The Annals of Applied Probability
1994, Vol. 4, No. 2, 255–286

DYNAMIC ALLOCATION PROBLEMS IN CONTINUOUS TIME

BY NICOLE EL KAROUI AND IOANNIS KARATZAS[1]

Université Pierre et Marie Curie and Columbia University

We present an approach to the general, non-Markovian dynamic allocation (or multiarmed bandit) problem, formulated in continuous time as a problem of *stochastic control for multiparameter processes* in the manner of Mandelbaum. This approach is based on a direct, martingale study of auxiliary questions in optimal stopping. Using a methodology similar to that of Whittle and relying on simple time-change arguments, we construct Gittins-index-type strategies, verify their optimality, provide explicit expressions for the values of dynamic allocation and associated optimal stopping problems, explore interesting dualities and derive various characterizations of Gittins indices. This paper extends results of our recent work on discrete-parameter dynamic allocation to the continuous time setup; it can be read independently of that work.

1. Introduction. We study in this paper the general, non-Markovian version of the dynamic allocation (or multiarmed bandit) problem in continuous time. There are d independent projects ("investigations," "arms") among which effort has to be allocated. By engaging in any one of these projects, a random reward is accrued which depends on time, on the history of the project and on the rate at which effort is being allocated to the project (a rate of zero corresponds to not engaging in that project at all, a rate of 1 to engaging in that project only). The objective is to find an "allocation strategy" that maximizes total expected reward over an infinite horizon, discounted at the rate $\alpha > 0$.

The *discrete-time* version of the problem, in which only one project can be engaged at a time, is by now well understood. Beginning with the seminal work of Gittins and Jones (1974) and culminating with the papers and books by Gittins (1979, 1989) and Whittle (1980, 1982), it is quite well known that, in a Markovian framework, one can associate to each project a deterministic function of its state, now called the "Gittins index," and such that the following strategy is optimal: *at any given time, engage a project with maximal current Gittins index*. These insights were then extended to the general, non-Markovian case by Varaiya, Walrand and Buyukkoc (1985), by

Received May 1993; revised November 1993.

[1]Research supported by U.S. Army Research Office Grant DAAL 03-91-G-0170. Part of this work was carried out during the Fall of 1992, while the author was visiting the Departments of Statistics and Finance at the Wharton School, University of Pennsylvania. Their hospitality is gratefully acknowledged.

AMS 1991 *subject classifications*. Primary 93E20, 60G60; secondary 60G40, 90B85, 62L10.

Key words and phrases. Multiarmed bandit problem, optimal stopping, stochastic control, multiparameter random time-change, Gittins index, martingales, Brownian local time.

Mandelbaum (1986), who was the first to cast the problem as one of controlling a process with multidimensional time-parameter, and recently by El Karoui and Karatzas (1993) and Cairoli and Dalang (1993). A random "Gittins index sequence," representing the "equitable surrender value" of the particular project at any given time, replaces the index function, and the same rule as above is again optimal.

Attempts to extend these insights and results to their *continuous-time* framework had to face up to the task of providing a proper definition of allocation strategy in the continuum. Here again, the best and most insightful formulation comes by treating the problem as one of stochastic control for processes with multidimensional parameter $\underset{\sim}{s} = (s_1, s_2, \dots, s_d) \in [0, \infty)^d$, where each s_i represents the total amount of time the particular project has been engaged up to the total (calendar) time $t = \sum_{i=1}^d s_i$. This was the formulation offered by Mandelbaum (1987), who treated the general, non-Markovian problem in this framework. Related results in the Markovian case were obtained by Karatzas (1984), Eplett (1986) and Menaldi and Robin (1990).

We present in this paper an approach to the problem in its general, non-Markovian version, by combining the powerful formulation of Mandelbaum (1987) with the methodology of Whittle (1980). This approach is based on a direct, "martingale-type" study of auxiliary questions in optimal stopping; in terms of these, *it computes fairly explicitly the value of the bandit problem* in the manner of Whittle (1980). The approach also offers a constructive proof for the optimality of strategies that always engage in projects with maximal current Gittins index, suggests interesting dualities and leads to various characterizations of Gittins index processes. At the same time it avoids boundedness assumptions on the reward processes, removes some of the strong (and hard to verify) conditions imposed in Mandelbaum (1987) and does not rely on discretizations. The arguments used in the proofs are novel and quite straightforward. A key new fact is the martingale property of the process in (3.9), which is introduced for the first time in the study of these problems. When specializing the results in this paper to the Markovian framework, one obtains those of Karatzas (1984) and Menaldi and Robin (1990). Furthermore, the paper extends to the continuous-time framework the results of our recent work [El Karoui and Karatzas (1993)]; the two papers can be read independently of each other.

The primary results can be described, very roughly, as follows. In terms of the auxiliary questions of optimal stopping, one *constructs* the Gittins index processes $M_i(t)$ (with the same interpretation as "equitable surrender value for the ith project at time t"), the lower envelopes $\underline{M}_i(t) = \inf_{0 \le u \le t} M_i(u)$, $t \ge 0$, and their inverses $\underline{M}_i^{-1}(\cdot)$, $i = 1, \dots, d$. Then $\mathbf{M}(t) = (\sum_{i=1}^d \underline{M}_i^{-1})^{-1}(t)$ can be interpreted as "equitable surrender value for the entire collection of d projects" at time t, and any *index-type strategy* [which engages project(s) with maximal $\underline{M}_i(t)$, at all times $t \ge 0$] *is optimal* (Theorem 8.1). We construct such a strategy $\underset{\sim}{I} = (I_1, \dots, I_d)$ in Proposition 7.3 and show that, under certain conditions, it is given by $I_j(t) = \underline{M}_j^{-1}(\mathbf{M}(t))$, where $I_j(t)$ denotes the total amount of time spent on the jth project by $\underset{\sim}{I}(\cdot)$ during $[0, t]$. Index-type strategies need not be unique. In contrast to the discrete-time situation, they need not be *pure*, that is, they may engage several projects at the same time (Remark 8.3), but that is not necessary in the case of diffusions (Section 9). However, the *value* of the multiarmed bandit problem (supremum of achievable expected reward) is always the same over pure as over general allocation strategies, and is given by

$$\Phi = \int_0^\infty \left(1 - E\left[\exp\left\{-\alpha \sum_{i=1}^d M_i^{-1}(u)\right\}\right]\right) du = \int_0^\infty \alpha e^{-\alpha t} E\mathbf{M}(t)\, dt.$$

The paper is organized as follows. We start in Section 2 with a detailed study of a one-parameter family of optimal stopping problems, based on martingale techniques. In terms of this analysis, we are then able to introduce the so-called Gittins index process in Section 3, to study its properties and to offer interesting alternative representations for it. Section 4 has a "motivational" function. It poses two stochastic control problems related to the optimal stopping problem, one "equivalent" and one "dual," which are then studied in a way that presages *questions of dynamic allocation*. These questions are formulated in Section 5, both in their "original" form and in a "parametric" one suggested by the optimal stopping problem. Section 6 offers an explicit expression (6.9) for the value of the dynamic allocation problem; a key observation here is the reduction property (6.3). Section 7 introduces a special class of "index-type" allocation strategies for the dynamic allocation problem, as well as for a dual minimization problem first formulated by Mandelbaum (1986, 1987). Section 8 proves the optimality of index-type strategies for the dynamic allocation problem, as well as for a dual minimization problem first formulated by Mandelbaum (1986, 1987). We discuss in Section 9 the case of two independent "Brownian" arms with identical reward structures, for which very explicit computations of the optimal policy and value are possible in terms of Brownian local time. All proofs are collected in the Appendix.

2. A family of optimal stopping problems. Consider a complete probability space $(\Omega, \mathscr{F}, P)$ equipped with a quasi-left-continuous filtration $\{\mathscr{F}(t)\}_{0 \le t \le \infty}$ that satisfies the usual conditions and $\mathscr{F}(0) = \{\varnothing, \Omega\}$, $\mathscr{F}(\infty) = \sigma(\bigcup_{0 \le t < \infty} \mathscr{F}(t))$, mod P. We shall denote by $\mathscr{S}$ the class of stopping times of this filtration (with values in $[0, \infty]$), and for any given $u \in \mathscr{S}$, we shall denote by $\mathscr{S}(u)$ the class of stopping times $\sigma \in \mathscr{S}$ with $P[u \le \sigma] = 1$.

On this space, let $H = \{h(t), \mathscr{F}(t);\ 0 \le t < \infty\}$ be a nonnegative, progressively measurable process. It will be assumed that $E\int_0^\infty e^{-\alpha t} h(t)\, dt < \infty$ and that, for any given $t \in \mathscr{S}$,

$$E\left[\int_t^\tau e^{-\alpha s} h(s)\, ds \middle| \mathscr{F}(t)\right] > 0 \quad \text{a.s.,} \quad \forall\, \tau \in \mathscr{S}(t) \text{ such that } P[\tau > t] = 1.$$

Here $\alpha \in (0, \infty)$, the *discount* rate for the problem of (2.1), is a given constant.

We can then introduce the *family of optimal stopping problems*

$$V(t;m) := \operatorname*{ess\,sup}_{\tau\in\mathscr{S}(t)} E\left[\int_t^\tau e^{-\alpha(u-t)}h(u)\,du + me^{-\alpha(\tau-t)}\middle|\mathscr{F}(t)\right], \quad 0 \le t < \infty, \tag{2.1}$$

with $V(\infty; m) \equiv m$, indexed by the parameter $m \in [0,\infty)$. This latter has an obvious interpretation as "reward-upon-stopping," while $h(t)$ represents the "instantaneous reward for continuation" at time t; the effects of both these factors are discounted at the rate α. From standard theory on the optimal stopping problem [e.g., Fakeev (1970, 1971), Bismut and Skalli (1977) and El Karoui (1981)], we know that

$$Z(t;m) := e^{-\alpha t}V(t;m) + \int_0^t e^{-\alpha s}h(s)\,ds, \qquad 0 \le t \le \infty, \tag{2.2}$$

is the *Snell envelope* of (i.e., the smallest supermartingale with RCLL paths [right-continuous, with limits from the left], that dominates) the process

$$Y(t;m) := me^{-\alpha t} + \int_0^t e^{-\alpha s}h(s)\,ds, \qquad 0 \le t \le \infty; \tag{2.3}$$

$$\begin{aligned} \sigma_t(m) \equiv \sigma(t;m) &:= \inf\{\theta \ge t \mid Z(\theta;m) = Y(t;m)\} \\ &= \inf\{\theta \ge t \mid V(\theta;m) = m\} \end{aligned} \tag{2.4}$$

is an optimal stopping time for the problem of (2.2);

$$\{Z(\theta \wedge \sigma_t(m); m), \mathscr{F}(\theta);\, t \le \theta \le \infty\} \text{ is a martingale;} \tag{2.5}$$

and the *dynamic programming equation* (2.6) holds for any $\tau \in \mathscr{S}$ and $\sigma \in \mathscr{S}(\tau)$:

$$\begin{aligned} &e^{-\alpha\tau}V(\tau;m) \\ &\quad = \max\left[me^{-\alpha\tau}, E\left\{V(\sigma;m)e^{-\alpha\sigma} + \int_\tau^\sigma e^{-\alpha s}h(s)\,ds\middle|\mathscr{F}(\tau)\right\}\right] \quad \text{a.s.} \end{aligned} \tag{2.6}$$

2.1. REMARK. The processes $Y(\cdot;m)$ and $Z(\cdot;m)$ of (2.3) and (2.2) are nonnegative, agree at $t = \infty$, and $Y(\cdot;m)$ has continuous paths with $E(\sup_{0\le t\le\infty} Y(t;m)) < \infty$. Thus $Z(\cdot;m)$ is *of class D, regular, and hence quasi-left-continuous*, thanks to the quasi-left-continuity of the filtration [Dellacherie (1972), pages 85 and 119; Bismut and Skalli (1977)].

Based on this theory, we can establish the following properties for the processes of (2.1) and (2.4); the proofs are collected in the Appendix. In what follows we consider appropriate measurable versions of the mappings $(t, m, \omega) \mapsto \sigma_t(m;\omega)$, $(t, m, \omega) \mapsto V(t; m, \omega)$ and of their (indicated) derivatives.

2.2. LEMMA. *For every fixed $t \in [0,\infty)$, the mapping $m \mapsto \sigma_t(m)$ is decreasing and right-continuous with $\sigma_t(0) = \infty$ almost surely.*

2.3. LEMMA. *For every fixed* $t \in [0, \infty]$, *the mapping* $m \mapsto V(t; m)$ *is convex and increasing, with* $\lim_{m\to\infty}[V(t;m) - m] = 0$ *and right-hand derivative*

$$\frac{\partial^+}{\partial m} V(t;m) := \lim_{\delta \downarrow 0} \frac{1}{\delta}[V(t; m+\delta) - V(t;m)] = E\left[e^{-\alpha(\sigma_t(m)-t)} | \mathscr{F}(t)\right]. \tag{2.7}$$

2.4. LEMMA. *For every fixed* $m \in [0, \infty)$, *the mapping* $t \mapsto \sigma_t(m)$ *is a.s. increasing and quasi-left-continuous, and* $t \mapsto e^{-\alpha t}(\partial^+/\partial m)\, V(t;m)$ *defines a quasi-left-continuous supermartingale.*

3. Gittins index processes and their lower envelopes. Let us introduce for each fixed $t \in [0, \infty)$, the *decreasing*, *right-continuous* inverse

$$\begin{aligned} \underline{M}(t, \theta) &:= \sup\{m \geq 0 \mid \sigma(t;m) > \theta\} \\ &= \inf\{m \geq 0 \mid \sigma(t;m) \leq \theta\}, \qquad t \leq \theta < \infty, \end{aligned} \tag{3.1}$$

of the mapping $m \mapsto \sigma_t(m)$ of Lemma 2.2. The process $\underline{M}(t, \cdot)$ satisfies

$$\sigma(t;m) > \theta \;\Leftrightarrow\; V(s;m) > m \quad \forall\, s \in [t, \theta] \;\Leftrightarrow\; m < \underline{M}(t, \theta) \tag{3.2}$$

for all $m \geq 0$, $\theta \geq t$, and is adapted to $\{\mathscr{F}(\theta)\}_{t \leq \theta < \infty}$. In particular, we have $\{m > 0 \mid V(t;m) = m\} = [M(t), \infty) \bmod P$, where

$$M(t) := \underline{M}(t,t) = \lim_{\theta \downarrow t} \underline{M}(t, \theta), \qquad 0 \leq t < \infty. \tag{3.3}$$

This $\mathscr{F}(t)$-measurable random variable satisfies $M(t) > 0$ a.s.,

$$\begin{aligned} M(t) &= \operatorname{ess\,inf}\{\xi \in P\mathscr{F}(t) \mid V(t;\xi) = \xi \text{ a.s.}\} \\ &= \operatorname{ess\,sup}\{\xi \in P\mathscr{F}(t) \mid V(t;\xi) > \xi \text{ a.s.}\}, \end{aligned} \tag{3.4)(i}$$

$$\begin{aligned} M(t) = \operatorname{ess\,inf}\Bigg\{ \xi \in P\mathscr{F}(t) \Bigg| &\operatorname{ess\,sup}_{\tau \in \mathscr{S}(t)} \\ &\times E\left[\int_t^\tau e^{-\alpha s}(h(s) - \alpha\xi)\, ds \middle| \mathscr{F}(t)\right] = 0 \text{ a.s.}\Bigg\} \end{aligned} \tag{3.4)(ii}$$

and admits the interpretations of Whittle (1980) [see also Weber (1992)] as "the smallest value of the reward-upon-stopping m that makes immediate stopping optimal (equitable surrender value) at time t"; here $P\mathscr{F}(t)$ is the class of all positive, $\mathscr{F}(t)$-measurable random variables.

