The Annals of Probability
1983, Vol. 11, No. 2, 442–450

THE ADVANTAGE OF USING NON-MEASURABLE STOP RULES

BY THEODORE P. HILL AND VICTOR C. PESTIEN

Georgia Institute of Technology and University of Miami

Comparisons are made between the expected returns using measurable and non-measurable stop rules in discrete-time stopping problems. In the independent case, a natural sufficient condition ("preservation of independence") is found for the expected return of every bounded non-measurable stopping function to be equal to that of a measurable one, and for that of every unbounded non-measurable stopping function to be arbitrarily close to that of a measurable one. For non-negative and for uniformly-bounded independent random variables, universal sharp bounds are found for the advantage of using non-measurable stopping functions over using measurable ones. Partial results for the dependent case are obtained.

1. Introduction. In classical optimal stopping problems a player is faced with a fixed sequence of random variables whose distributions he knows, and realizations of which will be shown him sequentially. His objective usually is to determine a strategy for stopping (stop rule) which will make the expected value at the time he stops as large, or small, as possible. In nearly all classical formulations of such problems, (e.g. Chow, Robbins, and Siegmund (1971)), the player is restricted to stopping only on *measurable* sets. The purpose of this paper is to analyze the situation if the player is allowed to stop on *arbitrary* sets, and compare his expected gain with the gain of measurable stop rules. That is, the player is again faced with a fixed sequence of known distributions, but now he may label the probabilities of non-measurable sets in any consistent manner (i.e. the result must be a finitely additive probability which agrees with the original probability distribution on the Borel sets), and may then select a stopping function allowing him to stop at any stage for any set (measurable or not) of (real) values he wishes.

In Section 2 these notions are made precise, and it is shown (Proposition 2.6) that for integrable sequences, the expected gain using a non-measurable stopping function is uniquely determined by the extension of the probability distribution.

In Section 3 it is shown (Theorem 3.4) that for finite sequences of independent random variables, if the extension "preserves independence" then there is no advantage to using non-measurable stopping functions, in fact, for every non-measurable stopping function there is a measurable one with exactly the same expectation. For arbitrary extensions, universal sharp bounds are found for the advantage of using non-measurable versus measurable stopping functions in the cases of non-negative (Theorem 3.11) and of uniformly bounded (Theorem 3.12) independent random variables.

In Section 4 it is shown (Theorem 4.1) that also in the unbounded case if the extension preserves independence there is again no advantage to using nonmeasurable stopping functions, but in a slightly weaker sense: for every nonmeasurable unbounded stopping function there is a measurable one with arbitrarily close expectation; in general equality is not attainable.

Section 5 discusses a few aspects of the case of arbitrarily-dependent random variables, and derives universal sharp bounds (Theorem 5.1) for the non-negative finite stage problem.

2. Eudoxus integration and stopping functions. Throughout this paper, X_1, X_2, $\cdots$ will be a sequence of integrable random variables on a probability triple $(\Omega, \mathscr{A}, \mu)$, and in all but the last section, will be assumed to be (mutually) independent.

Received October 1981; revised April 1982.
AMS 1980 *subject classification.* Primary 60G40; secondary 28A20, 90C39.
Key words and phrases. Optimal stopping theory, non-measurable stopping function, stop rule.

$\mathbb{N}$ will denote the natural numbers, $\mathbb{R}^n$ Euclidean n-space, and $\mathcal{B}^n$ the Borel σ-algebra on $\mathbb{R}^n$. I_S will denote the indicator function of an arbitrary set S.

EX_j will denote the integral of X_j with respect to μ. $\tilde{P}$ will be a (finitely additive) extension of the distribution P of $X_1, X_2, \cdots$ to all subsets of $\mathbb{R}^\infty$; that is, $\tilde{P}$ is a finitely additive probability measure, defined on the power set $\mathscr{P}(\mathbb{R}^\infty)$ of $\mathbb{R}^\infty$, which agrees with P on $\mathcal{B}^\infty(\tilde{P}(B) = P(B) = \mu(\{\omega: (X_1(\omega), X_2(\omega), \cdots) \in B\})$ for all $B \in \mathcal{B}^\infty)$. The Hahn-Banach Theorem guarantees that such a $\tilde{P}$ (and in general many) exists.

After Dubins and Savage [(1976), page 10] a function $f: \mathbb{R}^\infty \to \mathbb{R}$ will be called *Eudoxus integrable* (relative $\tilde{P}$) if f is either $\mathcal{B}^\infty$-measurable and $E|f| < \infty$, or non-measurable and there is only one possible value for the integral of f (with respect to $\tilde{P}$) based on the linearity and order-preserving properties of integration.

DEFINITION 2.1. $\mathscr{E}$ is the class of all functions from $\mathbb{R}^\infty$ to $\mathbb{R}$ which are Eudoxus integrable relative $\tilde{P}$ for all extensions $\tilde{P}$ of P; and for each $f \in \mathscr{E}$, $\tilde{E}f$ is the integral of f with respect to $\tilde{P}$. (Of course, the integral of f may differ for different extensions of P if f is non-measurable).

