_n = \sum_1^n X_m$, then $D(nT) = D + Y_n$. Let $C = G - D$. If $Y_n \leqslant C$, then $D(nT) \leqslant G$. Define N^* as the first n such that $Y_n \leqslant C$. Since $N^*T \geqslant T^*$, to prove $ET^* < \infty$ we need only show

$$EN^* < \infty. \tag{5.9}$$

The random variables X_n and Y_n have the following properties:

$$| X_n | < B, \tag{5.10}$$

where $B = 2JTb$, and Lemma 2 implies that

$$E[X_n \mid \mathscr{G}_{n-1} \,.\, Y_{n-1} > C] < -2. \tag{5.11}$$

To prove that (5.10) and (5.11) imply (5.9), we use an inequality due to Freedman [1]. Let

$$W_n = (X_n + B)/2B,$$

$$S_n = \sum_1^n W_m = \frac{Y_n + nB}{2B},$$

$$R_n = E[W_n \mid \mathscr{G}_{n-1}].$$

Suppose

$$n > -C; \tag{5.12}$$

if $N^* > n$, then $Y_m > C$ and $R_m < (-2 + B)/2B$ for $m = 1,..., n$, so that $S_n > (C + nB)/2B \equiv a_n$ and $\sum_1^n R_m < n(-2 + B)/2B \equiv b_n$. Since (5.12) implies $a_n > b_n$, we may use (4.b) of Freedman [1, p. 911] to conclude that

$$\mathbf{P}\{N^* > n\} \leqslant \exp[-(a_n - b_n)^2/2a_n]. \tag{5.13}$$

Note that

$$(a_n - b_n)^2/2a_n = (C + 2n)^2/4B(C + nB) \geqslant n^2/4nB^2 = n/4B^2,$$

because of (5.12). Thus (5.13) may be replaced by

$$\mathbf{P}\{N^* > n\} \leqslant \exp[-n/4B^2].$$

This implies $EN^* < \infty$, which completes the proof of Lemma 3.

Proposition 2 is an immediate consequence of Lemma 3, which states that, if $\{C_I(t)\}$ ever leaves the finite set of states such that $D(t) \leqslant G$, then the expected time to return is finite. This implies $\{C_I(t)\}$ is positive recurrent [3, pp. 98, 99, 135, 143] (recall that it has a single class).

If there are only two activities, the above arguments can be straightforwardly adapted to prove that the chain $\{C_I(t)\}$ is positive recurrent for all values of $\bar{\zeta}$. It follows that, for putting out fires with two activities, in the long run both activities grow—or decline—at the same average rate, namely, $\bar{\zeta}$.

The conclusions of Theorem 3 do not necessarily hold if $\bar{\zeta} < 0$. Suppose that there are three activities, and let $\bar{\zeta}_{12}$ be the growth rate of the natural weighted average of the first two activities considered by themselves. If all effort were allocated to these two activities and allocated between them by putting-out-fires behavior, then each of the first two activities would have asymptotic rates of growth equal to $\bar{\zeta}_{12}$. Suppose that $\bar{\zeta}_{12} < -\xi_3 < 0$. Then with positive probability $U_1(t)$ and $U_2(t)$ could sink below $U_3(t)$ and remain there forever, even though effort were concentrated entirely on activities 1 and 2. Were this to happen, $U_1(t)$ and $U_2(t)$ would decline at the rate $\bar{\zeta}_{12}$, while $U_3(t)$ would decline at the (distinct) rate $-\xi_3$.

Theorems 1 and 3 may be combined to yield the following simple result.

THEOREM 4. *If* (5.3) *is satisfied, then survival is possible with positive probability if and only if survival is possible with putting-out-fires behavior.*

6. "Staying with a Winner"

A behavior that is diametrically different from putting out fires is one that allocates all effort at any date to the best performing activity or activities. Following some colloquial practice, we shall call such behavior *staying with a winner*; its defining property is

$$a_i(t) = 0 \text{ for all } i \text{ such that } U_i(t) < M^*(t) \equiv \max_j U_j(t). \tag{6.1}$$

The asymptotic performance of staying with a winner is analyzed in Theorem 5, below; a rough description is presented at this point. Under staying-with-a-winner behavior, the allocation of effort will eventually concentrate on a single activity, but which activity that will be cannot be predicted with certainty in advance. In other words, for some T and J, all effort will be allocated to activity J from date T on, but T and J are random variables. The rate of growth of $M^*(t)$ will be η_J, that is, the mean of $Z_J(t)$ if all effort is allocated to activity J. The rate of growth of each activity i other than J will be $-\xi_i$; in particular, the performance levels of all of these other activities will eventually become negative.

We may contrast staying with a winner with the constant proportions behavior that allocates all effort to the (or an) activity for which η_i is maximum. The rate of growth of $M^*(t)$ for this latter activity is equal to $\eta^* = \max_i \eta_i$, whereas the rate of growth of $M^*(t)$ for staying with a winner (which is η_J, as already noted) cannot exceed η^* and will be strictly less than η^* with positive probability. If the manager were trying by experimentation to guess the activity with the highest η_i, then he would face what is known as a two-armed-bandit problem. It is interesting to note that the asymptotic behavior of the optimal strategy in some two-armed-bandit problems is identical to that of staying with a winner (see Rothschild [6]).