By analogy with Whittle (1980, 1982), Varaiya, Walrand and Buyukkoc (1985) and Mandelbaum (1986, 1987), we shall call $M(t)$ of (3.4) the *Gittins index at time* t, and the progressively measurable process $\{M(t), \mathscr{F}(t); 0 \leq t < \infty\}$ of (3.3) the *Gittins index process.*

3.1. LEMMA. *For each fixed* $t \in [0, \infty)$, *the process* $\theta \mapsto \underline{M}(t, \theta)$ *of* (3.1) *is the lower envelope* $\underline{M}(t, \theta) = \inf_{t \leq u \leq \theta} M(u)$ *of the Gittins index process,* *namely*,

$$(3.5)\quad m < \underline{M}(t,\theta) \quad \Leftrightarrow \quad m < M(u) \quad \forall\, u \in [t,\theta], 0 \le t < \theta < \infty.$$

The equivalences of (3.2) allow us to represent the value $V(t;m)$ of the optimal stopping problem (2.1), in terms of the lower envelope $\underline{M}(t,\theta)$ of the Gittins index process.

3.2. PROPOSITION. *For every $t \in [0,\infty)$, $m \in [0,\infty)$, we have the a.s. representations*

$$(3.6)\quad E\left[\int_t^{\sigma_t(m)} e^{-\alpha\theta} h(\theta)\, d\theta \middle| \mathscr{F}(t)\right] = E\left[\int_t^{\sigma_t(m)} \alpha e^{-\alpha\theta} \underline{M}(t,\theta)\, d\theta \middle| \mathscr{F}(t)\right],$$

$$(3.7)\quad \begin{aligned} e^{-\alpha t} V(t;0) &= E\left[\int_t^{\infty} e^{-\alpha\theta} h(\theta)\, d\theta \middle| \mathscr{F}(t)\right] \\ &= E\left[\int_t^{\infty} \alpha e^{-\alpha\theta} \underline{M}(t,\theta)\, d\theta \middle| \mathscr{F}(t)\right], \end{aligned}$$

$$(3.8)\quad e^{-\alpha t} V(t;m) = E\left[\int_t^{\infty} \alpha e^{-\alpha\theta} (m \vee \underline{M}(t,\theta))\, d\theta \middle| \mathscr{F}(t)\right]$$

and the process $\{U(\theta), \mathscr{F}(\theta);\ \theta \ge t\}$ is a martingale with RCLL paths, where

$$(3.9)\quad \begin{aligned} U(\theta) &:= e^{-\alpha\theta}\left[V(\theta;\underline{M}(t,\theta)) - \underline{M}(t,\theta)\right] \\ &\quad + \int_t^{\theta} e^{-\alpha u}\left[h(u) - \alpha \underline{M}(t,u)\right] du. \end{aligned}$$

3.3. REMARK. For any given $t \in [0,\infty)$, the function $m \mapsto (\partial^+/\partial m)V(t;m)$ is right-continuous and increasing, with

$$\left\{\frac{\partial^+}{\partial m} V(t;m) = 1\right\} = \{\sigma(t;m) = t\} = \{V(t;m) = m\} = \{M(t) \le m\}, \quad \text{mod } P.$$

These properties are immediate from Lemmas 2.2 and 2.4, 3.2 and 3.3. On the other hand, from (3.1), (3.2) and (3.5) we have the a.s. representations

$$\begin{aligned} \sigma(t;m) &= \inf\{\theta \ge t \mid \underline{M}(t,\theta) \le m\} = \inf\{\theta \ge t \mid M(\theta) \le m\} \\ &= \inf\{\theta \ge t \mid V(\theta;m) = m\}. \end{aligned}$$

Consider now the two new classes of stopping times $\mathscr{S}^*(t) := \{\tau \in \mathscr{S}(t) \mid \tau > t \text{ a.s.}\}$ and $\mathscr{S}^\dagger(t) := \{\tau \in \mathscr{S}^*(t) \mid \tau = \sigma_t(\xi)$, for some $\xi \in P\mathscr{F}(t)$ with $\xi < M(t)$ a.s.$\}$. In terms of them we have the following representations of $M(t)$; these will *not* be used in the remaining Sections 4–8.

3.4. PROPOSITION. *For any given* $t \in [0, \infty)$ *and with*

$$Q(t, \tau) := \frac{E\left[\int_t^\tau e^{-\alpha\theta} h(\theta)\, d\theta \middle| \mathscr{F}(t)\right]}{E\left[\int_t^\tau e^{-\alpha\theta}\, d\theta \middle| \mathscr{F}(t)\right]}, \qquad \tau \in \mathscr{S}^*(t), \tag{3.10}$$

we have

$$\begin{aligned} \alpha M(t) &= \operatorname*{ess\,sup}_{\tau \in \mathscr{S}^*(t)} Q(t, \tau) \\ &= \operatorname*{ess\,sup}_{\tau \in \mathscr{S}^\dagger(t)} Q(t, \tau) = \lim_{n \uparrow \infty} Q(t, \sigma_t(\xi_n)) \quad a.s. \end{aligned} \tag{3.11}$$

for any strictly increasing sequence $\{\xi_n\}_{n \in \mathbb{N}}$ *of positive,* $\mathscr{F}(t)$*-measurable random variables with* $\lim_{n \to \infty} \xi_n = M(t)$ *a.s.*

3.5. REMARKS. The first equality in (3.11) corresponds to the *forward induction characterization* of the dynamic allocation index $M(t)$ [cf. Gittins (1979, 1989) for the discrete-time, Markovian case], as the essential supremum of "conditional expected discounted reward per unit of conditional expected discounted time"; see also Morimoto (1991) for continuous-time results related to Proposition 3.4. On the other hand, the second equality in (3.7) extends the formula (7.3) of Mandelbaum (1987).

Let us discuss some simple examples, inspired by Whittle [(1982), pages 221–223].

3.6. EXAMPLE. *The "improving" case.* Suppose that we have almost surely $V(s; m) \le V(t; m)$, $\forall\ 0 \le s < t < \infty$, $m \ge 0$ [for this, it suffices that $t \mapsto h(t)$ be increasing]. Then from (3.4), $V(t; M(s)(1 - 1/n)) \ge V(s; M(s)(1 - 1/n)) > M(s)(1 - 1/n)$ and $M(t) > M(s)(1 - 1/n)$, $\forall\ n \in \mathbf{N}$, whence $M(t) \ge M(s)$ almost surely. Therefore, $t \mapsto M(t)$ is increasing, and from (3.5) and (3.8), $\underline{M}(t, \theta) = M(t)$ for $\theta \ge t$, $V(t; m) = m \vee M(t)$ and

$$M(t) = V(t; 0) = E\left[\int_t^\infty e^{-\alpha(\theta - t)} h(\theta)\, d\theta \middle| \mathscr{F}(t)\right].$$

3.7. EXAMPLE. *The "deteriorating" case.* If $t \mapsto h(t)$ is a.s. decreasing, then it is quite easy to see [from (2.1) or (3.11)] that $M(t) = h(t+)/\alpha$, and we have from this and (3.8),

$$\underline{M}(t, \theta) = M(\theta) = \frac{h(\theta +)}{\alpha}, \qquad \theta \ge t$$

and

$$V(t; m) = E\left[\int_t^\infty e^{-\alpha(\theta - t)}(\alpha m \vee h(\theta))\, d\theta \middle| \mathscr{F}(t)\right].$$

This example also leads to the following observation.

3.8. PROPOSITION. *For any given $t \geq 0$, the value of the optimal stopping problem*

$$V'(t;m) := \operatorname*{ess\,sup}_{\tau \in \mathscr{S}(t)} E\left[\int_t^\tau e^{-\alpha(\theta-t)} \alpha \underline{M}(t,\theta)\, d\theta + m e^{-\alpha(\tau-t)} \middle| \mathscr{F}(t)\right],$$

as in (2.1) *with decreasing, right-continuous reward* $h_t'(\theta) := \alpha \underline{M}(t,\theta)$, $\theta \geq t$, *is the same as the value of* (2.1): $V'(t;m) = V(t;m)$ *a.s.*

The next two paragraphs deal with the important case where the process H is a deterministic function of an underlying n-dimensional Markov process $\mathbf{X}$.

3.9. THE MARKOVIAN CASE [Menaldi and Robin (1990)]. Suppose now that $\mathbf{X} = (\{X(t),\ 0 \leq t < \infty\}, \{\mathscr{F}(t),\ 0 \leq t < \infty\}, \{\mathbf{P}^x\}_{x \in \mathscr{R}^n})$ is a time-homogeneous, strong Markov family:

$$P\left[X(\sigma + t_j) \in \Gamma_j, 1 \leq j \leq m \middle| \mathscr{F}(\sigma)\right] = \mathbf{P}^x\left[X(t_j) \in \Gamma_j, 1 \leq j \leq m\right]\Big|_{x = X(\sigma)},$$

P-a.s. on $\{\sigma < \infty\}$, for any $\sigma \in \mathscr{S}$, $m \in \mathbf{N}$, $\Gamma_j \in \mathscr{B}(\mathscr{R}^n)$ and $(t_1, \dots, t_m) \in [0,\infty)^m$. Assume that the associated transition semigroup

$$(S_t f)(x) := \mathbf{E}^x f(X(t)) \qquad \forall\, 0 \leq t < \infty, \tag{3.12}$$

maps the space of bounded, Borel measurable functions $f\colon \mathscr{R}^n \to \mathscr{R}$ into itself, and suppose that

$$\begin{aligned} &h(t) = \eta(X(t)), \qquad t \geq 0; \\ &r(x) := \mathbf{E}^x \int_0^\infty e^{-\alpha t} \eta(X(t))\, dt \in (0,\infty) \qquad \forall\, x \in \mathscr{R}^n, \end{aligned} \tag{3.13}$$

for some Borel measurable function $\eta\colon \mathscr{R}^n \to [0,\infty)$.

Then the optimal stopping reward $V(t;m)$ of (2.1) and the Gittins index $M(t)$ of (3.4) take the form [cf. Fakeev (1971)]

$$V(t;m) = \nu(X(t);m), \qquad M(t) = \mu(X(t)), \qquad 0 \leq t < \infty. \tag{3.14}$$

Here we have set

$$\begin{aligned} \nu(x;m) &:= \sup_{\tau \in \mathscr{S}} \mathbf{E}^x\left[\int_0^\tau e^{-\alpha\theta} \eta(X(\theta))\, d\theta + m e^{-\alpha\tau}\right] \\ &= \mathbf{E}^x\left[\int_0^{\sigma_0(m)} e^{-\alpha\theta} \eta(X(\theta))\, d\theta + m e^{-\alpha\sigma_0(m)}\right], \end{aligned} \tag{2.1$'$}$$

with $\sigma_0(m) = \inf\{\theta \geq 0 \mid \nu(X(\theta);m) = m\}$. The function of (2.1′) satisfies the dynamic programming equation

$$\begin{aligned} \nu(x;m) = \max\Big[& m, e^{-\alpha t}(S_t \nu(\cdot;m))(x) \\ & + \int_0^t e^{-\alpha\theta}(S_\theta \nu(\cdot;m))(x)\, d\theta\Big], \qquad 0 \leq t < \infty, \end{aligned} \tag{2.6$'$}$$

and the *Gittins index function*

$$\mu(x) := \inf\{m \ge 0 \mid \nu(x;m) = m\} \tag{3.4'}$$

also has the forward induction characterizations

$$\begin{aligned} \alpha\mu(x) &= \sup_{\tau \in \mathscr{S}^*(0)} \frac{\mathbf{E}^x \int_0^\tau e^{-\alpha\theta}\eta(X(\theta))\,d\theta}{\mathbf{E}^x \int_0^\tau e^{-\alpha\theta}\,d\theta} \\ &= \lim_{m \uparrow \mu(x)} \frac{\mathbf{E}^x \int_0^{\sigma_0(m)} e^{-\alpha\theta}\eta(X(\theta))\,d\theta}{\mathbf{E}^x \int_0^{\sigma_0(m)} e^{-\alpha\theta}\,d\theta}. \end{aligned} \tag{3.15}$$

3.10. THE ONE-DIMENSIONAL DIFFUSION CASE. Now suppose that the process $\mathbf{X}$ of Section 3.9 is one-dimensional ($n = 1$) and has continuous paths, that the function g of (3.18) is of class $C^1(\mathscr{R})$ with $g' < 0$, that the function $\eta: \mathscr{R} \to [0,\infty)$ is *increasing* and continuous, and that the resolvent $r(\cdot)$ of $\eta(\cdot)$ as in (3.13) is of class $C^1(\mathscr{R})$. Then $\nu(\cdot\,; m)$ of (2.1′) is increasing and of class $C^1(\mathscr{R})$, $\mu(\cdot)$ of (3.4′) is increasing and continuous [see (3.18) and (3.20)], we have

$$\begin{aligned} \sigma_0(m) &\equiv \inf\{t \ge 0 \mid \nu(X(t);m) = m\} \\ &= \inf\{t \ge 0 \mid \mu(X(t)) \le m\} = \tau(\mu^{-1}(m)) \end{aligned} \tag{3.16}$$

with the notation $\tau(y) := \inf\{t \ge 0 \mid X(t) \le y\}$ and thus (3.15) becomes

$$\mu(x) = \lim_{y \uparrow x} \frac{\mathbf{E}^x \int_0^{\tau(y)} e^{-\alpha t}\eta(X(t))\,dt}{1 - \mathbf{E}^x e^{-\alpha\tau(y)}}, \tag{3.17}$$

as noted by Mandelbaum [(1987), relation (4.6)]. Now define the continuous, strictly decreasing function $g: \mathscr{R} \to (0,1)$ via the relationships

$$\mathbf{E}^x e^{-\alpha\tau(y)} = \frac{g(x)}{g(y)} \quad \text{for } x > y \tag{3.18}$$

and $g(0) = 1$. Note that the numerator on the right-hand side of (3.17) is equal to $r(x) - r(y)(g(x)/g(y))$ and observe that (3.17) leads to the explicit computation

$$\begin{aligned} \mu(x) &= \lim_{y \uparrow x} \frac{r(x) - r(y)(g(x)/g(y))}{1 - (g(x)/g(y))} \\ &= r(x) - g(x)\frac{r'(x)}{g'(x)} \end{aligned} \tag{3.19}$$

of Karatzas [(1984), page 180] for the Gittins index function of one-dimensional diffusion processes. In particular, for $0 \le t \le \theta < \infty$,

$$M(t) = \mu(X(t)), \qquad \underline{M}(t,\theta) = \min_{t \le u \le \theta} M(u) = \mu\Big(\min_{t \le u \le \theta} X(u)\Big). \tag{3.20}$$

It can also be seen then, in conjunction with (3.19), that the function $\nu(x;m)$ of (2.1′) is given by

$$\nu(x;m) = \begin{cases} r(x) - g(x)\dfrac{r'(\mu^{-1}(m))}{g'(\mu^{-1}(m))}, & x > \mu^{-1}(m), \\ m, & x \le \mu^{-1}(m), \end{cases} \tag{3.21}$$

$$= (m \vee \mu(x)) - g(x)\int_{x\wedge\mu^{-1}(m)}^{x} \frac{d\mu(\xi)}{g(\xi)},$$

so that

$$\frac{\partial}{\partial m}\nu(x;m) = \frac{g(x)}{g(\mu^{-1}(m))} \wedge 1.$$

In the special case of Brownian motion with constant local drift and variance coefficients b and σ^2, respectively, and with $\gamma_\pm = (1/\sigma^2)(\sqrt{b^2 + 2\alpha\sigma^2} \pm b)$, the functions g, r and μ become, respectively [cf. Karatzas (1984), pages 181–182],

$$g(x) = \exp(-x\gamma_+),$$

$$r(x) = \frac{2}{\sigma^2(\gamma_+ + \gamma_-)}\left[\int_{-\infty}^{x} \exp[-(x-y)\gamma_+]\eta(y)\,dy + \int_{x}^{\infty} \exp[-(y-x)\gamma_-]\eta(y)\,dy\right],$$

$$\mu(x) = \frac{1}{\alpha}\int_0^\infty \eta\left(x + \frac{z}{\gamma_-}\right)e^{-z}\,dz. \tag{3.22}$$

4. Embedding, separation principle and duality. In order to help motivate further developments, we shall show in this section that the optimal stopping problem of (2.1) can be embedded in (and is, in fact, equivalent to) an apparently more general optimization problem that involves both *control* (i.e., choice of a process in continuous-time) and *stopping* (i.e., choice of a stopping time), as follows.