By definition, all $\mathcal{B}^\infty$-measurable, P-integrable functions are in $\mathscr{E}$, and it is easy to see that all arbitrary simple functions (functions with finite range) are in $\mathscr{E}$; in fact, $\tilde{E} \sum_{i=1}^n a_i I_{S_i} = \sum_{i=1}^n a_i \tilde{P}(S_i)$. The next lemma generalizes these two facts.

LEMMA 2.2 (a) *If $f: \mathbb{R}^\infty \to \mathbb{R}$ is bounded, then $f \in \mathscr{E}$.*
(b) *If g and h are integrable random variables on $(\mathbb{R}^\infty, \mathcal{B}^\infty, P)$, and if $f: \mathbb{R}^\infty \to \mathbb{R}$ satisfies $h \le f \le g$, then $f \in \mathscr{E}$.*

PROOF. Part (a) is routine, its statement without proof is in Dubins and Savage [(1976), page 10]. For part (b), begin by fixing an extension $\tilde{P}$, and first assume $f \ge 0$. By part (a), $f \wedge n \in \mathscr{E}$ for all n, so $E(f \wedge n)$ is an increasing sequence bounded above by $E(g)$, and $\lim_{n\to\infty} E(f \wedge n) = \alpha$ exists and is finite. It will be shown that $\tilde{E}f$ exists and $= \alpha$. Denoting the outer integral with respect to $\tilde{P}$ by E^* (i.e. $E^*(f) = \inf\{\tilde{E}(\phi): f \le \phi$, and $\phi \in \mathscr{E}\}$) and the inner integral by E_*, note first that $E_*(f) \ge \alpha$. Fix $\varepsilon > 0$ and pick m such that $E(g \cdot I_{(g>m)}) < \varepsilon$. Since $f \cdot I_{(f>m)} \le g \cdot I_{(g>m)}$, it follows that $E^*(f \cdot I_{(f<m)}) < \varepsilon$. Since $f = f \cdot I_{(f\le m)} + f \cdot I_{(f>m)} \le f \wedge m + f \cdot I_{(f>m)}$, the subadditivity of E^* implies $E^* f \le E^*(f \wedge m) + E^*(f \cdot I_{(f>m)}) = \tilde{E}(f \wedge m) + E^*(f \cdot I_{(f>m)}) \le \alpha + \varepsilon$. Since ε was arbitrary, $E^* f = E_* f = \alpha$, and hence $f \in \mathscr{E}$. For general f, apply the above argument to the positive and negative parts of f. □

It should perhaps be mentioned that for functions not in $\mathscr{E}$, that is, where $\tilde{E}f$ is not uniquely determined by linearity and order preserving properties alone, various alternative definitions of the $\tilde{P}$-integral of f have been studied extensively, e.g. Dunford and Schwartz [(1958), Section III.2] and Purves and Sudderth [(1976), Section 4].

DEFINITION 2.3. A function $s: \mathbb{R}^\infty \to \mathbb{N}$ is a *stopping function* if $s(r_1', r_2', \cdots) = n$ whenever $s(r_1, r_2 \cdots) = n$ and $r_i' = r_i$ for all $i = 1, 2, \cdots, n$.

Notice two differences between this definition and the conventional definition of a *stop rule*. First, no mention is made of random variables or of measurability, and second, a stopping function is always defined in terms of subsets of $\mathbb{R}$, rather than subsets of Ω. This approach seems more natural to the authors, since *implementation* of stop rules invariably involves only sets of real values with which the player is content to stop, not observation of the underlying subsets of Ω. The stopping functions defined here are essentially the "stop rules" of Dubins and Savage (1976).

DEFINITION 2.4. $\mathscr{S}$ is the set of all stopping functions, and $\mathscr{S}_n \subset \mathscr{S}$ is the set of all stopping functions which stop no later than n (i.e. $\mathscr{S}_n = \{s \in \mathscr{S}: s \le n$ everywhere$\}$). $\mathscr{T} \subset \mathscr{S}$ is the set of measurable stopping functions (i.e. $\mathscr{T} = \{s \in \mathscr{S}: s^{-1}(n) \in \mathcal{B}^n \times \mathbb{R}^\infty$ for all $n\}$), and $\mathscr{T}_n = \mathscr{T} \cap \mathscr{S}_n$ is the set of measurable stopping functions which stop no later than n. (Without ambiguity, the domain of $s \in \mathscr{S}_n$ will sometimes be taken as $\mathbb{R}^n$.)

DEFINITION 2.5. If $X_1, X_2, \cdots$ are random variables and $s \in \mathscr{S}$, then X_s: $\Omega \to \mathbb{R}$ is the function defined by $X_s(\omega) = X_n(\omega)$ for all ω with $s(X_1(\omega), X_2(\omega), \cdots) = n$. (For integration purposes, X_s will be identified with the function π_s: $\mathbb{R}^\infty \to \mathbb{R}$ defined by $\pi_s(r_1, r_2, \cdots) = r_{s(r_1, r_2, \cdots)}$. For example, $\tilde{E}X_s$ means $\tilde{E}\pi_s$, and $X_s \in \mathscr{E}$ means $\pi_s \in \mathscr{E}$.)