We turn now to a formal analysis of staying with a winner. We shall consider any (finite) number of activities. At any date at which only one activity attains the maximum performance $M^*(t)$, all of the effort is allocated to that activity. At any date at which more than one activity attains a performance equal to $M^*(t)$, all of the effort is allocated to the activity in that subset with the lowest index i. (Any well-defined rule for breaking ties would be acceptable, as far as the following analysis goes.)

We shall rule out cases in which an activity has no possibility of "catching up" once it falls behind. For this purpose, we shall say that an *activity i can compete* if, for any date t, any allocation at that date, and any other activity j, there is a positive probability that $Z_i(t) - Z_j(t)$ is strictly positive.

THEOREM 5. *If every activity can compete and is aperiodic, then the following statements are true with probability one, for staying with a winner:* (*a*) *There exists an activity J and a* (*finite*) *date T such that, at all dates subsequent to T, all effort is allocated to activity J; note that J and T are random variables.* (*b*) *For every activity i, there is a positive probability p_i that $J = i$.* (*c*) *If $i = J$, then the rate of growth of $U_i(t)$ is η_i, and there is a positive conditional probability that $U_i(t)$ never becomes negative; if $i \neq J$, then the rate of growth of $U_i(t)$ is $-\xi_i$, and the conditional probability that*

$U_i(t)$ eventually becomes and stays negative is unity. (d) The rate of growth of $M^*(t)$ is η_J, a random variable, and the probability is positive that $M^*(t)$ never becomes negative.

It follows that (e) the expected rate of growth of $M^*(t)$, which is $\sum_i p_i \eta_i$, is strictly less than $\max_i \eta_i$, *if the η_i are not all equal; indeed, $\eta_J \leqslant \eta^*$ a.s.*

Proof. Consider any given vector $U(0)$ of initial performance levels and consider any activity i. Since i can compete, for any number d there exists some date, say s, such that, with positive probability, the performance of activity i at date s exceeds that of any other activity by at least d. One can choose d sufficiently large so that, for every $j \neq i$, the conditional probability that $U_i(t) \leqslant U_j(t)$ for some $t \geqslant s$, given $U_i(s) \geqslant U_j(s) + d$, does not exceed $\frac{1}{2}(I - 1)$, where I is the number of activities.[5] Therefore, for such a d, the conditional probability that $U_i(t) \leqslant U_j(t)$ for some $j \neq i$ and some $t \geqslant s$, given $U_i(s) \geqslant U_j(s) + d$ for all $j \neq i$, does not exceed $\frac{1}{2}$. Let $m(t)$ denote the activity to which all effort is allocated at date t. It follows that, for any given vector $U(0)$ of initial performance levels, and any activity i, there is some date s such that the probability is positive that $m(t) = i$ for all $t \geqslant s$.

Suppose now that $m(0) = i$, and define

$$Q_i[U(0)] = \mathbf{P}\{m(t) = i \text{ for all } t \mid U(0), m(0) = i\},$$

$$q_i = \inf\{Q_i[U(0)]: m(0) = i, \text{ all } U(0) \text{ such that } U_i(0) = M^*(0)\}.$$

An examination of the preceding argument shows that $q_i > 0$; let $q = \min_i q_i$. Since $\{m(t), U(t)\}$ is Markovian, it follows that if $m(t) = i$ then the conditional probability that $m(n) = i$ for all $n \geqslant t$, given $[m(t), U(t)]$, is at least q. Consider any date t such that either $t = 0$ or $m(t - 1) \neq m(t)$. Call (t_k) the sequence of such dates t; $t_0 \equiv 0$. Either $m(n) = m(t_k)$ for all $n \geqslant t_k$, or there is some first (finite) t' after t_k such that $m(t') \neq m(t_k)$; the first case will occur with probability at least q. Since $q > 0$, the probability is one that the first case will eventually occur after some date t_k. This completes the proof of parts (a) and (b). The remainder of the proof is straightforward.

References

1. D. Freedman, Another note on the Borel–Cantelli lemma and the strong law, *Ann. Probability* **1** (1973), 910–925.
2. D. Freedman, On tail probabilities for martingales. *Ann. Probability* **3** (1975), 100–118.

[5] See footnote 3, above.

3. J. G. KEMENY, J. L. SNELL, AND A. W. KNAPP, "Denumerable Markov Chains," Van Nostrand, Princeton, NJ, 1966.
4. J. NEVEU, "Martingales à Temps Discret," Masson, Paris, 1972.
5. R. RADNER, Aspiration, bounded rationality and control, in preparation.
6. M. ROTHSCHILD, A two-armed-bandit theory of market pricing, *J. Econ. Theory* **9** (1974), 185–202.
7. F. SPITZER, "Principles of Random Walk," Van Nostrand, Princeton, NJ, 1964.