For any $s \in [0,\infty)$, let $\mathscr{A}(s)$ denote the class of increasing adapted processes $T = \{T(u), \mathscr{F}(u);\ 0 \le u < \infty\}$ with RCLL paths and $s \le T(u) \le s + u$, $\forall\ 0 \le u < \infty$. Let $\mathscr{P}(s)$ denote the class of pairs (T, ρ), where $T \in \mathscr{A}(s)$ and ρ is a stopping time of the filtration $\{\mathscr{F}(T(u))\}_{u\ge 0}$, and let $\mathscr{P}^*(s)$ denote the class of pairs $(T,\rho) \in \mathscr{P}(s)$ such that $T(\rho) = \sigma(s;m)$ a.s. in the notation of (2.4). (One interprets T as a "random time change" that never exceeds the regular clock $u \mapsto s + u$, and ρ as an associated "retirement time".)

The *mixed control/stopping problem* consists of maximizing the conditional expected reward

$$\mathscr{J}(T,\rho;s,m) := E\left[\int_0^\rho e^{-\alpha u}h(T(u))\,dT(u) + me^{-\alpha(T(\rho)-s)}\middle|\mathscr{F}(s)\right] \tag{4.1}$$

over $(T, \rho) \in \mathscr{P}(s)$. It can be seen easily that the pair $(I_s, \rho_s) \in \mathscr{P}^*(s)$ given by

$$I_s(u) := s + u, \qquad 0 \le u < \infty \quad \text{and} \quad \rho_s := \sigma(s; m) - s \tag{4.2}$$

is optimal for this problem, thus establishing a *separation principle* between stopping and control and the aforementioned equivalence with the problem of (2.2).

4.1. PROPOSITION. *The pair $(I_s, \rho_s) \in \mathscr{P}^*(s)$ of (4.2) attains the value*

$$W(s; m) := \operatorname*{ess\,sup}_{(T, \rho) \in \mathscr{P}(s)} \mathscr{J}(T, \rho; s, m). \tag{4.3}$$

In fact, with $W^*(s; m) := \operatorname{ess\,sup}_{(T, \rho) \in \mathscr{P}^*(s)} \mathscr{J}(T, \rho; s, m)$, we have

$$\begin{aligned} W(s; m) &= W^*(s; m) = V(s; m), \\ \rho_s &= \inf\{u \ge 0 \mid W(T(u); m) = m\} \quad a.s. \end{aligned} \tag{4.4}$$

Furthermore, it turns out that $V(s; m)$ is also the value of a *dual stochastic control* problem, as follows.

4.2. PROPOSITION. *The function $I_s(\cdot)$ of (4.2) is optimal for the dual minimization problem*

$$\tilde{W}(s; m) := \inf_{T \in \mathscr{A}(s)} \tilde{\mathscr{J}}(T; s, m) \tag{4.5}$$

with $\tilde{\mathscr{J}}(T; s, m) := E[\int_0^\infty \alpha e^{-\alpha u}(m \vee \underline{M}(s, T(u)))\, du | \mathscr{F}(s)]$. *In particular, we have*

$$\tilde{W}(s; m) = V(s; m), \tag{4.6}$$

so that for every $(T, \rho) \in \mathscr{P}(s)$, $T' \in \mathscr{A}(s)$,

$$\begin{aligned} \mathscr{J}(T, \rho; s, m) &\le \mathscr{J}(I_s, \rho_s; s, m) = V(s; m) \\ &= \tilde{\mathscr{J}}(I_s; s, m) \le \tilde{\mathscr{J}}(T'; s, m). \end{aligned} \tag{4.7}$$

5. The dynamic allocation problem. Consider now, on our probability space $(\Omega, \mathscr{F}, \mathscr{P})$, quasi-left-continuous filtrations $\mathbf{F}_i = \{\mathscr{F}_i(t);\ 0 \le t \le \infty\}$, $i = 1, \dots, d$, and corresponding nonnegative, progressively measurable processes $H_i = \{h_i(t), \mathscr{F}_i(t);\ 0 \le t < \infty\}$ satisfying the conditions of Section 2. From the filtrations $\{\mathbf{F}_i\}_{i=1}^d$, we construct a new filtration $\mathbf{F} = \{\mathscr{F}(\underset{\sim}{s})\}_{\underset{\sim}{s} \in S}$ on the orthant $S := [0, \infty)^d$ by setting

$$\mathscr{F}(\underset{\sim}{s}) = \bigvee_{i=1}^d \mathscr{F}_i(s_i), \qquad \underset{\sim}{s} = (s_1, \dots, s_d) \in S. \tag{5.1}$$

Obviously $\mathscr{F}(\underset{\sim}{s}) \subseteq \mathscr{F}(\underset{\sim}{r})$ if $\underset{\sim}{0} \le \underset{\sim}{s} \le \underset{\sim}{r}$ in the usual partial ordering $\underset{\sim}{s} \le \underset{\sim}{r} \Leftrightarrow_{\text{def}} s_i \le r_i, \forall\, i \in \{1, \dots, d\}$.

5.1. DEFINITION [Mandelbaum (1987)]. For any given $\underset{\sim}{s} = (s_1, \dots, s_d) \in S$, a random process $\underset{\sim}{T}(\cdot) = \{\underset{\sim}{T}(t),\ 0 \le t < \infty\}$ with values in S is called a *dynamic allocation strategy for* $\underset{\sim}{s}$ if it has the following properties:

$$T_i(\cdot) \text{ is increasing, right-continuous with } T_i(0) = s_i \quad \forall\, i = 1, \dots, d, \tag{5.2}$$

$$\sum_{i=1}^{d} (T_i(t) - s_i) = t \qquad \forall\, 0 \le t < \infty \quad \text{a.s.,} \tag{5.3}$$

$$\{\underset{\sim}{T}(t) \le \underset{\sim}{r}\} = \bigcap_{i=1}^{d} \{T_i(t) \le r_i\} \in \mathscr{F}(\underset{\sim}{r}) \qquad \forall\, \underset{\sim}{r} = (r_1, \dots, r_d) \in S,\ 0 \le t < \infty. \tag{5.4}$$

We shall denote by $\mathscr{A}(\underset{\sim}{s})$ the class of such strategies. These can be thought of as multiparameter random time-changes and they coincide with the "optional increasing paths" of Walsh (1981).

From (5.2), (5.3) and for every $i \in \{1, \dots, d\}$ we have, almost surely,

$$0 \le T_i(t) - T_i(u) \le \sum_{j=1}^{d} \big(T_j(t) - T_j(u)\big) = t - u \qquad \forall\, 0 \le u \le t < \infty.$$

In particular, almost every path of each $T_i(\cdot)$ is actually *absolutely continuous* with respect to Lebesgue measure, and thus, without loss of generality, we may assume

(5.5) $T_i(t) = s_i + \int_0^t \chi_i(u)\,du$, $0 \le t < \infty$ holds almost surely, where $\chi_i(\cdot)$ is a progressively measurable [with respect to the filtration $\mathbf{G}_{\underset{\sim}{T}}$ of (5.10)] process with values in $[0, 1]$, $i = 1, \dots, d$, and $\sum_{i=1}^{d} \chi_i(t) = 1 \quad \forall 0 \le t < \infty$ a.s.

5.2. DEFINITION. An allocation strategy $\underset{\sim}{T}(\cdot) \in \mathscr{A}(\underset{\sim}{s})$ will be called *pure* if the corresponding processes $\chi_i(\cdot)$, $1 \le i \le d$, of (5.5) take values in $\{0, 1\}$, or, equivalently, if there exists a progressively measurable [with respect to the filtration $\mathbf{G}_{\underset{\sim}{T}}$ of (5.10)] process $\{e(t),\ 0 \le t < \infty\}$ with values in $\{1, \dots, d\}$ such that $\chi_i(t) = 1_{\{e(t) = i\}}$, $0 \le t < \infty$, $i = 1, \dots, d$, almost surely. We shall denote by $\mathscr{A}_p(\underset{\sim}{s})$ the class of pure strategies.

Intuitively, suppose that there are d projects to be engaged in (respectively, "arms" to be "pulled") by the decision-maker. The processes $H_i(\cdot)$, $1 \le i \le d$, represent the instantaneous rewards obtained by "engaging in a particular one of the d available projects," and the filtrations $\mathbf{F}_i$ record the information accumulated on the independent evolution of different projects. The increment $T_i(t) - T_i(u)$ of the process $T_i(\cdot)$ in (5.4) measures "the total amount of time during $[u, t]$ of engaging in the ith project" (respectively, pulling the ith arm), and the process $\chi_i(t)$ of (5.5) measures "the relative intensity of engaging in the ith project at time t." Of course, for pure strategies, one engages in only one project at any given time t, and that is denoted by $e(t)$ as in Definition 5.2. Finally, according to (5.4), the decision about which arm(s) to pull at any given time, and with what intensity, is to be made only on the basis of the collective information $\mathscr{F}_1(T_1(t)) \vee \cdots \vee \mathscr{F}_d(T_d(t))$ from previous engagements, accumulated up to that time.

5.3. Definition. The *dynamic allocation problem* consists of maximizing the conditional expected total discounted reward $E[\mathscr{R}(\underset{\sim}{T})|\mathscr{F}(\underset{\sim}{s})]$, where

$$\mathscr{R}(\underset{\sim}{T}) := \int_0^\infty e^{-\alpha t}\left(\sum_{i=1}^d h_i(T_i(t))\, dT_i(t)\right) \tag{5.6}$$

over $\underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s})$. The value of this problem will be denoted by

$$\Phi(\underset{\sim}{s}) := \operatorname*{ess\,sup}_{T \in \mathscr{A}(\underset{\sim}{s})} E\big[\mathscr{R}(\underset{\sim}{T})\,|\,\mathscr{F}(\underset{\sim}{s})\big], \qquad \underset{\sim}{s} \in S. \tag{5.7}$$

In subsequent sections we shall compute fairly explicitly the random field of (5.7) and shall exhibit strategies $\underset{\sim}{T}^*(\cdot) \in \mathscr{A}(\underset{\sim}{s})$ such that $\Phi(\underset{\sim}{s}) = E[\mathscr{R}(\underset{\sim}{T}^*)|\mathscr{F}(\underset{\sim}{s})]$ a.s. in the case of *independent* filtrations $\mathbf{F}_i$, $i = 1, \cdots, d$. In so doing we shall adopt the method of Whittle (1980), which embeds the dynamic allocation problem of Definition 5.3 into a family of mixed stochastic control/optimal stopping problems of the same sort, but with the additional option of "retiring" (i.e., of abandoning all projects) and receiving a reward $M \geq 0$. This parametrization is the same as that in the problems of (2.1) and (4.3)

To carry out this embedding we shall need to extend the notion of allocation strategy to accommodate a stopping time. We shall say that a measurable function $\underset{\sim}{\nu}: \Omega \to S$ is *stopping point of* $\mathbf{F}$, if $\{\underset{\sim}{\nu} \leq \underset{\sim}{s}\} \in \mathscr{F}(\underset{\sim}{s})$ holds for every $\underset{\sim}{s} \in S$, and for any such $\underset{\sim}{\nu}$ we consider the σ-field

$$\mathscr{F}(\underset{\sim}{\nu}) := \{A \in \mathscr{F}/A \cap \{\underset{\sim}{\nu} \leq \underset{\sim}{s}\} \in \mathscr{F}(\underset{\sim}{s}), \forall\, \underset{\sim}{s} \in S\}. \tag{5.8}$$

From the Definition 5.1 of an allocation strategy, it follows that for $\underset{\sim}{T}(\cdot) \in \mathscr{A}(\underset{\sim}{s})$,

$$\underset{\sim}{T}(t) \text{ is a stopping point of } \mathbf{F} \qquad \forall\, 0 \leq t < \infty, \tag{5.9}$$

and thus

$$\begin{gathered} \mathscr{G}_T(t) := \mathscr{F}(\underset{\sim}{T}(t)) \text{ is well defined for all } 0 \leq t < \infty \\ \text{and } \mathbf{G}_T := \{\mathscr{G}_T(t)\}_{0 \leq t < \infty} \text{ is a filtration.} \end{gathered} \tag{5.10}$$

5.4. Definition. For any given $M \in [0, \infty)$, the mixed *dynamic allocation / stopping problem* consists of maximizing the conditional expected total discounted reward $E[\mathscr{R}(\underset{\sim}{T},\tau;M)|\mathscr{F}(\underset{\sim}{s})]$, where

$$\mathscr{R}(\underset{\sim}{T},\tau;M) := \sum_{i=1}^{d}\int_0^{\tau} e^{-\alpha t}h_i(T_i(t))\,dT_i(t) + Me^{-\alpha\tau}, \tag{5.11}$$

over the class of *policies*

$$\mathscr{P}(\underset{\sim}{s}) := \{(\underset{\sim}{T},\tau) \mid \underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s}),\ \tau \text{ is a } \mathbf{G}_{\underset{\sim}{T}} \text{ stopping time}\} \tag{5.12}$$

in the notation of (5.10). We shall denote the corresponding value random field by

$$\Phi(\underset{\sim}{s};M) := \operatorname*{ess\,sup}_{(\underset{\sim}{T},\tau)\in\mathscr{P}(\underset{\sim}{s})} E[\mathscr{R}(\underset{\sim}{T},\tau;M)|\mathscr{F}(\underset{\sim}{s})], \qquad \underset{\sim}{s} \in S. \tag{5.13}$$

Evidently $\Phi(\underset{\sim}{s}) = \Phi(\underset{\sim}{s};0)$ in the notation of (5.7) and (5.13). On the other hand, the problem of (5.13) is an extension of the problem of (4.3) to the multiparameter case. Let us also note that for every $(\underset{\sim}{T},\tau) \in \mathscr{P}(\underset{\sim}{s})$ we have

$$\begin{aligned}\mathscr{R}(\underset{\sim}{T},\tau;M) - M &= \sum_{i=1}^{d}\int_0^{\tau} e^{-\alpha t}[h_i(T_i(t)) - \alpha M]\,dT_i(t)\\ &=: \tilde{\mathscr{R}}(\underset{\sim}{T},\tau;M).\end{aligned} \tag{5.14}$$

5.5. REMARK. For every $i \in \{1,\dots,d\}$, we may consider for the filtration $\mathbf{F}_i$ the family of optimal stopping problems, analogues of (2.1), with

$$V_i(\infty;m) = m$$

and

$$V_i(t;m) = \operatorname*{ess\,sup}_{\tau\in\mathscr{S}_i(t)} E\left[\int_t^{\tau} e^{-\alpha(\theta-t)}h_i(\theta)\,d\theta + me^{-\alpha(\tau-t)}\middle|\mathscr{F}_i(t)\right], \qquad 0 \le t < \infty, \tag{5.15}$$

parametrized by $m \in [0,\infty)$. [Just as in (2.2), $\mathscr{S}_i(t)$ is the class of stopping times τ of $\mathbf{F}_i$ with $\tau \ge t$ a.s.] We can also consider the associated Gittins index processes $M_i(t)$, their lower envelopes $\underline{M}_i(t,\theta)$ and the martingales $\{U_i(\theta), \theta \ge t\}$ as in Section 3. All the results of Sections 2 and 3 remain in force for these individual optimal stopping problems.