PROPOSITION 2.6 (a) *If $X_1, X_2, \cdots$ are random variables satisfying*

$$-\infty < E(\inf X_j) \le E(\sup X_j) < \infty,$$

then $X_s \in \mathscr{E}$ for each $s \in \mathscr{S}$.
(b) *If $X_1, X_2, \cdots, X_n$ are integrable random variables then $X_s \in \mathscr{E}$ for each $s \in \mathscr{S}_n$.*

PROOF. The first assertion follows from Lemma 2.2(b); the second assertion is an immediate consequence of the first. □

3. The finite-stage stopping problem. Recall that $X_1, X_2, \cdots$ are (mutually) independent integrable random variables on $(\Omega, \mathscr{A}, \mu)$.

DEFINITION 3.1. An extension $\tilde{P}$ of P *preserves the independence of* $X_1, X_2, \cdots$ if $\tilde{P}(A_1 \times A_2 \times \mathbb{R}^\infty) = \tilde{P}(A_1 \times \mathbb{R}^\infty) \cdot \tilde{P}(\mathbb{R}^k \times A_2 \times \mathbb{R}^\infty)$ for all $k \ge 1$ and $n \ge 1$, and all $A_1 \subseteq \mathbb{R}^k$ and $A_2 \subseteq \mathbb{R}^n$.

PROPOSITION 3.2. *There always exists an extension $\tilde{P}$ of P which preserves the independence of $X_1, X_2, \cdots$.*

PROOF. For $i = 1, 2, \cdots$, let P_{X_i} be the distribution of X_i, that is, P_{X_i} is the countably additive probability measure on $(\mathbb{R}, \mathbb{B})$ satisfying $P_{X_i}(B) = \mu(X_i^{-1}(B)) = P(\mathbb{R}^{i-1} \times B \times \mathbb{R}^\infty)$ for all $B \in \mathbb{B}$. For each i, fix one (finitely additive) extension $\tilde{P}_{X_i}$ of P_{X_i} on $(\mathbb{R}, \mathscr{P}(\mathbb{R}))$, the existence of which is guaranteed by the Hahn-Banach theorem.

Let V be the vector space of all finite linear combinations of indicator functions of sets in $\mathbb{R}^\infty$, and define a sequence of subspaces $V_1 \subseteq W_2 \subseteq V_2 \subseteq W_3 \subseteq V_3 \subseteq \cdots$ of V as follows. V_j and W_j are the sets of all finite linear combinations of indicator functions of sets of the form $\mathbb{R}^k \times A_j \times \mathbb{R}^\infty$ where $k \ge 0$ and $A_j \subseteq \mathbb{R}^j$, and respectively, of the form $\mathbb{R}^k \times A_1 \times \cdots \times A_n \times \mathbb{R}^\infty$, where $k \ge 0$, $A_i \subseteq \mathbb{R}^{d(i)}$, and $\sum_{i=1}^n d(i) = j$.

For each $j = 1, 2, \cdots$ a linear functional $\tilde{L}_j$ on V_j is defined inductively as follows. $\tilde{L}_1(I_{\mathbb{R}^k \times A_1 \times \mathbb{R}^\infty}) = \tilde{P}_{X_{k+1}}(A_1)$, and $\tilde{L}_1$ is extended to all of V_1 by linearity. Assume $\tilde{L}_1, \cdots, \tilde{L}_{j-1}$ have been defined, and let L_j be the extension of $\tilde{L}_{j-1}$ to W_j given by

$$L_j(I_{\mathbb{R}^k \times A_1 \times \cdots \times A_n \times \mathbb{R}^\infty}) = \prod_{i=1}^n \tilde{L}_{d(i)}(I_{\mathbb{R}^{k+d(0)+d(1)+\cdots+d(i-1)} \times A_i \times \mathbb{R}^\infty}), \tag{1}$$

where $A_i \subseteq \mathbb{R}^{d(i)}$, and $d(0) = 0$. (To check that L_j is well defined, use the fact that Cartesian product representations are essentially unique.) By the Hahn-Banach theorem (using the outer-measure (integral) of L_j as the subadditive function), L_j may be extended to a linear functional $\tilde{L}_j$ on V_j.

Next define a linear functional L on $V_\infty = \cup_{n=1}^\infty V_n$ by $L(I_{A \times \mathbb{R}^\infty} = \tilde{L}_k(I_{A \times \mathbb{R}^\infty})$ if $A \subseteq \mathbb{R}^k$. Applying the Hahn-Banach theorem again, L may be extended to a linear functional $\tilde{L}$ on V. Define $\tilde{P}$ on $\mathscr{P}(\mathbb{R}^\infty)$ by $\tilde{P}(E) = \tilde{L}(I_E)$ for all $E \subseteq \mathbb{R}^\infty$. It is clear from (1) that $\tilde{P}$ preserves the independence of $X_1, X_2, \cdots$. □

An alternate proof of Proposition 3.2 may be based on the notions of strategy and inductively integrable function, as in Dubins and Savage (1976), Chapter 2.