6. Computation of the value. The purpose of this section is to provide the explicit computation (6.9) for the value random field $\Phi(\underset{\sim}{s};M)$ of (5.15) in terms of the values $\{V_i(s_i;m)\}_{i=1,\dots,d,\ 0\le m<\infty}$ for all the individual optimal stopping problems of (5.18), under the assumption that

$$\text{the filtrations } \mathbf{F}_i,\ i = 1,\cdots,d \text{ are } \textit{independent}. \tag{6.1}$$

This computation relies on a certain "reduction property" (6.3) of the value random field, which goes back to Whittle (1980). To state the property, let us denote by $\Phi^{(i)}(\underset{\sim}{s}^{(i)};M)$, $\underset{\sim}{s}^{(i)} = (s_1,\dots,s_{i-1},s_{i+1},\dots,s_d) \in [0,\infty)^{d-1}$, the value random field for the dynamic allocation problem of Definition 5.4 *with the i-th* *project deleted*, that is, for the $d-1$ projects $1,\dots,i-1,i+1,\dots,d$ only. This random field is adapted to the filtration

$$\text{(6.2)}\quad \mathbf{F}^{(i)} = \{\mathscr{F}^{(i)}(\underset{\sim}{s}^{(i)})\}_{\underset{\sim}{s}\in S} \quad\text{with}\quad \mathscr{F}^{(i)}(\underset{\sim}{s}^{(i)}) := \sigma\left(\bigcup_{\substack{j=1\\ j\neq 1}}^{d} \mathscr{F}_j(s_j)\right),\ \underset{\sim}{s}\in S.$$

6.1. Proposition. *For any fixed $i\in\{1,\dots,\text{d}\}$ and $\underset{\sim}{s}\in S$, we have under the condition* (6.1):

$$\text{(6.3)}\quad \Phi(\underset{\sim}{s};M) = \Phi^{(i)}(\underset{\sim}{s}^{(i)};M)\quad a.s.\ on\ \{V_i(s_i;M)=M\} = \{M_i(s_i)\le M\}.$$

In other words, the property (6.3) states that "any given project should be abandoned as soon as its Gittins index has fallen below the retirement reward M." In particular, $\{\Phi(\underset{\sim}{s};M)=M\} = \bigcap_{i=1}^{d}\{V_i(s_i;M)=M\} = \{\mathscr{M}(\underset{\sim}{s})\le M\}$ mod P, from (6.3) and Remark (3.3), where $\mathscr{M}(\underset{\sim}{s})$ is the so-called *index random field*

$$\text{(6.4)}\qquad \mathscr{M}(\underset{\sim}{s}) := \max_{1\le j\le d} M_j(s_j),\qquad \underset{\sim}{s}\in S.$$

Proposition 6.1 suggests considering the class $\mathscr{P}^*(\underset{\sim}{s},M)$ of "write-off" policies $(\underset{\sim}{T},\tau)$ as in Whittle (1980), with the following properties:

1. The policy $(\underset{\sim}{T},\tau)$ abandons any given project i at the $\{\mathscr{F}_i(T_i(t))\}$-stopping time

$$\text{(6.5)}\qquad \begin{aligned}\hat{\tau}_i &:= \inf\{t\ge 0 \mid V_i(T_i(t);M)=M\}\\ &= \inf\{t\ge 0 \mid \underline{M}_i(s_i,T_i(t))\le M\}\\ &= \inf\{t\ge 0 \mid T_i(t)\ge \sigma_i(s_i;M)\},\end{aligned}$$

that is, "as soon as the allocation strategy $\underset{\sim}{T}(\cdot)$ *has engaged in it for a length $T_i(\hat{\tau}_i)-T_i(0)=\sigma_i(s_i;M)-s_i$ of its own time.*" Thus, $T_i(\tau)=T_i(\hat{\tau}_i)=\sigma_i(s_i;M)$ a.s.

2. The $\mathbf{G}_{\underset{\sim}{T}}$-stopping time τ is given as $\tau=\hat{\tau}(\underset{\sim}{s};M)$, where

$$\text{(6.6)}\qquad \hat{\tau}(\underset{\sim}{s};M) := \sum_{i=1}^{d}[\sigma_i(s_i;M)-s_i] = \sum_{i=1}^{d}[T_i(\hat{\tau}_i)-T_i(0)],$$

that is, "*the policy retires when all projects have been abandoned, and only then.*"

All these policies have the *same* retirement time as in (6.6), which also satisfies

$$\text{(6.7)}\qquad \hat{\tau}=\hat{\tau}(\underset{\sim}{s},M) = \inf\{t\ge 0 \mid \Phi(\underset{\sim}{T}(t);M)=M\}\quad a.s.\quad \forall\,(\underset{\sim}{T},\tau)\in\mathscr{P}^*(\underset{\sim}{s},M),$$

by analogy with (4.4). This fact suggests the following reduction.

6.2. SEPARATION PRINCIPLE. *Suppose that* (6.1) *holds. Then for an arbitrary policy* $(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})$, *and any* $\varepsilon > 0$, *there is a policy* $(\underset{\sim}{T}^{\varepsilon}, \hat{\tau}) \in \mathscr{P}^*(\underset{\sim}{s}, M)$ *such that* $E[\mathscr{R}(\underset{\sim}{T}, \tau; M)|\mathscr{F}(\underset{\sim}{s})] - \varepsilon \le E[\mathscr{R}(\underset{\sim}{T}^{\varepsilon}, \hat{\tau}; M)|\mathscr{F}(\underset{\sim}{s})]$ *a.s. In particular,*

$$\Phi(\underset{\sim}{s}; M) = \operatorname*{ess\,sup}_{\underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s})} E\big[\mathscr{R}(T, \hat{\tau}; M)|\mathscr{F}(\underset{\sim}{s})\big] \quad a.s. \tag{6.8}$$

The next step is to notice that (6.8) leads to an explicit expression for $\Phi(\underset{\sim}{s}; M)$ in terms of the values of the optimal stopping problems in (5.15).

6.3. THE WHITTLE (1980) COMPUTATION. *Under the condition* (6.1), *the reduction property* (6.3) *leads to the representation for the value random field of* (5.13) *as*

$$\begin{aligned} \Phi(\underset{\sim}{s}; M) - M &= \int_M^{\infty} \Big(1 - E\big[\exp\left[-\alpha\hat{\tau}(\underset{\sim}{s}; m)\right] \big| \mathscr{F}(\underset{\sim}{s})\big]\Big)\, dm \\ &= \int_M^{\infty} \left(1 - \prod_{i=1}^{d} \frac{\partial^+}{\partial m} V_i(s_i; m)\right) dm. \end{aligned} \tag{6.9}$$

Derivations of (6.3), (6.8) and (6.9) are presented in the Appendix. In the next section we shall use the computation (6.9) to discern allocation policies $\underset{\sim}{T}^*(\cdot)$ that achieve the supremum in both (5.7) and (6.8).

7. Strategies of index type. We introduce in this section a class of "index-type" allocation strategies. Using simple time-change arguments, we prove in Proposition 7.4 their optimality for an auxiliary dynamic allocation problem with *decreasing* rewards, which has the same value as the problem of (5.13). In the next section, we shall show that these index-type strategies are also optimal for our original problem of (5.13), as well as for a dual minimization problem [under the assumption (6.1)]. We start by introducing the *lower envelope index random field*

$$\underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{r}) := \max_{1 \le j \le d} \underline{M}_j(s_j, r_j), \qquad \underline{\mathscr{M}}(\underset{\sim}{r}) \equiv \underline{\mathscr{M}}(\underset{\sim}{0}, \underset{\sim}{r}), \qquad \underset{\sim}{0} \le \underset{\sim}{s} \le \underset{\sim}{r}. \tag{7.1}$$

7.1. DEFINITION. A strategy $\underset{\sim}{T}^*(\cdot) \in \mathscr{A}(\underset{\sim}{s})$ is said to be of *index type* if, for every $i = 1, \dots, d$,

$$T_i^*(\cdot) \text{ is flat off the set } \big\{t \ge 0 \mid \underline{M}_i(s_i, T_i^*(t)) = \underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{T}^*(t))\big\} \quad \text{a.s.} \tag{7.2}$$

In other words, at any given time $t \ge 0$, an allocation strategy of index type engages only projects from among those with *maximal* current lower envelope of the Gittins index $\underline{M}_i(s_i, T_i^*(t))$. Let us consider now the decreasing, right-continuous inverse

$$\mathbf{M}(t, \underset{\sim}{s}) := \inf\left\{ m \geq 0 \Big/ \sum_{i=1}^{d} (\sigma_i(s_i; m) - s_i) \leq t \right\}, \qquad 0 \leq t < \infty \tag{7.3}$$

of the (decreasing, right-continuous) mapping

$$m \mapsto \hat{\tau}(\underset{\sim}{s}; m) = \sum_{i=1}^{d} (\sigma_i(s_i; m) - s_i)$$

in (6.6). By analogy with the interpretation of the Gittins index in Section 3, the random variable $\mathbf{M}(t, \underset{\sim}{s})$ has the significance of an "equitable surrender value" at time t, for the entire collection of projects $i = 1, \dots, d$. Let us also note the equivalence

$$\hat{\tau}(\underset{\sim}{s}; m) > t \quad \Leftrightarrow \quad \mathbf{M}(t, \underset{\sim}{s}) > m \quad \forall\, m \geq 0, t \geq 0. \tag{7.4}$$

7.2. Lemma. *For any strategy* $\underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s})$ *we have*

$$\mathbf{M}(t, \underset{\sim}{s}) \leq \underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{T}(t)), \qquad 0 \leq t < \infty. \tag{7.5}$$

The following result provides a construction of index-type strategies.

7.3. Proposition. *The S-valued process* $\underset{\sim}{I}(\cdot) = (I_1, \dots, I_d)$ *given by*

$$I_j(t) := \sigma_j\big(s_j; \mathbf{M}(t, \underset{\sim}{s})\big), \qquad 0 \leq t < \infty, j = 1, \dots, d, \tag{7.6}$$

satisfies the properties (5.2), (5.4), (7.2), $\underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{I}(t)) = \mathbf{M}(t, \underset{\sim}{s})$, *and* $\sum_{j=1}^{d} (I_j(t) - s_j) \leq t$, $0 \leq t < \infty$. *If it fails to satisfy* (5.3), $\underset{\sim}{I}(\cdot)$ *can be modified into an S-valued process* $\underset{\sim}{I}^*(\cdot) = (I_1^*, \dots, I_d^*)$ *that satisfies all the above properties as well as* (5.3), *and is thus an allocation strategy in* $\mathscr{A}(\underset{\sim}{s})$ *of index type.*

7.4. Proposition. *For fixed* $\underset{\sim}{s} \in S$, *consider the problem of Definition* 5.4, *under the condition* (6.1) *and with the reward processes* $h_i(\theta)$, $\theta \geq s_i$, *replaced by the decreasing, right-continuous processes* $h_i'(\theta) := \alpha \underline{M}_i(s_i; \theta)$, $\theta \geq s_i$ *for* $i = 1, \dots, d$. *The value of this problem*

$$\boldsymbol{\Psi}(\underset{\sim}{s}; M) := \operatorname*{ess\,sup}_{(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})} E\left[\sum_{i=1}^{d} \int_0^{\tau} \alpha e^{-\alpha t} \underline{M}_i(s_i, T_i(t))\, dT_i(t) + M e^{-\alpha \tau} \middle| \mathscr{F}(\underset{\sim}{s}) \right]$$

is the same as that of (5.13):

$$\boldsymbol{\Psi}(\underset{\sim}{s}; M) = \boldsymbol{\Phi}(\underset{\sim}{s}; M) \quad a.s. \tag{7.7}$$

and we have

$$\boldsymbol{\Psi}(\underset{\sim}{s}; M) = E\left[\int_0^{\infty} \alpha e^{-\alpha t} (M \vee \mathbf{M}(t, \underset{\sim}{s}))\, dt \middle| \mathscr{F}(\underset{\sim}{s}) \right]. \tag{7.8}$$

Furthermore, every index-type strategy $\underset{\sim}{T}^*$ *satisfies*

$$\underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{T}^*(t)) = \mathbf{M}(t, \underset{\sim}{s}), \qquad 0 \leq t < \infty, \tag{7.9}$$

and $(\underset{\sim}{T}^*, \hat{\tau}(\underset{\sim}{s}; M))$ *is optimal for this new problem*:

$$\Psi(\underset{\sim}{s}; M) = E\Bigg[\sum_{i=1}^{d} \int_0^{\hat{\tau}(\underset{\sim}{s};M)} \alpha e^{-\alpha t} \underline{M}_i(s_i, T_i^*(t))\, dT_i^*(t) + M e^{-\alpha\hat{\tau}(\underset{\sim}{s};M)} \Bigg| \mathscr{F}(\underset{\sim}{s})\Bigg]. \tag{7.10}$$

8. Optimality and duality.

We are now in a position to establish the main result.