As the next example shows, not all extensions $\tilde{P}$ of P preserve the independence of X_1, $X_2, \cdots$.

EXAMPLE 3.3 Let $X_1, X_2, \cdots$ be i.i.d. uniformly distributed on [0, 1]. Let S be any non-Lebesgue-measurable subset of [0, 1] with inner measure zero and outer measure 1, and let $\tilde{P}$ be any extension of P satisfying $\tilde{P}(S \times S \times \mathbb{R}^\infty) = \tilde{P}(S^c \times S^c \times \mathbb{R}^\infty) = 1/2$. Clearly $\tilde{P}$ does not preserve the independence of X_1 and X_2.

THEOREM 3.4. *If $\tilde{P}$ preserves the independence of $X_1, \cdots, X_n$, then for every $s \in \mathscr{S}_n$ there is a $t \in \mathscr{T}_n$ satisfying $EX_t = \tilde{E}X_s$. Moreover, t may be chosen in $A(s)$, where $A(s)$ is defined below.*

The proof of Theorem 3.4 relies upon four lemmas.

LEMMA 3.5. *If $\tilde{P}$ preserves the independence of $X_1, X_2, \cdots, X_n$ and if $f: \mathbb{R}^{n-k} \to \mathbb{R}$ is a Borel measurable function with $E[f(X_{k+1}, \cdots, X_n)] < \infty$, then*

$$\tilde{E}[f(X_{k+1}, \cdots, X_n) \cdot I_B(X_1, \cdots, X_k)] \tag{2}$$
$$= \tilde{P}(B \times \mathbb{R}^\infty) \cdot E[f(X_{k+1}, \cdots, X_n)] \quad \textit{for all subsets } B \text{ of } \mathbb{R}^k.$$

PROOF. Since $-|f| \leq f \cdot I_B \leq |f|$, then by Lemma 2.2(b), $f \cdot I_B \in \mathscr{E}$. To establish (2), the standard argument for the case where B is Borel and $X_1, \cdots, X_n$ are independent is easily extended. □

DEFINITION 3.6. For $s \in \mathscr{S}_n$, $A(s)$ is the set of measurable stopping functions (stopping no later than n) which stop on precisely the same P-atoms as s; that is, $A(s) = \{t \in \mathscr{T}_n$: if $P(\{r_1\} \times \cdots \times \{r_j\} \times \mathbb{R}^\infty) > 0$, then $t(r_1, r_2, \cdots) = j$ if an only if $s(r_1, r_2, \cdots) = j\}$.

LEMMA 3.7. *If $\tilde{P}$ preserves the independence of $X_1, \cdots, X_n$, then for every $s \in \mathscr{S}_n$ there exist t_1 and t_2 in $A(s)$ satisfying*

$$EX_{t_1} \leq \tilde{E}X_s \leq EX_{t_2}. \tag{3}$$

PROOF. Fix $s \in \mathscr{S}_n$. By Proposition 2.6 (b), $X_s \in \mathscr{E}$. Let $\alpha(r_1, r_2, \cdots) = \inf\{k : P(\{r_1\} \times \cdots \times \{r_k\} \times \mathbb{R}^\infty) = 0\}$, and let $t_0 = \alpha \wedge s$. It is easy to check that not only is t_0 measurable, but in fact $t_0 \in A(s)$.

Let $V_n = EX_n$, and define non-decreasing real numbers $V_{n-1}, \cdots, V_1$ inductively by $V_j = E[\max(X_j, V_{j+1})]$. It is well-known [e.g. Chow, Robbins, and Siegmund (1971), page 50] that the measurable stopping function $t^* \in \mathscr{T}_n$ defined inductively by $t^*(r_1, r_2, \cdots) = j$ if and only if $t^*(r_1, r_2, \cdots) > j - 1$ and $r_j > V_{j+1}$ and $j \geq t_0(r_1, r_2, \cdots)$ is "optimal" in the class $\{t \in \mathscr{T}_n : t \geq t_0\}$. The optimality of t^* extends to the class of *non*-measurable stopping functions as well; indeed, Lemma 3.5 (together with backward induction) shows that if $G = \{s \geq \alpha\}$, then $\tilde{E}(X_s \cdot I_G) \leq \tilde{E}(X_{t^*} \cdot I_G)$. Define t_2 by $t_2 = s$ if $s < \alpha$, and $= t^*$ otherwise, and check that $t_2 \in A(s)$. Then

$$\tilde{E}X_s = \tilde{E}(X_s \cdot I_G) + \tilde{E}(X_s \cdot I_{G^c}) \leq \tilde{E}(X_{t^*} \cdot I_G) + \tilde{E}(X_s \cdot I_{G^c}) = EX_{t_2},$$

and the upper bound for (3) is established. To obtain the lower bound, use symmetry, replacing X_i by $-X_i$. □

The next lemma, a result of Liapounoff [Diestel and Uhl (1977), page 261], is stated here for ease of reference.

LEMMA 3.8. (Liapounoff Convexity Theorem). *The range of every non-atomic, countably additive, finite-dimensional, vector-valued measure is convex.*

LEMMA 3.9. *For every $s \in \mathscr{S}_n$, the mapping defined on $A(s)$ by $t \to EX_t$ has a convex range.*

PROOF. Fix $s \in \mathscr{S}_n$. First, using an argument similar in part to a proof of Blackwell (1951), it will be shown that

(4) If $\mu_1, \cdots, \mu_n$ are countably additive, finite signed measures on $(\mathbb{R}^n, \mathbb{B}^n)$, each of which is absolutely continuous with respect to P_n (the distribution of $X_1, \cdots, X_n$), then the mapping defined on $A(s)$ by $t \mapsto \sum_{j=1}^n \mu_j[t^{-1}(j)]$ has a convex range.

To establish (4), let t_1 and t_2 be in $A(s)$, let $\gamma \in (0, 1)$, and for $j = 1, 2, \cdots, n$, let $A_j = t_1^{-1}(j)$ and $B_j = t_2^{-1}(j)$. The aim is to find $t \in A(s)$ satisfying

$$\sum_{j=1}^n \mu_j[t^{-1}(j)] = \gamma \sum_{j=1}^n \mu_j(A_j) + (1-\gamma) \sum_{j=1}^n \mu_j(B_j). \tag{5}$$

If $1 \le k \le n$, let $\mathscr{F}_k$ be the sub σ-algebra of $\mathbb{B}^n$ consisting of those sets of the form $G \times \mathbb{B}^{n-k}$, where $G \in \mathbb{B}^k$. Let $\mathscr{G}$ be the σ-algebra of events prior to time $t_1 \wedge t_2$. That is, define

$$\mathscr{G} = \{G \in \mathbb{B}^n\colon \text{ for each } \ k\ (1 \le k \le n), \qquad G \cap [t_1 \wedge t_2 = k] \in \mathscr{F}_k\}.$$

Let ρ be the $n \times n$-matrix-valued measure defined on $(\mathbb{R}^n, \mathscr{G})$ by

$$(\rho(G))_{jk} = (\mu_j - \mu_k)(G \cap A_j \cap B_k), \qquad 1 \le j \le n, \qquad 1 \le k \le n.$$

To show that ρ is non-atomic, it is first claimed that every C (in $\mathscr{G}$) of the form

$$C = \{r_1\} \times \{r_2\} \times \cdots \times \{r_\ell\} \times \mathbb{R}^{n-\ell}, \text{ where } (r_1, r_2, \cdots, r_n) \in \mathbb{R}^n \text{ and } (t_1 \wedge t_2)(r_1, r_2, \cdots, r_n) = \ell \tag{6}$$

has ρ-measure 0. To see this, notice first that the (j, j)th entry of the matrix $\rho(C)$ is zero for each j. Also observe that $P_n(C \cap A_j \cap B_k) = 0$ for $j \ne k$, because t_1 and t_2 are both in $A(s)$ (and thus stop on precisely the same P_n-atoms). Then use the absolute continuity of $\mu_1, \mu_2, \cdots, \mu_n$ to conclude $\rho(C)$ is the zero matrix.

Suppose now, by way of contradiction, that $\mathscr{G}$ has a ρ-atom A; that is: $A \in \mathscr{G}$, $\rho(A) \ne 0$, and if $B \subseteq A$ for some $B \in \mathscr{G}$, then $\rho(B) = \rho(A)$ or $\rho(B) = 0$. For each $n \ge 1$, let $\{I_{n,m}\}_{m=-\infty}^{\infty}$ be the (countable) collection of closed intervals of the form $[m/2^n, (m+1)/2^n]$. Since A is an atom, and $A = \bigcup_{m=-\infty}^{\infty} A \cap [I_{n,m} \times \mathbb{R}^{n-1}]$, there exist $m_1, m_2, \cdots$ with $I_{1,m_1} \supseteq I_{2,m_2} \supseteq \cdots$ satisfying $\rho(A \cap [I_{i,m_i} \times \mathbb{R}^{n-1}]) = \rho(A)$ for all i. Hence there exists $r_1 \in \mathbb{R}$ with $\rho(A \cap [\{r_1\} \times \mathbb{R}^{n-1}]) = \rho(A)$. If $(t_1 \wedge t_2)(r_1, r_2, \cdots) = 1$, then $\{r_1\} \times \mathbb{R}^{n-1} \in \mathscr{G}$ is of the form (6), and so has ρ-measure zero, contradicting the assumption that $\rho(A) \ne 0$. If $(t_1 \wedge t_2)(r_1, r_2, \cdots) > 1$, proceeding as above one has the existence of $r_2 \in \mathbb{R}$ with $\rho(A) = \rho(A \cap [\{r_1\} \times \{r_2\} \times \mathbb{R}^{n-2}])$. If $(t_1 \wedge t_2)(r_1, r_2, \cdots) = 2$, then again $\{r_1\} \times \{r_2\} \times \mathbb{R}^{n-2} \in \mathscr{G}$ is of the form (6), so $\rho(A) = 0$. Otherwise continue, if necessary, until $(t_1 \wedge t_2)(r_1, r_2, \cdots, r_n) = n$, concluding that $\rho(A) = 0$, and thus that ρ is non-atomic.