8.1. THEOREM. *Under the condition* (6.1), *any index-type allocation strategy* $\underset{\sim}{T}^* \in \mathscr{A}(\underset{\sim}{s})$ *is optimal for the problem of Definition* 5.3 *and the pair* $(\underset{\sim}{T}^*, \hat{\tau}(\underset{\sim}{s}; M)) \in \mathscr{P}(\underset{\sim}{s})$ *achieves the supremum in* (5.13) *for any given* $M \geq 0$, $\underset{\sim}{s} \in S$.

8.2. COROLLARY. *With* $\hat{\tau}(m) = \sum_{i=1}^{d} \sigma_i(0; m)$ *as in* (6.6), $\mathbf{M}(t) = \mathbf{M}(t, \underset{\sim}{0})$ *as in* (7.3) *and any index-type strategy* $\underset{\sim}{T}^*$, *we have under condition* (6.1),

$$\begin{aligned} \Phi(\underset{\sim}{0}) &= \int_0^\infty [1 - Ee^{-\alpha\hat{\tau}(m)}]\, dm \\ &= \int_0^\infty \alpha e^{-\alpha t} E\mathbf{M}(t)\, dt \\ &= E\mathscr{R}(\underset{\sim}{T}^*) \geq E\mathscr{R}(\underset{\sim}{T}) \qquad \forall\, \underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{0}). \end{aligned} \tag{8.1}$$

8.3. REMARKS. The expression (8.1) shows that the value of the dynamic allocation problem of (5.7), with $\underset{\sim}{s} = \underset{\sim}{0}$, does not depend on the choice of probability space $(\Omega, \mathscr{F}, \mathscr{P})$, $\{\mathbf{F}_i\}_{i=1}^d$, where the problem is formulated. On the other hand, it can be shown that

$$\Phi(\underset{\sim}{s}) = \operatorname*{ess\,sup}_{\underset{\sim}{T} \in \mathscr{A}_p(\underset{\sim}{s})} E[\mathscr{R}(\underset{\sim}{T}) | \mathscr{F}(\underset{\sim}{s})], \qquad \underset{\sim}{s} \in S, \tag{8.2}$$

where, in contrast to (5.7), the supremum is over the class $\mathscr{A}_p(\underset{\sim}{s})$ of *pure* allocation strategies (Definition 5.2). However, *this class will often fail to contain the index-type allocation strategies* $\underset{\sim}{T}^* \in \mathscr{A}(\underset{\sim}{s})$ *of Definition* 7.1.

To see this, let us consider the following very special case, which extends Example 1.3 of Mandelbaum (1987). Suppose that the independent processes H_i, $i = 1, \cdots, d$, have *continuous and strictly decreasing paths* with $h_i(0) \in (0, \infty)$ and $h_i(\infty) = 0$ almost surely. Then with $\underset{\sim}{s} = \underset{\sim}{0}$, we have $M_i(\theta) = \underline{M}_i(\theta) \equiv \underline{M}_i(0, \theta) = h_i(\theta)/\alpha$ from Example 3.7, $\sigma_i(m) \equiv \sigma_i(0, m) = h_i^{-1}(\alpha m)$ from Remark 3.3 and $\mathbf{M}(t) \equiv \mathbf{M}(t; \underset{\sim}{0}) = h(t)/\alpha$, where $h(\cdot)$ is defined as the inverse of $m \mapsto \sum_{j=1}^{d} h_j^{-1}(m)$. Thus, the index-type strategy $\underset{\sim}{I}(\cdot)$ of (7.6) now becomes

$$I_j(t) = \sigma_j(\mathbf{M}(t)) = h_j^{-1}(h(t)), \qquad 0 \leq t < \infty, 1 \leq j \leq d. \tag{8.3}$$

It satisfies $\sum_{j=1}^{d} I_j(t) \equiv t$, is the unique index-type allocation strategy, and *maintains equal indices for all arms at all times* since

$$\alpha M_j\big(I_j(t)\big) = h(t), \qquad 0 \le t < \infty, j = 1, \dots, d. \tag{8.4}$$

Each of the processes $I_j(\cdot)$ in (8.3) is *strictly increasing* for all $t \in (0, \infty)$. In other words, the strategy $\underset{\sim}{I}(\cdot)$ "pulls all arms at all times," albeit with rates $\dot{I}_j(t) \in (0, 1)$ which satisfy $\sum_{j=1}^{d} \dot{I}_j(t) = 1$.

We also have the following *duality result*, an analogue of Proposition 4.2.

8.4. Proposition. *Under* (6.1), *the random field* $\Phi(\underset{\sim}{s}; M)$ *of* (5.13) *and* (6.9) *is also the value of the dual minimization problem*

$$\begin{aligned} &\operatorname*{ess\,inf}_{\underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s})} E\big[\hat{\mathscr{R}}(\underset{\sim}{T}; M) \big| \mathscr{F}(\underset{\sim}{s})\big] = \Phi(\underset{\sim}{s}; M) \\ &\qquad \text{with } \hat{\mathscr{R}}(\underset{\sim}{T}; M) := \int_0^\infty \alpha e^{-\alpha\theta}\big(M \vee \underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{T}(\theta))\big)\, d\theta. \end{aligned} \tag{8.5}$$

More precisely, for any $(\underset{\sim}{T}', \tau) \in \mathscr{P}(\underset{\sim}{s})$, $\underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s})$ *and any two index-type allocation strategies* $\underset{\sim}{T}^*, \underset{\sim}{I}'$ *in* $\mathscr{A}(\underset{\sim}{s})$, *we have almost surely*

$$\begin{aligned} E\big[\mathscr{R}(\underset{\sim}{T}', \tau; M) \big| \mathscr{F}(\underset{\sim}{s})\big] &\le E\big[\mathscr{R}(\underset{\sim}{I}', \hat{\tau}; M) \big| \mathscr{F}(\underset{\sim}{s})\big] \\ &= \Phi(\underset{\sim}{s}; M) = E\big[\hat{\mathscr{R}}(\underset{\sim}{T}^*; M) \big| \mathscr{F}(\underset{\sim}{s})\big] \\ &\le E\big[\hat{\mathscr{R}}(\underset{\sim}{T}; M) \big| \mathscr{F}(\underset{\sim}{s})\big], \end{aligned} \tag{8.6}$$

where $\hat{\tau} \equiv \hat{\tau}(\underset{\sim}{s}; M)$ *is the stopping time of* (6.6).

9. On the nature of the optimal strategy in a special case. The purpose of this section is to illustrate the *nature of the optimal allocation strategy* $\underset{\sim}{I} = (I_1, I_2)$ *of index type*, for the problem of (5.7) with $\underset{\sim}{s} = \underset{\sim}{0}$ and $d = 2$, when it is assumed that each "arm" has the diffusion process dynamics of Section 3.10. General existence results for such strategies were obtained in Mazziotto and Millet (1987) and Dalang (1990), and by Karatzas (1984) and Mandelbaum (1987) in the diffusion case. To obtain more concrete results, here we shall make these dynamics of Section 3.10 as simple as possible; that is, we shall take $b_1 = b_2 = 0$ and $\sigma_1 = \sigma_2 = 1$ so that $X_i \equiv W_i$, $i = 1,2$, are *independent* Brownian motions.

If we also assume identical reward structures, $\eta_1(\cdot) = \eta_2(\cdot) = \eta(\cdot)$ for some function $\eta(\cdot)$ as in Section 3.10, then $M_i(\theta) = \mu(W_i(\theta))$ and $\underline{M}_i(\theta) \equiv \underline{M}_i(0, \theta) = \mu(A_i(\theta))$, for $i = 1, 2$, $\theta \ge 0$, where $\mu(\cdot) = (1/\alpha)\int_0^\infty \eta(\cdot + z/\sqrt{2\alpha})e^{-z}\, dz$ is the index function of (3.19) and $A_i(\theta) := \min_{0 \le u \le \theta} W_i(u)$. We shall seek an allocation strategy $\underset{\sim}{I} = (I_1, I_2)$ that satisfies almost surely

$$I_1(t) = \int_0^t 1_{\{W_1(I_1(s)) > W_2(I_2(s))\}}\, ds, \tag{9.1}$$

$$I_2(t) = \int_0^t 1_{\{W_1(I_1(s)) \le W_2(I_2(s))\}}\, ds \qquad \forall\, 0 \le t < \infty, \tag{9.2}$$

and

$$\{I_1(t) \le r_1\} \cap \{I_2(t) \le r_2\} \in \mathscr{F}_1(r_1) \vee \mathscr{F}_2(r_2) \qquad \forall\, t \ge 0, \forall\, \underset{\sim}{r} = (r_1, r_2) \in S, \tag{9.3}$$

where $\mathbf{F}_i = \{\mathscr{F}_i(u)\}_{u \ge 0}$ is the P-augmentation of $\bar{\mathscr{F}}_i(u) := \sigma(W_i(s),\ s \le u)$, $u \ge 0$. Obviously then the requirements of (5.2)–(5.4) are all established; the strategy $\underset{\sim}{I} = (I_1, I_2)$ is of index type (Definition 7.1) and *belongs to the class* $\mathscr{A}_p(\underset{\sim}{0})$ *of "pure" allocation strategies* (Definition 5.2) with

$$e(t) = \begin{cases} 1, & \text{if } W(t) > 0 \\ 2, & \text{if } W(t) \le 0 \end{cases} \quad \text{and} \quad W(t) = W_1(I_1(t)) - W_2(I_2(t)).$$

To accomplish this, let us start with a standard Brownian motion W on some probability space $(\Omega, \mathscr{F}, P), \{\mathscr{F}(t)\}$ and recall the Tanaka formulae

$$\begin{aligned} W^{\pm}(t) &= L(t) + N_{\pm}(t), \\ |W(t)| &= 2L(t) + N(t), \qquad 0 \le t < \infty \end{aligned} \tag{9.4}$$

[e.g., Karatzas and Shreve (1991), page 205]. Here $W^+ := \max(W, 0)$, $W^- := \max(-W, 0)$,

$$L(t) := \lim_{\varepsilon \downarrow 0} \frac{1}{4\varepsilon} \operatorname{meas}\{0 \le s \le t / |W(s)| \le \varepsilon\}$$

is the local time at the origin for the process W and

$$\begin{aligned} N_+(t) &:= \int_0^t 1_{(0,\infty)}(W(s))\, dW(s), \\ N_-(t) &:= \int_0^t 1_{(-\infty,0]}(W(s))\, dW(s), \\ N(t) &:= N_+(t) + N_-(t) \end{aligned} \tag{9.5}$$

are continuous $\{F(t)\}$-martingales with quadratic variations

$$\begin{aligned} \langle N_+\rangle(t) &= \Gamma_+(t) := \int_0^t 1_{(0,\infty)}(W(s))\, ds, \\ \langle N_-\rangle(t) &= \Gamma_-(t) := t - \Gamma_+(t), \\ \langle N\rangle(t) &= t, \end{aligned} \tag{9.6}$$

respectively. In particular, N is a Brownian motion. With references to Karatzas and Shreve (1991), from these formulae, from the Skorohod equation (pages 210–212) and from the Knight theorem (page 179), we know that (pages 418–421):

1. $L(t) = -\min_{0 \le s \le t} N_{\pm}(s)$ and $2L(t) = -\min_{0 \le s \le t} N(s)$, $0 \le t < \infty$.
2. $N_{\pm}(t) = B_{\pm}(\Gamma_{\pm}(t))$, $0 \le t < \infty$, where $B_{\pm}(\tau) := N_{\pm}(\Gamma_{\pm}^{-1}(\tau))$, $0 \le \tau < \infty$, are *independent* standard Brownian motions.

3. $R_{\pm}(\tau) := \pm W(\Gamma_{\pm}^{-1}(\tau)) = W^{\pm}((\Gamma_{\pm}^{-1}(\tau)), 0 \le \tau < \infty$, are *independent reflecting* Brownian motions, with local times $L_{\pm}(\tau) := \lim_{\varepsilon \downarrow 0}(1/2\varepsilon)\text{meas}\{0 \le \sigma \le \tau \mid R_{\pm}(\sigma) \le \varepsilon\}$ given by

$$L_{\pm}(\tau) = L\big(\Gamma_{\pm}^{-1}(\tau)\big) = -\min_{0 \le \sigma \le \tau} B_{\pm}(\sigma). \tag{9.7}$$

4. The inverse occupation times $\Gamma_{\pm}^{-1}(\tau) := \inf\{t \ge 0 \mid \Gamma_{\pm}(t) > \tau\}$ are given by the Williams formula

$$\Gamma_{\pm}^{-1}(\tau) = \tau + H_{\mp}(L_{\pm}(\tau)). \tag{9.8}$$

Here $H_{\pm}(b) := \inf\{t \ge 0 \mid B_{\pm}(t) \le -b\}$, $b \ge 0$, are independent, Brownian first-passage-time processes, with inverses $H_{\pm}^{-1}(\tau) = -\min_{0 \le \sigma \le \tau} B_{\pm}(\sigma) = L_{\pm}(\tau)$.

With this notation, we take now $W_1 \equiv B_+$, $W_2 \equiv B_-$, we take $\mathbf{F}_i = \{\mathscr{F}_i(u)\}_{u \ge 0}$ to be the P-augmentation of $\bar{\mathscr{F}}_i(u) \equiv \sigma(W_i(s), s \le u)$, $u \ge 0$ for $i = 1, 2$, and we let $\underset{\sim}{I} = (I_1, I_2) \equiv (\Gamma_+, \Gamma_-)$. We get from (9.4) and (9.7):

$$\begin{aligned} W(t) &= W^+(t) - W^-(t) = N_+(t) - N_-(t) \\ &= W_1(I_1(t)) - W_2(I_2(t)), \\ A_1(\theta) &= -L_+(\theta), \qquad A_2(\theta) = -L_-(\theta), \\ A_1(I_1(t)) &= A_2(I_2(t)) = -L(t) \end{aligned} \tag{9.9}$$

and thus $I_1(t) = \int_0^t 1_{\{W(s) > 0\}}\, ds$, $I_2(t) = \int_0^t 1_{\{W(s) \le 0\}}\, ds$ satisfy the requirements (9.1) and (9.2). It remains to verify (9.3), which holds trivially if $t > r_1 + r_2$. Thus, we take $t \le r_1 + r_2$ and note that we have then from (9.7) and (9.8),

$$\begin{aligned} \{I_1(t) \le r_1\} &= \{t \le \Gamma_+^{-1}(r_1)\} = \{t - r_1 \le H_-(L_+(r_1))\} \\ &= \Big\{\min_{0 \le u \le t - r_1} W_2(u) \ge -L_+(r_1)\Big\} \\ &= \{A_2(t - r_1) \ge A_1(r_1)\} \mod P. \end{aligned}$$

Similarly,

$$\{I_2(t) \le r_2\} = \{A_1(t - r_2) \ge A_2(r_2)\} \mod P$$

and thus (9.3) is satisfied since $t \le r_1 + r_2$.