Thus Lemma 3.8 applies, and there exists $D \in \mathscr{G}$ such that $\rho(D) = \gamma\rho(\mathbb{R}^n)$.

It is easy to see that the map $t\colon \mathbb{R}^n \to \mathbb{N}$ defined by $t = j$ on $(D \cap A_j) \cup (D^c \cap B_j)$ is in $\mathscr{T}_n$, and in fact even in $A(s)$. To establish (5), calculate

$$\begin{aligned}
\sum_{j=1}^n \mu_j[t^{-1}(j)] &= \sum_j \mu_j(D \cap A_j) + \sum_k \mu_k(D^c \cap B_k) \\
&= \sum_j \sum_k \mu_j(D \cap A_j \cap B_k) + \sum_k \sum_j \mu_k(D^c \cap A_j \cap B_k) \\
&= \sum_j \sum_k (\mu_j - \mu_k)(D \cap A_j \cap B_k) + \sum_k \sum_j \mu_k(A_j \cap B_k) \\
&= \sum_j \sum_k \gamma(\mu_j - \mu_k)(A_j \cap B_k) + \sum_k \sum_j \mu_k(A_j \cap B_k) \\
&= \sum_j \sum_k \gamma\mu_j(A_j \cap B_k) + \sum_j \sum_k (1-\gamma)\mu_k(A_j \cap B_k) \\
&= \gamma\left(\sum_{j=1}^n \mu_j(A_j)\right) + (1-\gamma)\left(\sum_{k=1}^n \mu_k(B_k)\right).
\end{aligned}$$

To complete the proof of the lemma, notice that the measures $\mu_1, \cdots, \mu_n$ defined on $(\mathbb{R}^n, \mathbb{B}^n)$ by $\mu_j(A) = E[X_j \cdot I_A(X_1, \cdots, X_n)]$ are each absolutely continuous with respect to P_n, and apply (4) and the fact that $EX_t = \sum_{j=1}^n \mu_j[t^{-1}(j)]$ for all $t \in \mathscr{T}_n$. □

PROOF OF THEOREM 3.4. Fix $s \in \mathscr{S}_n$ and find $t_1, t_2 \in A(s)$ as in Lemma 3.7. By Lemma 3.9, there exists $t \in \mathscr{T}_n$ (in fact $t \in A(s)$) satisfying $EX_t = \tilde{E}X_s$. □

It is not necessary that $\tilde{P}$ preserve the independence of $X_1, \cdots, X_n$ in order for the conclusion of Theorem 3.4 to hold, as the following example shows.

EXAMPLE 3.10. Let X_1, X_2, and $\tilde{P}$ be as in Example 3.3, and take $\hat{X}_1 = X_1$, $\hat{X}_2 = X_2 + 1$. Clearly $E\hat{X}_2 = \sup\{\tilde{E}\hat{X}_s : s \in \mathscr{S}_2\}$ and $E\hat{X}_1 = \inf\{\tilde{E}\hat{X}_s : s \in \mathscr{S}_2\}$, so the conclusion of Theorem 3.4 holds by Lemma 3.9.

On the other hand, it is possible that $\tilde{E}X_s > \sup\{EX_t : t \in \mathscr{T}_n\}$ for some $s \in \mathscr{S}_n$ if $\tilde{P}$ does not preserve the independence of $X_1, \cdots, X_n$. This will be seen in the proof of the next theorem, which states that if the $\{X_i\}$ are non-negative, the optimal expected gain using measurable stopping functions is always at least half that using arbitrary stopping functions, regardless of the extension $\tilde{P}$.

THEOREM 3.11. *If $X_1, X_2, \cdots, X_n$ are non-negative independent random variables, then $\sup\{\tilde{E}X_s : s \in \mathscr{S}_n\} \le 2 \sup\{EX_t : t \in \mathscr{T}_n\}$ for all n and all extensions $\tilde{P}$ of P. Moreover, this bound is best possible for all $n > 1$.*

PROOF. Since $X_s \le \max\{X_1, \cdots, X_n\}$ for all $s \in \mathscr{S}_n$, it follows from the measurability and integrability of $\max\{X_1, \cdots, X_n\}$ that $\tilde{E}X_s \le \tilde{E}(\max\{X_1, \cdots, X_n\}) = E(\max\{X_1, \cdots, X_n\})$ for all extensions $\tilde{P}$ of P. The desired inequality is now easily derived from the following "prophet" inequality [Krengel and Sucheston (1978), Hill and Kertz (1981a)]:

$$E(\max\{X_1, \cdots, X_n\}) \le 2 \sup\{EX_t : t \in \mathscr{T}_n\}.$$

To show this bound "2" is best possible for $n > 1$, fix $\varepsilon \in (0, 1)$ and let $X_1, X_2, \cdots, X_n$ be independent, X_1 uniform on $[1, 1 + \varepsilon]$, X_2 discrete with $P(X_2 = 1/\varepsilon) = 1 - P(X_2 = 0) = \varepsilon$, and $X_i \equiv 0$ for $i > 2$. Let S be any non-Lebesgue-measurable subset of $[1, 1 + \varepsilon]$ with inner measure 0 and outer measure ε, and let $\tilde{P}$ be any extension of P satisfying $\tilde{P}(S \times \{0\} \times \mathbb{R}^\infty) = 1 - \varepsilon$ and $\tilde{P}(S^c \times \{1/\varepsilon\} \times \mathbb{R}^\infty) = \varepsilon$. Since $EX_2 = 1$, is clear that $\sup\{EX_t : t \in \mathscr{T}_n\} = EX_1 = 1 + \varepsilon/2$. Let $s \in \mathscr{S}_2$ be defined by $s = 1$ on $S \times \mathbb{R}^{n-1}$, and $= 2$ otherwise. Then $\tilde{E}X_s = \tilde{E}[X_1 \cdot I_S(X_1)] + \tilde{E}[X_2 \cdot I_{S^c}(X_1)] \ge 1(1 - \varepsilon) + (1/\varepsilon) = 2 - \varepsilon$. Letting $\varepsilon \to 0$ shows the bound "2" is best possible. □

A simple modification of an example in Hill and Kertz (1981a) shows that if the $\{X_j\}$ are not non-negative, the conclusion of Theorem 3.11 does not hold in general; in fact, for each $M > 0$, one may find an example with $n = 2$ satisfying $\tilde{E}X_s > M \sup\{EX_t : t \in \mathscr{T}_2\}$ for some $s \in \mathscr{S}_2$.

If the independent random variables $X_1, X_2, \cdots, X_n$ are uniformly bounded, the differences between the optimal expected gains of non-measurable and measurable stopping functions is no more than one-fourth the "spread".

THEOREM 3.12. *If $X_1, X_2, \cdots, X_n$ are independent and take on values only in $[a, b]$, then $\sup\{\tilde{E}X_s : s \in \mathscr{S}_n\} - \sup\{EX_t : t \in \mathscr{T}_n\} \le (b - a)/4$ for all extensions $\tilde{P}$ of P, and this bound is best possible for all $n > 1$.*

PROOF. The inequality follows, as in the proof of Theorem 3.11, from another "prophet" inequality [Hill and Kertz (1981b), Theorem A], namely,

$$E(\max\{X_1, \cdots, X_n\}) - \sup\{EX_t : t \in \mathscr{T}_n\} \le (b - a)/4.$$

To show this bound is best possible for $n > 1$, fix ε in (0, 1), and let $X_1, X_2, \cdots, X_n$ be independent with X_1 uniform on $[½, ½ + \varepsilon]$, X_2 discrete with $P(X_2 = 0) = P(X_2 = 1) = ½$, and $X_i \equiv 0$ for $i > 2$. Let S be any non-(Lebesgue)-measurable subset of $[½, ½ + \varepsilon]$ with inner measure 0 and outer measure ε, and let $\tilde{P}$ be any extension of P satisfying $\tilde{P}(S \times \{0\}$

$\times\, \mathbb{R}^\infty) = \tilde{P}(S^c \times \{1\} \times \mathbb{R}^\infty) = 1/2$. Since $EX_2 = 1/2$, clearly $\sup\{EX_t : t \in \mathcal{T}_n\} = EX_1 = 1/2 + \varepsilon/2$. Define $s \in \mathcal{S}_n$ by $s = 1$ on $S \times \mathbb{R}^{n-1}$, and $= 2$ otherwise. Then $\tilde{E}X_s = \tilde{E}[X_1 \cdot I_S(X_1)] + \tilde{E}[X_2 \cdot I_{S^c}(X_1)] \geq 1/2\,(1/2) + 1/2 = 3/4$. Since ε was arbitrary, taking $a = 0$ and $b = 1$ completes the proof. □

4. The infinite-stage stopping problem. As in the previous section, $X_1, X_2, \cdots$ are independent integrable random variables on $(\Omega, \mathcal{A}, \mu)$.