9.1. REMARK. The above construction also implies

$$\begin{aligned} W_1(I_1(t)) \wedge W_2(I_2(t)) &= N_+(t) \wedge N_-(t) \\ &= (W^+(t) \wedge W^-(t)) - L(t) = -L(t) \end{aligned} \tag{9.10}$$

and in conjunction with (9.9) and (9.4),

$$W_1(I_1(t)) \vee W_2(I_2(t)) = |W(t)| - L(t) = L(t) + N(t), \qquad 0 \le t < \infty. \tag{9.11}$$

9.2. REMARK. As we mentioned already, we have from (3.19):

$$(9.12)\quad M_i(\theta) = \mu(W_i(\theta)), \qquad \underline{M}_i(\theta) = \min_{0\le u\le\theta} M_i(u) = \mu(A_i(\theta)),$$

and thus with $H_1 \equiv H_+$, $H_2 \equiv H_-$, $H \equiv H_1 + H_2$, the expressions of (3.16), (6.6) and (7.3) become

$$\sigma_i(m) := \inf\{t \ge 0 \mid \mu(W_i(t)) \le m\} = H_i(-\mu^{-1}(m))1_{\{\mu^{-1}(m)<0\}},$$

$$(9.13)\quad \hat{\tau}(m) := \sigma_1(m) + \sigma_2(m) = H(-\mu^{-1}(m))1_{\{\mu^{-1}(m)<0\}},$$

$$\mathbf{M}(t) := \inf\{m \ge 0 \mid \hat{\tau}(m) \le t\} = \mu(-L(t)).$$

Consequently, the index random field of (6.4) takes the form

$$(9.14)\quad \mathscr{M}(\underset{\sim}{s}) = \max_{i=1,2} M_i(s_i) = \mu(W_1(s_1) \vee W_2(s_2)),\ \underset{\sim}{s} = (s_1, s_2) \in S.$$

We also note the identities $\mathscr{M}(\underset{\sim}{I}(t)) = \mu(L(t)+N(t))$, $\underline{M}_i(I_i(t)) = \mu(A_i(I_i(t))) = \mu(-L(t))$, whence the decreasing process $t \mapsto \mathscr{M}(\underset{\sim}{I}(t)) \equiv \mathscr{M}(\underset{\sim}{0}, \underset{\sim}{I}(t)) = \max_{i=1,2} \underline{M}_i(I_i(t))$ becomes $\mathscr{M}(\underset{\sim}{I}(t)) = \mathbf{M}(t) = \mu(-L(t))$, $0 \le t < \infty$. Obviously, the strategy $\underset{\sim}{I}(\cdot)$ *maintains equal lower envelopes of indices and engages a project with maximal current index* $\mu(W_i(I_i(\cdot)))$ *at all times.*

9.3. REMARK. In conjunction with (8.1) and the distribution $P[2L(t) \in db] = 2(2\pi t)^{-1/2}\exp(-b^2/2t)\,db$, $b > 0$, of Brownian local time [e.g., Karatzas and Shreve (1991), (3.6.28) and (2.8.3)], (9.13) leads to the explicit computation

$$(9.15)\quad \begin{aligned}\Phi(\underset{\sim}{0}) &= E\int_0^\infty \alpha e^{-\alpha t}\mu(-L(t))\,dt\\ &= 4\int_0^\infty\int_0^\infty \eta(x-\xi)e^{-(x+2\xi)\sqrt{2\alpha}}\,dx\,d\xi\end{aligned}$$

for the value of the dynamic allocation problem in (5.7). The expression (9.15) depends on the discount rate $\alpha > 0$ and on the reward function $\eta(\cdot)$, whereas the optimal (index-type) strategy $\underset{\sim}{I}(\cdot) \in \mathscr{A}_p(\underset{\sim}{0})$ of (9.1)–(9.3) *does not.*

APPENDIX

Proofs of selected results.

PROOF OF LEMMA 2.2. The decrease follows from the decrease of the convex function

$$(A.1)\quad \begin{aligned} m \mapsto \varphi(t;m) &:= e^{-\alpha t}[V(t;m) - m]\\ &= \operatorname*{ess\,sup}_{\tau\in\mathscr{S}(t)} E\left[\int_t^\tau e^{-\alpha s}(h(s) - \alpha m)\,ds \,\middle|\, \mathscr{F}(t)\right].\end{aligned}$$

Indeed, with $m_1 > m_2$ we have $0 \le \varphi(\sigma_t(m_2); m_1) \le \varphi(\sigma_t(m_2); m_2) = 0$ and, therefore, $\sigma_t(m_1) \le \sigma_t(m_2)$ almost surely. Thus, for any strictly decreasing sequence $\{m_k\}_{k\in N} \subseteq (m,\infty)$ with $\lim_{k\to\infty} m_k = m$, we have that $\sigma_* := \lim_{k\to\infty}\sigma_t(m_k)$ exists and $\sigma_* \le \sigma_t(m)$ almost surely. Similarly, the decrease of $\varphi(t,\cdot\,)$ gives, for $l > k$, $\varphi(\sigma_t(m_l); m_k) = 0$ almost surely. Now letting $l \uparrow \infty$, we obtain, from the quasi-left-continuity of $Z(\cdot\,;m)$ [and hence of $V(\cdot\,;m)$, $\varphi(\cdot\,;m)$ as well]: $\varphi(\sigma_*;m_k)=0$ a.s., $\forall\ k\in\mathbf{N}$. Finally, the continuity of $m\mapsto\varphi(t;m)$ yields $\varphi(\sigma_*;m)=0$ and thus $\sigma_*\ge\sigma_t(m)$. It follows that $m\mapsto\sigma_t(m)$ is a.s. right-continuous. The remaining properties follow from the assumptions on H. □

Proof of Lemma 2.3. From (A.1) we have $0\le\varphi(t;m)\le E[\int_t^\infty e^{-\alpha s}(h(s)-\alpha m)^+\,ds|\mathscr{F}(t)]\to 0$ as $m\to\infty$ a.s. from the dominated convergence theorem. For (2.7), fix $0\le t\le\infty$, $0\le m<\infty$ and recall the a.s. equality

$$V(t;m)=E\left[m\exp\left[-\alpha(\sigma_t(m)-t)\right]+\int_t^{\sigma_t(m)}e^{-\alpha(s-t)}h(s)\,ds\,\middle|\,\mathscr{F}(t)\right]. \tag{A.2}$$

With $0<\delta<m$, we obtain $E[Z(\sigma_t(m+\delta);m)|\mathscr{F}(t)]=Z(t;m)$ a.s. from (2.5), and this yields

$$\begin{aligned}V(t;m)&=E\Bigg[V(\sigma_t(m+\delta);m)\exp\left[-\alpha(\sigma_t(m+\delta)-t)\right]\\&\qquad+\int_t^{\sigma_t(m+\delta)}e^{-\alpha(s-t)}h(s)\,ds\,\Bigg|\,\mathscr{F}(t)\Bigg]\\&\ge E\big[m\exp\left[-\alpha(\sigma_t(m+\delta)-t)\right]+V(t;m+\delta)\\&\qquad-(m+\delta)\exp\left[-\alpha(\sigma_t(m+\delta)-t)\right]\big|\mathscr{F}(t)\big]\end{aligned} \tag{A.3}$$

in conjunction with $V(\cdot\,;m)\ge m$ and (A.2) with m replaced by $m+\delta$. From Lemma 2.2, (A.3) and bounded convergence, we obtain

$$\begin{aligned}&\limsup_{\delta\downarrow 0}\frac{V(t;m+\delta)-V(t;m)}{\delta}\\&\qquad\le E\big[\exp\left[-\alpha(\sigma_t(m)-t)\right]\big|\mathscr{F}(t)\big]\quad\text{a.s.}\end{aligned} \tag{A.4}$$

On the other hand, the supermartingale property of $Z(\cdot\,;m+\delta)$ gives $Z(t;m+\delta)\ge E[Z(\sigma_t(m);m+\delta)|\mathscr{F}(t)]$, whence

$$\begin{aligned}V(t;m+\delta)&\ge E\Bigg[V(\sigma_t(m);m+\delta)\exp\left[-\alpha(\sigma_t(m)-t)\right]\\&\qquad+\int_t^{\sigma_t(m)}e^{-\alpha(s-t)}h(s)\,ds\,\Bigg|\,\mathscr{F}(t)\Bigg]\\&\ge E\big[(m+\delta)\exp\left[-\alpha(\sigma_t(m)-t)\right]+V(t;m)\\&\qquad-m\exp\left[-\alpha(\sigma_t(m)-t)\right]\big|\mathscr{F}(t)\big]\end{aligned}$$

almost surely, again in conjunction with $V(\cdot\,;m+\delta)\ge m+\delta$ and (A.2). We conclude that

$$\begin{aligned}&\liminf_{\delta\downarrow 0}\frac{V(t;m+\delta)-V(t;m)}{\delta}\\&\qquad\ge E\big[\exp\left[-\alpha(\sigma_t(m)-t)\right]\big|\mathscr{F}(t)\big]\quad\text{a.s.}\end{aligned} \tag{A.5}$$

holds and (2.7) follows from (A.4) and (A.5). □

PROOFS OF (3.4) AND (3.5). Recall the continuous, decreasing function $m \mapsto \varphi(t; m)$ of (A.1). In terms of it, we have from (3.2) the almost sure equivalences, for any $\theta \geq t$,

$$
\begin{aligned}
m \geq \underline{M}(t,\theta) \quad &\Leftrightarrow \quad \sigma(t;m) \leq \theta \quad \Leftrightarrow \quad \varphi(u;m) = 0, \quad \text{for some } u \in [t,\theta] \\
&\Leftrightarrow \quad m \geq M(u), \quad \text{for some } u \in [t,\theta],
\end{aligned}
$$

and obtain (3.5). These same equivalences with $\theta = t$ lead to (3.4)(i) and (3.4)(ii). □

PROOF OF PROPOSITION 3.2. From (3.2) we have almost surely

$$
\begin{aligned}
\exp[-\alpha(\sigma_t(\lambda) - t)] &= \alpha \int_0^\infty e^{-\alpha s} 1_{\{\sigma_t(\lambda) \leq s+t\}}\, ds \\
&= \alpha \int_t^\infty e^{-\alpha(\theta - t)} 1_{\{\lambda \geq \underline{M}(t,\theta)\}}\, d\theta.
\end{aligned} \tag{A.6}
$$

Integrating with respect to λ over $[m, K]$ and then taking $\mathscr{F}(t)$-conditional expectations, we obtain, in conjunction with Fubini's theorem and (2.7),

$$
\begin{aligned}
&V(t;K) - V(t;m) \\
&\quad = \alpha \cdot E\left[\int_t^\infty e^{-\alpha(\theta-t)}\{K - (m \vee \underline{M}(t,\theta))\}\, d\theta \middle| \mathscr{F}(t)\right],
\end{aligned} \tag{A.7}
$$

and (3.8) follows from (A.7) and $\lim_{K\to\infty}[V(t;K) - K] = 0$.

With (3.8) established, we just let $m = 0$ [in (3.8) and (A.2)] to obtain (3.7). Now the right-hand side of (3.8) is equal to

$$
E\left[m \exp[-\alpha\sigma_t(m)] + \alpha \int_t^{\sigma_t(m)} e^{-\alpha\theta} \underline{M}(t,\theta)\, d\theta \middle| \mathscr{F}(t)\right],
$$

whereas, thanks to (A.2), the left-hand side is equal to

$$
E\left[m \exp[-\alpha\sigma_t(m)] + \int_t^{\sigma_t(m)} e^{-\alpha\theta} h(\theta)\, d\theta \middle| \mathscr{F}(t)\right] \quad \text{a.s.}
$$

Thus (3.6) follows by comparing these two expressions.

The RCLL regularity for the paths of the process in (3.9) follows from the RCLL regularity of $t \mapsto V(t;m)$, $\theta \mapsto \underline{M}(t,\theta)$ and from the continuity of $m \mapsto V(t;m)$. To prove the martingale property, write (3.8) in the form

$$
e^{-\alpha\theta}[V(\theta;m) - m] = E\left[\int_\theta^\infty \alpha e^{-\alpha u}(\underline{M}(\theta,u) - m)^+\, du \middle| \mathscr{F}(\theta)\right]. \tag{A.8}
$$

Also write (A.8) with m replaced by $\underline{M}(t,\theta)$, and subtract the resulting equation memberwise from (A.8), observing in particular that on $\{\theta < \sigma_t(m)\} = \{\underline{M}(t,\theta) > m\}$ we have, for $t \leq \theta \leq u$,

$$
\begin{aligned}
(\underline{M}(\theta,u) - m)^+ - (\underline{M}(\theta,u) - \underline{M}(t,\theta))^+ &= (\underline{M}(t,\theta) \wedge \underline{M}(\theta,u) - m)^+ \\
&= (\underline{M}(t,u) - m)^+.
\end{aligned}
$$

It follows that we have on $\{t \le \theta < \sigma_t(m)\}$:

$$e^{-\alpha\theta}\big[V(\theta;m) - m - V(\theta;\underline{M}(t,\theta)) + \underline{M}(t,\theta)\big]$$
$$= E\bigg[\int_\theta^\infty \alpha e^{-\alpha u}(\underline{M}(t,u) - m)^+ \,du \bigg| \mathscr{F}(\theta)\bigg],$$

and therefore the process

$$\begin{aligned}\exp\big[-\alpha(\theta\wedge\sigma_t(m))\big]\big[V(\theta\wedge\sigma_t(m);m) - m \\ -V(\theta\wedge\sigma_t(m);\underline{M}(t,\theta\wedge\sigma_t(m))) + \underline{M}(t,\theta\wedge\sigma_t(m))\big] \\ + \int_t^{\theta\wedge\sigma_t(m)} \alpha e^{-\alpha u}(\underline{M}(t,u) - m)\,du, \qquad \theta \ge t,\end{aligned} \tag{A.9}$$

is a martingale. However, from (2.5),

$$e^{-\alpha(\theta\wedge\sigma_t(m))}\big[V(\theta\wedge\sigma_t(m);m) - m\big] + \int_t^{\theta\wedge\sigma_t(m)} e^{-\alpha u}(h(u) - \alpha m)\,du,\ \theta \ge t,$$

is also a martingale. By subtraction from (A.9), the martingale property is established for the process $\{U(\theta\wedge\sigma_t(m)),\ \theta \ge t\}$ and *any* $m \ge 0$. □

PROOF OF PROPOSITION 3.4. For any $\tau \in \mathscr{S}^*(t)$, we have

$$M(t) = V(t;M(t)) \ge E\bigg[\int_t^\tau e^{-\alpha(\theta-t)}h(\theta)\,d\theta + M(t)e^{-\alpha(\tau-t)}\bigg|\mathscr{F}(t)\bigg]$$

and thus $\alpha M(t) \ge \operatorname{ess\,sup}_{\tau\in\mathscr{S}^*(t)} Q(t,\tau)$ a.s. On the other hand, with $\{\xi_n\}_{n\in\mathbf{N}}$ a strictly increasing sequence of $\mathscr{F}(t)$-measurable, $(0,\infty)$-valued random variables such that $\lim_{n\to\infty}\xi_n = M(t)$, a.s., we have $\sigma_t(\xi_n) \in \mathscr{S}^\dagger(t)$ for all $n \in \mathbf{N}$, as well as

$$\xi_n < V(t;\xi_n) = E\bigg[\int_t^{\sigma_t(\xi_n)} e^{-\alpha(\theta-t)}h(\theta)\,d\theta$$
$$+ \xi_n \exp\big[-\alpha(\sigma_t(\xi_n) - t)\big]\bigg|\mathscr{F}(t)\bigg],$$

whence $\alpha\xi_n < Q(t,\sigma_t(\xi_n))$ almost surely. Letting $n \to \infty$ we deduce from this $\alpha M(t) \le \lim_{n\to\infty} Q(t,\sigma_t(\xi_n)) \le \operatorname{ess\,sup}_{\tau\in\mathscr{S}^\dagger(t)} Q(t,\tau)$ a.s., and (3.11) follows. □

PROOF OF PROPOSITION 3.8. From Example 3.7, the lower envelope of the Gittins index process associated with this new problem is given by $\underline{M}'(t,\theta) = (1/\alpha)h_t'(\theta) = \underline{M}(t,\theta)$. The conclusion now follows from (3.8). □

PROOF OF PROPOSITION 4.1. With every $\tau \in \mathscr{S}(s)$ we can associate the pair $(I_s, \tau - s) \in \mathscr{P}(s)$, which has the same conditional expected total reward, namely,

$$\mathscr{J}(I_s, \tau - s; s, m) = E\bigg[\int_s^\tau e^{-\alpha(u-s)}h(u)\,du + me^{-\alpha(\tau-s)}\bigg|\mathscr{F}(s)\bigg] \quad \text{a.s.}$$

In particular, to the optimal stopping time $\tau^* = \sigma(s;m)$ for the problem $V(s;m)$ of (2.1), we can associate $\rho^* = \tau^* - s = \sigma(s;m) - s$, such that $I_s(\rho^*) = s + \rho^* = \sigma(s;m)$. It follows that $V(s;m) \le W^*(s;m) \le W(s;m)$.

On the other hand, for every $(T, \rho) \in \mathscr{P}(s)$, let $l(\theta) := \inf\{u \geq 0 \mid T(u) > \theta\}$, $s \leq \theta < \infty$, be the increasing, right-continuous inverse of $T(\cdot)$, with $l(s) = 0$ and $l(\theta) \geq u \Leftrightarrow T(u) \leq \theta$, whence $l(\theta) \geq \theta - s, \forall\, s \leq \theta < \infty$. The change-of-variable formula gives

$$\mathscr{J}(T, \rho; s, m) = E\left[\int_s^{T(\rho)} e^{-\alpha l(\theta)} h(\theta)\, d\theta + e^{-\alpha(T(\rho)-s)} m \middle| \mathscr{F}(s)\right]$$
$$\leq E\left[\int_s^{T(\rho)} e^{-\alpha(\theta-s)} h(\theta)\, d\theta + e^{-\alpha(T(\rho)-s)} m \middle| \mathscr{F}(s)\right] \leq V(s; m)$$

for any $(T, \rho) \in \mathscr{P}(s)$, and thus $W(s; m) \leq V(s; m)$, proving (4.4) and the remaining claims. □

Proof of Proposition 4.2. For any $T' \in \mathscr{A}(s)$ we have $T'(u) \leq I_s(u) = s + u$, and thus

$$\tilde{\mathscr{J}}(T'; s, m) \geq \tilde{\mathscr{J}}(I_s; s, m)$$
$$= \alpha \cdot E\left[\int_s^{\infty} e^{-\alpha(\theta-s)}\left(m \vee \underline{M}(s, \theta)\right) d\theta \middle| \mathscr{F}(s)\right] = V(s; m),$$

from (3.8). Both (4.6) and (4.7) follow readily from this. □

Proof of Proposition 6.1. Clearly, (6.3) is implied by the double a.s. inequality

$$\text{(A.10)}\qquad \begin{aligned} \Phi^{(i)}\left(\underset{\sim}{s}^{(i)}; M\right) \vee V_i(s_i; M) &\leq \Phi(\underset{\sim}{s}; M) \\ &\leq \Phi^{(i)}\left(\underset{\sim}{s}^{(i)}; M\right) + V_i(s_i; M) - M, \end{aligned}$$

valid for every given $i \in \{1, \dots, d\}$, $\underset{\sim}{s} \in S$. [This inequality was first obtained by Tsitsiklis (1986) in the discrete-time Markovian case.] The first inequality in (A.10) is obvious. For the second, recall (5.14), as well as (5.13) in the form

$$\text{(A.11)}\qquad \Phi(\underset{\sim}{s}; M) - M = \operatorname*{ess\,sup}_{(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})} E\left[\tilde{\mathscr{R}}(\underset{\sim}{T}, \tau; M) \mid \mathscr{F}(\underset{\sim}{s})\right],$$

and observe (from the independence of the $\mathbf{F}_j$'s and time-change arguments) the analogue

$$\text{(A.12)}\qquad \Phi^{(i)}\left(\underset{\sim}{s}^{(i)}; M\right) - M = \operatorname*{ess\,sup}_{(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})} E\left[\tilde{\mathscr{R}}^{(i)}(\underset{\sim}{T}, \tau; M) \mid \mathscr{F}(\underset{\sim}{s})\right]$$

of (A.11), where

$$\tilde{\mathscr{R}}^{(i)}\left(\underset{\sim}{T}^{(i)}, \tau; M\right) := \sum_{j \neq i} \int_0^{\tau} e^{-\alpha u}\left[h_j(T_j(u)) - \alpha M\right] dT_j(u).$$

In other words, $\Phi^{(i)}(\underset{\sim}{s}^{(i)}; m)$ is the value random field for the problem of Definition 5.4, but with $h_i(\cdot) \equiv 0$. Similarly,

$$\text{(A.13)}\quad V_i(s_i; M) - M = \operatorname*{ess\,sup}_{(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})} E\left[\int_0^{\tau} e^{-\alpha u} \{ h_i(T_i(u)) - \alpha M \} \, dT_i(u) \middle| \mathscr{F}(\underset{\sim}{s}) \right].$$

Using (A.11)–(A.13) and

$$\tilde{\mathscr{R}}(\underset{\sim}{T}, \tau; M) = \int_0^{\tau} e^{-\alpha u} [h_i(T_i(u)) - \alpha M] \, dT_i(u) + \tilde{\mathscr{R}}^{(i)}(\underset{\sim}{T}^{(i)}, \tau; M),$$

we obtain

$$\begin{aligned} E\big[\tilde{\mathscr{R}}(\underset{\sim}{T}, \tau; M) | \mathscr{F}(\underset{\sim}{s})\big] &\le E\left[\sum_{j \ne i} \int_0^{\tau} e^{-\alpha u} \{ h_j(T_j(u)) - \alpha M \} \, dT_j(u) \middle| \mathscr{F}(\underset{\sim}{s}) \right] \\ &\quad + E\left[\int_0^{\tau} e^{-\alpha u} \{ h_i(T_i(u)) - \alpha M \} \, dT_i(u) \middle| \mathscr{F}(\underset{\sim}{s}) \right] \\ &\le \left[\Phi^{(i)}(\underset{\sim}{s}^{(i)}; M) - M \right] + \left[V_i(s_i; M) - M \right]. \end{aligned}$$

Now take supremum over $(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})$ to obtain from (A.11) the second inequality in (A.10). □

Proof of the separation principle (6.2). Take an arbitrary $(\underset{\sim}{T}, \tau) \in \mathscr{P}(\underset{\sim}{s})$, and consider the following cases.

(i) $V_i(T_i(t); M) = M$ a.s. *for some* $t \in [0, \infty)$, *some* $i \in \{1, \dots, d\}$. Then (6.3) gives $\Phi(\underset{\sim}{T}(t); M) = \Phi^{(i)}(\underset{\sim}{T}^{(i)}(t); M)$ a.s., which means that it is then optimal to abandon the ith project at time t.

(ii) $V_i(T_i(t); M) = M$ a.s. *for some* $t \in [0, \infty)$ *and all* $i \in \{1, \dots, d\}$. Then from (6.4) we have $\Phi(\underset{\sim}{T}(t); M) = M$ a.s. In other words, it is then optimal to retire at time t.

(iii) $V_i(T_i(t); M) > M$ a.s. *for some* $t \in [0, \infty)$ and *and some* $i \in \{1, \dots, d\}$. Then $\Phi(\underset{\sim}{T}(t); M) \ge V_i(T_i(t); M) > M$ a.s. Thus it is not optimal to retire, as long as there is at least one project that has not been abandoned.

Putting together these three observations we arrive at the first claim, from which (6.8) follows. □

Proof of (6.9). With $\hat{\tau} \equiv \hat{\tau}(\underset{\sim}{s}; M)$, take conditional expectations with respect to $\mathscr{F}(\underset{\sim}{s})$ in $\mathscr{R}(\underset{\sim}{T}, \hat{\tau}; M \pm \delta) = \mathscr{R}(\underset{\sim}{T}, \hat{\tau}; M) \pm \delta e^{-\alpha \hat{\tau}}$ with $0 < \delta < M$, to obtain

$$\Phi(\underset{\sim}{s}; M \pm \delta) \ge E\big[\mathscr{R}(\underset{\sim}{T}, \hat{\tau}; M) | \mathscr{F}(\underset{\sim}{s})\big] \pm \delta \cdot E\big[e^{-\alpha \hat{\tau}} | \mathscr{F}(\underset{\sim}{s})\big] \quad \text{a.s.,}$$

and then take the essential supremum of the right-hand side over $\underset{\sim}{T} \in \mathscr{A}(\underset{\sim}{s})$ to get from (6.8):

$$\Phi(\underset{\sim}{s}; M \pm \delta) \ge \Phi(\underset{\sim}{s}; M) \pm \delta \cdot E\big[e^{-\alpha \hat{\tau}} | \mathscr{F}(\underset{\sim}{s})\big] \quad \text{a.s.}$$

Therefore, we have almost surely

$$\liminf_{\delta\downarrow 0}\frac{\Phi(\underset{\sim}{s};M+\delta)-\Phi(\underset{\sim}{s};M)}{\delta}\geq E[e^{-\alpha\hat{\tau}}|\mathscr{F}(\underset{\sim}{s})]$$

$$\geq \limsup_{\delta\downarrow 0}\frac{\Phi(\underset{\sim}{s};M)-\Phi(\underset{\sim}{s};M-\delta)}{\delta},$$

and this leads to the first equality in (6.9), thanks to the convexity of $M\mapsto\Phi(\underset{\sim}{s};M)$ and $\lim_{M\to\infty}[\Phi(\underset{\sim}{s};M-M]=0$ a.s. [because from (5.13), (5.14) and dominated convergence,

$$0\leq\Phi(\underset{\sim}{s};M)-M\leq E\left[\sum_{i=1}^{d}e^{-\alpha s_i}\int_{s_i}^{\infty}e^{-\alpha\theta}(h_i(\theta)-\alpha M)^+\,d\theta|\mathscr{F}(\underset{\sim}{s})\right]\to 0$$

as $M\to\infty$]. Finally, from (6.6), (2.7) and the independence of the $\mathbf{F}_i$'s, we have

$$\begin{aligned}E[e^{-\alpha\hat{\tau}}|\mathscr{F}(\underset{\sim}{s})]&=E\left[\prod_{i=1}^{d}\exp\left[-\alpha(\sigma_i(s_i;M)-s_i)\right]\middle|\mathscr{F}(\underset{\sim}{s})\right]\\&=\prod_{i=1}^{d}E\left[\exp\left[-\alpha(\sigma_i(s_i;M)-s_i)\right]|\mathscr{F}_i(s_i)\right]\\&=\prod_{i=1}^{d}\left(\frac{\partial^+}{\partial m}V_i(s_i;M)\right)\quad\text{almost surely.}\end{aligned}$$ □

PROOF OF LEMMA 7.2. From definitions (7.1) and (6.6) and properties (5.3), (3.2) and (7.4),

$$\begin{aligned}\{\underline{\mathscr{M}}(\underset{\sim}{s},\underset{\sim}{T}(t))\leq m\}&=\bigcap_{j=1}^{d}\{\underline{M}_j(s_j,T_j(t))\leq m\}\\&=\bigcap_{j=1}^{d}\{\sigma_j(s_j;m)\leq T_j(t)\}\subseteq\{\hat{\tau}(\underset{\sim}{s};m)\leq t\}\\&=\{\mathbf{M}(t,\underset{\sim}{s})\leq m\}\quad\text{for any } m\geq 0,\end{aligned}$$

and (7.5) follows. □

PROOF OF PROPOSITION 7.3. We have from (7.6) the inequality

$$\underline{M}_j(s_j,I_j(t))=\underline{M}_j\big(s_j,\sigma_j(s_j;\mathbf{M}(t,\underset{\sim}{s}))\big)\leq\mathbf{M}(t,\underset{\sim}{s})\qquad\forall\, j=1,\dots,d,$$

which holds in fact as an *equality* unless $m\mapsto\sigma_j(s_j;m)$ is flat at $m=\mathbf{M}(t,\underset{\sim}{s})$. However, from the definition of $\mathbf{M}(t,\underset{\sim}{s})$ in (7.3), this cannot be the case for all the j's. Therefore, for every $t\geq 0$, there exists some $k\in\{1,\dots,d\}$ such that

$$\underline{M}_k(s_k,I_k(t))=\mathbf{M}(t,\underset{\sim}{s})=\max_{1\leq j\leq d}\underline{M}_j(s_j,I_j(t))=\underline{\mathscr{M}}(\underset{\sim}{s},\underset{\sim}{I}(t)).$$

Additionally, for any given $i \in \{1, \dots, d\}$, the process $I_i(\cdot)$ is flat on $\{t \geq 0 \mid \underline{M}_i(s_i, I_i(t)) < \mathbf{M}(t, \underset{\sim}{s})\}$. The properties of (5.2) and (5.4) are easily verified; in particular, for any $\underset{\sim}{r} \in S$ with $\underset{\sim}{s} \leq \underset{\sim}{r}$ and any $t \geq 0$, we have

$$\{\underset{\sim}{I}(t) \leq \underset{\sim}{r}\} = \bigcap_{j=1}^{d} \{\underline{M}_j(s_j, r_j) \leq \mathbf{M}(t, \underset{\sim}{s})\} = \{\underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{r}) \leq \mathbf{M}(t, \underset{\sim}{s})\} \in \mathscr{F}(\underset{\sim}{r}).$$

We also have

$$\sum_{j=1}^{d} \left(I_j(t) - s_j\right) = \hat{\tau}\left(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})\right) \leq t,$$

and this inequality is strict whenever $\hat{\tau}(\underset{\sim}{s}; \cdot\,)$ has a jump at $m = \mathbf{M}(t, \underset{\sim}{s})$ [in which case $I_j(t) = \sigma_j(s_j; m)$ for $\hat{\tau}(\underset{\sim}{s}; m) \leq t < \hat{\tau}(\underset{\sim}{s}; m\, -\,)$, $j = 1, \dots, d$].

Now the set $\mathbf{D}$ of jump points of $\hat{\tau}(\underset{\sim}{s}; \cdot\,)$ is at most countable, and so is the difference of the two sets $\mathscr{D} := \bigcup_{m \in \mathbf{D}} [\hat{\tau}(\underset{\sim}{s}; m), \hat{\tau}(\underset{\sim}{s}; m\, -\,))$ and $\mathscr{E} := \{t \geq 0 \mid \hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})) < t\}$. (In particular, this difference is not changed by the continuous, increasing functions.) Now for any $m \in \mathbf{D}$, let us partition the interval $L(m) := [\hat{\tau}(\underset{\sim}{s}; m), \hat{\tau}(\underset{\sim}{s}; m\, -\,))$ into disjoint intervals $L_j(m) = [y_{j-1}, y_j)$, $j = 1, \dots, d$, with $L(m) = \bigcup_{j=1}^{d} L_j(m)$, $y_0 = \hat{\tau}(\underset{\sim}{s}; m)$, $y_j - y_{j-1} = \Delta\sigma_j(s_j; m) := \sigma_j(s_j; m\, -\,) - \sigma_j(s_j; m)$ for $j \geq 1$. For any given $t \in L(m)$ find the unique $k \in \{1, \dots, d\}$ such that $t \in L_k(m)$, and let

$$I_j^*(t) = \begin{cases} \sigma_j(s_j; m\, -\,), & j = 1, \dots, k-1, \\ \sigma_j(s_j; m) + (t - y_{k-1}), & j = k, \\ \sigma_j(s_j; m), & j = k+1, \dots, d. \end{cases}$$

This defines the processes $I_j^*(\cdot)$ on $\mathscr{D}$. On $[0, \infty) \setminus \mathscr{D}$ we set $I_j^*(t) \equiv I_j(t) = \sigma_j(s_j; \mathbf{M}(t, \underset{\sim}{s}))$ as before. Clearly, for all $t \in L(m)$, $m \in \mathbf{D}$, we have

$$\sum_{j=1}^{d} \left(I_j^*(t) - s_j\right) = \sum_{j=1}^{d} \left(\sigma_j(s_j; m) - s_j\right) + \sum_{j=1}^{k-1} (y_j - y_{j-1}) + (t - y_{k-1}) = t.$$

Hence $\underset{\sim}{I}^*(\cdot)$ obeys (5.3) for every $t \in \mathbf{D}$, and thus also for every $t \geq 0$ that satisfies either $\hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})) \leq t < \hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})\, -\,)$ or $\hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})) = t = \hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})\, -\,)$. By continuity, this property can be extended to all $t \geq 0$ that satisfy $t = \hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s})\, -\,) > \hat{\tau}(\underset{\sim}{s}; \mathbf{M}(t, \underset{\sim}{s}))$, and thus to all of $[0, \infty)$. This construction preserves all the other properties (5.2), (5.4), (7.2), $\underline{\mathscr{M}}(\underset{\sim}{s}, \underset{\sim}{I}^*(t)) = \mathbf{M}(t, \underset{\sim}{s})$ and provides an allocation strategy of index type. □

Remarks. (i) The ordering of the intervals $L_j(m)$, $j = 1, \dots, d$, is arbitrary, so there can exist several different strategies of index type. The support of $I_j^*(\cdot)$ is included in $\{t \geq 0 \mid \underline{M}_j(s_j, I_j^*(t)) = \mathbf{M}(t, \underset{\sim}{s})\}$, but this inclusion can be strict, so, in general, we do *not* have

$$I_j^*(t) = \int_0^t 1_{\{\underline{M}_j(s_j, I_j^*(u)) = \mathbf{M}(u, \underset{\sim}{s})\}}\, du.$$

(ii) The above proof does not use the independence of the $\mathbf{F}_j$'s. If, however, these filtrations *are* independent, then the (independent) processes $m \mapsto$

$\sigma_j(s_j, m)$, $j = 1,\dots,d$, cannot have common jumps, so for any given $m \in \mathbf{D}$ there exists $k \in \{1,\dots,d\}$ such that $\hat{\tau}(\underset{\sim}{s}; m-) = \hat{\tau}(\underset{\sim}{s}; m) + \Delta\sigma_k(s_k, m)$.

PROOF OF PROPOSITION 7.4. The first claim (7.7) follows from (6.9) and Proposition 3.8. Now from (6.9), (7.7), (7.4) and Fubini's theorem,

$$\begin{aligned}\mathbf{\Psi}(\underset{\sim}{s}; M) - M &= E\left[\int_M^\infty \left(1 - e^{-\alpha\hat{\tau}(\underset{\sim}{s};m)}\right) dm \middle| \mathscr{F}(s)\right]\\ &= E\left[\int_M^\infty\int_0^\infty \alpha e^{-\alpha t} 1_{\{t<\hat{\tau}(\underset{\sim}{s};m)\}}\, dt\, dm \middle| \mathscr{F}(\underset{\sim}{s})\right]\\ &= E\left[\int_M^\infty\int_0^\infty \alpha e^{-\alpha t} 1_{\{m<\mathbf{M}(t,\underset{\sim}{s})\}}\, dt\, dm \middle| \mathscr{F}(\underset{\sim}{s})\right]\\ &= E\left[\int_0^\infty \alpha e^{-\alpha t}\left(\mathbf{M}(t,\underset{\sim}{s}) - M\right)^+ dt \middle| \mathscr{F}(\underset{\sim}{s})\right],\end{aligned}$$

which gives (7.8). On the other hand, for any index-type strategy $\underset{\sim}{T}^*$ we have, from (7.2) and (7.5),

$$\begin{aligned}&E\left[\int_0^{\hat{\tau}(\underset{\sim}{s};M)} \alpha e^{-\alpha t}\sum_{i=1}^d \left(\underline{M}_i(s_i, T_i^*(t)) - M\right) dT_i^*(t) \middle| \mathscr{F}(\underset{\sim}{s})\right]\\ &\quad= E\left[\int_0^{\hat{\tau}(\underset{\sim}{s};M)} \alpha e^{-\alpha t}\left(\underline{\mathscr{M}}(\underset{\sim}{s}, T^*(t)) - M\right) dt \middle| \mathscr{F}(s)\right]\\ &\quad\ge E\left[\int_0^{\hat{\tau}(\underset{\sim}{s};M)} \alpha e^{-\alpha t}\left(\mathbf{M}(t,\underset{\sim}{s}) - M\right) dt \middle| \mathscr{F}(\underset{\sim}{s})\right]\\ &\quad= E\left[\int_0^\infty \alpha e^{-\alpha t}\left(\mathbf{M}(t,\underset{\sim}{s}) - M\right)^+ dt \middle| \mathscr{F}(\underset{\sim}{s})\right] = \mathbf{\Psi}(\underset{\sim}{s}; M) - M.\end{aligned}$$

Since the reverse inequality, between the first and last of these expressions, is obvious [from (5.14)], we obtain the optimality of $(\underset{\sim}{T}^*, \hat{\tau})$ for the new problem, and the identity (7.9). □

PROOF OF THEOREM 8.1. A comparison of (7.7) and (7.10) with (5.13) makes clear that it suffices to show, for any index-type $\underset{\sim}{T}^* \in \mathscr{A}(\underset{\sim}{s})$,

$$E\left[\sum_{i=1}^d \int_0^{\hat{\tau}(\underset{\sim}{s};M)} e^{-\alpha t}\left\{h_i(T_i^*(t)) - \alpha\underline{M}_i(s_i, T_i^*(t))\right\} dT_i^*(t) \middle| \mathscr{F}(\underset{\sim}{s})\right] = 0 \quad \text{a.s.} \tag{A.14}$$

Now for any $i = 1,\dots,d$ and any point of increase t of $T_i^*(\cdot)$, we have $\underline{M}_i(s_i, T_i^*(t)) = \mathbf{M}(t,\underset{\sim}{s})$ a.s. from (7.2) and (7.9), and $\sigma_i(s_i; \mathbf{M}(t,\underset{\sim}{s})-) \ge T_i^*(t) \ge \sigma_i(s_i; \mathbf{M}(t,\underset{\sim}{s}))$ from (3.2). Therefore,

$$\begin{aligned}V_i\left(T_i^*(t); \underline{M}_i(s_i, T_i^*(t))\right) &= V_i\left(T_i^*(t); \mathbf{M}(t,\underset{\sim}{s})\right)\\ &= \mathbf{M}(t,\underset{\sim}{s}) = \underline{M}_i(s_i, T_i^*(t))\end{aligned} \tag{A.15}$$

from Remark 3.3. Let us recall the $\mathbf{F}_i$-martingale

$$U_i(\theta) = e^{-\alpha\theta}\left[V_i(\theta; \underline{M}_i(s_i, \theta)) - \underline{M}_i(s_i, \theta)\right] + \int_{s_i}^{\theta} e^{-\alpha u}\left[h_i(u) - \alpha \underline{M}_i(s_i, u)\right] du, \quad \theta \geq s_i,$$

as in (3.9) and note, using the independence of the $\mathbf{F}_i$'s, that

$$\text{(A.16)} \quad E\left[\sum_{i=1}^{d} \int_0^{\hat{\tau}(\underset{\sim}{s}; M)} \exp\left[-\alpha(t - T_i^*(t))\right] dU_i(T_i^*(t)) \middle| \mathscr{F}(\underset{\sim}{s})\right] = 0 \quad \text{a.s.}$$

From (A.15), we see that

$$dU_i(T_i^*(t)) = \exp\left[-\alpha T_i^*(t)\right]\left[h_i(T_i^*(t)) - \alpha \underline{M}_i(s_i, T_i^*(t))\right] dT_i^*(t).$$

Thus (A.14) follows from (A.16). □

PROOF OF PROPOSITION 8.4. Only the last two relations in (8.6) require discussion, and they follow directly from (7.8), (7.7) and (7.5). □

NOTE ADDED IN PROOF. We have remarked recently that the assumption (6.1) (independence of the filtrations $\mathbf{F}_i$, $i = 1, \cdots, d$) can be relaxed considerably. In particular, it is possible to show that Proposition 7.3 and the results (7.8)–(7.10) of Proposition 7.4 remain valid *for arbitrary such filtrations*. Furthermore, the results of Section 8, as well as (7.7) and the first inequality in (6.9), continue to hold if one imposes on the $\mathbf{F}_i$'s a *conditional independence* assumption reminiscent of "condition F.4" in Cairoli and Walsh (1975) and Walsh (1981). However, the second equality in (6.9) is not valid under this weaker condition and requires the full strength of (6.1). We hope to present the arguments for these extensions in a future publication.

Acknowledgments. We are grateful to Avi Mandelbaum for explaining to us, forcefully and tirelessly, the merits of the multiparameter approach to dynamic allocation *eight years ago*, and regret that his efforts did not bear more immediate fruit with us. We also thank Daniel Ocone, Michael Phelan, Walter Willinger and X. X. Xue for helpful suggestions, and Robert Dalang for a discussion that prompted us to find the explicit form (7.6) of an optimal strategy. The two referees and the Editor requested clarifications which led to a thorough rewriting, and hopefully to a much more readable manuscript.

REFERENCES

BISMUT, J. M. and SKALLI, B. (1977). Temps d' arrêt optimal, théorie générale de processus, et processus de Markov. *Z. Warsch. Verw. Gebiete* **39** 301–313.

CAIROLI, R. and DALANG, R. C. (1993). *Optimal Stopping and Sequential Control of Stochastic Process*. Wiley, New York. To appear.

CAIROLI, R. AND WALSH, J. B. (1975). Stochastic integrals in the plane. *Acta Math.* **134** 111–183.

DALANG, R. C. (1990). Randomization in the two-armed bandit problem. *Ann. Probab.* **18** 218–225.

DELLACHERIE, C. (1972). *Capacités et Processus Stochastiques.* Springer, Berlin.

EL KAROUI, N. (1981). Les aspects probabilistes du contrôle stochastique. *École d'Éte de Probabilites de Saint Flour IX 1979. Lecture Notes in Math.* **876** 73–238. Springer, Berlin.

EL KAROUI, N. and KARATZAS, I. (1993). General Gittins index processes in discrete time. *Proc. Nat. Acad. Sci. U.S.A.* **90** 1232–1236.

EPLETT, W. J. R. (1986). Continuous-time allocation indices and their discrete-time approximations. *Adv. in Appl. Probab.* **18** 724–746.

FAKEEV, A. G. (1970). Optimal stopping rules for processes with continuous parameter. *Theory Probab. Appl.* **15** 324–331.

FAKEEV, A. G. (1971). Optimal stopping of a Markov process. *Theory Probab. Appl.* **16** 694–696.

GITTINS, J. C. (1979). Bandit processes and dynamic allocation indices (with discussion). *J. Roy. Statist. Soc. Ser. B* **41** 148–164.

GITTINS, J. C. (1989). *Multi-Armed Bandits and Allocation Indices.* Wiley, Chichester.

GITTINS, J. C. and JONES, D. M. (1974). A dynamic allocation index for the sequential design of experiments. In *Progress in Statistics* (J. Gani, ed.) 241–266. North-Holland, Amsterdam.

KARATZAS, I. (1984). Gittins indices in the dynamic allocation problem for diffusion processes. *Ann. Probab.* **12** 173–192.

KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, New York.

MANDELBAUM, A. (1986). Discrete multi-armed bandits and multi-parameter processes. *Probab. Theory Related Fields* **71** 129–147.

MANDELBAUM, A. (1987). Continuous multi-armed bandits and multi-parameter processes. *Ann. Probab.* **15** 1527–1556.

MAZZIOTTO, G. and MILLET, A. (1987). Stochastic control of two-parameter processes, with application to the two-armed bandit problem. *Stochastics* **22** 251–288.

MENALDI, J. L. and ROBIN, M. (1990). On the optimal reward function of the continuous-time multi-armed bandit problem. *SIAM J. Control Optim.* **28** 97–112.

MORIMOTO, H. (1991). On average cost stopping time problems. *Probab. Theory Related Fields* **90** 469–490.

TSITSIKLIS, J. (1986). A lemma on the multi-armed bandit problem. *IEEE Trans. Automat. Control* **AC-31** 576–577.

VARAIYA, P., WALRAND, J. and BUYUKKOC, C. (1985). Extensions of the multi-armed bandit problem: The discounted case. *IEEE Trans. Automat. Control* **AC-30** 426–439.

WALSH, J. B. (1981). Optional increasing paths. *Processus Aléatoire a Deux Indices. Lecture Notes in Math.* **863** 172–201. Springer, Berlin.

WEBER, R. (1992). On the Gittins index for multi-armed bandits. *Ann. in Appl. Probab.* **2** 1024–1033.

WHITTLE, P. (1980). Multi-armed bandits and the Gittins index. *J. Roy. Statist. Soc. Ser. B* **42** 143–149.

WHITTLE, J. B. (1982). *Optimization over Time: Dynamic Programming and Stochastic Control* **1**. Wiley, New York.

LABORATOIRE DE PROBABILITÉS
UNIVERSITÉ PIERRE ET MARIE CURIE
4, PLACE JUSSIEU-TOUR 56
75252 PARIS CEDEX 05
FRANCE

DEPARTMENT OF STATISTICS
619 MATHEMATICS BUILDING
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027