THEOREM 4.1. *If $-\infty < E(\inf X_j) \leq E(\sup X_j) < \infty$ and $\tilde{P}$ preserves the independence of $X_1, X_2, \cdots$, then for every s in $\mathcal{S}$ and every $\varepsilon > 0$, there exists t in $\mathcal{T}$ satisfying*

$$|\tilde{E}X_s - EX_t| < \varepsilon.$$

For the proof of Theorem 4.1, a weakened version of a dominated convergence theorem is helpful:

LEMMA 4.2. *If $s \in \mathcal{S}$ and $-\infty < E(\inf X_j) \leq E(\sup X_j) < \infty$, then*

$$\liminf_{t\to\infty} \tilde{E}X_{s\wedge t} \leq \tilde{E}X_s \leq \limsup_{t\to\infty} \tilde{E}X_{s\wedge t}, \tag{7}$$

where the liminf *and* limsup *are taken over the directed set of bounded, measurable stopping functions t.*

PROOF. By Proposition 2.6, $X_s \in \mathcal{E}$ and $X_{s\wedge t} \in \mathcal{E}$ for each bounded, measurable stopping function t. Fix $\varepsilon > 0$, and let t_0 be a bounded, measurable stopping function. Since $\sup X_j$ is integrable, there exists $\delta > 0$ such that if $A \in \mathcal{A}$ and $\mu(A) < \delta$ then $E(|\sup X_j| \cdot I_A) < \varepsilon/4$. By the independence of $X_1, X_2, \cdots$ and the Kolmogorov Zero-One Law, the random variable $\limsup_{n\to\infty} X_n$ is constant almost-surely; let L^* denote this constant. Choose M in $\mathbb{N}$ such that $t_0 \leq M$ and

$$\mu[\sup_{j\geq M} X_j < L^* + \varepsilon/4] > 1 - \delta/2.$$

Also, choose $N \geq M$ such that

$$\mu[\sup_{M\leq j\leq N} X_j > L^* - \varepsilon/4] > 1 - \delta/2.$$

Let $t'(r_1, r_2, \cdots) = \inf\{j : M \leq j \text{ and } r_j > L^* - \varepsilon/4\}$, and let $t = t' \wedge N$.

Then $t \in \mathcal{T}_N$ and $t \geq t_0$. Letting $B = \{\omega : s(X_1(\omega), X_2(\omega), \cdots) < t(X_1(\omega), X_2(\omega), \cdots)\}$, calculate

$$\begin{aligned}\tilde{E}X_s &= \tilde{E}(X_s \cdot I_B) + \tilde{E}(X_s \cdot I_{B^c}) \leq \tilde{E}(X_s \cdot I_B) + \tilde{E}([\sup_{j\geq M} X_j] \cdot I_{B^c}) \\ &\leq \tilde{E}(X_s \cdot I_B) + \tilde{E}(X_t \cdot I_{B^c \cap [\sup_{j\geq M} X_j < L^* + \varepsilon/4] \cap [\sup_{M\leq j\leq N} X_j > L^* - \varepsilon/4]}) + 3\varepsilon/4 \\ &\leq \tilde{E}(X_s \cdot I_B) + \tilde{E}(X_t \cdot I_{B^c}) + \varepsilon = \tilde{E}(X_{s\wedge t}) + \varepsilon.\end{aligned}$$

This proves the second inequality of (7). The first inequality follows from the second by replacing X_j with $-X_j$, for each j in $\mathbb{N}$. □

REMARK. If it is further assumed in Lemma 4.2 that $s \wedge t \to s$ in $\tilde{P}$-measure as $t \to \infty$, then the net $\{\tilde{E}X_{s\wedge t}\}$ converges to $\tilde{E}X_s$. This conclusion, which is stronger than (7), follows from Dunford and Schwartz [(1958), Theorem III.3.7]. However, in general, $s \wedge t$ need not converge to s in $\tilde{P}$-measure, and $\{\tilde{E}X_{s\wedge t}\}$ need not converge, as $t \to \infty$.

PROOF OF THEOREM 4.1. Let $s \in \mathcal{S}$, $\varepsilon > 0$. Using the hypothesis, there exists $\delta > 0$ such that if $A \in \mathcal{A}$ and $\mu(A) < \delta$, then $E(|\sup X_j| \cdot I_A) < \varepsilon/4$ and $E(|\inf X_j| \cdot I_A) < \varepsilon/4$. Let $G_k = \{(r_1, r_2, \cdots) \in \mathbb{R}^\infty : s(r_1, r_2, \cdots) > k \text{ and } P(\{r_1\} \times \cdots \times \{r_k\} \times \mathbb{R}^\infty) > 0\}$. It is easy to see that $G_k \in B^\infty$ for all k (since it is a subset of the countable collection of P-atoms), and that the sets $F_k \in \mathscr{A}$ defined by $F_k = \{\omega : (X_1(\omega), X_2(\omega), \cdots) \in G_k\}$ decrease to $\varnothing$, so there exists $k_0 \in N$ such that $\mu(F_{k_0}) < \delta$, and n in N such that $n \geq t_2 \geq t_1 \geq k_0$ and

$$\tilde{E}X_{s\wedge t_1} - \varepsilon/2 < \tilde{E}X_s < \tilde{E}X_{s\wedge t_2} + \varepsilon.$$

Now modify $s \wedge t_1$ slightly (seeking a stopping function s' having "the same atoms as" $s \wedge t_2$) by defining: $s'(r_1, r_2, \cdots) = (s \wedge t_1)(r_1, r_2, \cdots)$ if $(s \wedge t_1)(r_1, r_2, \cdots) = k$ and $P(\{r_1\} \times \cdots \times \{r_k\} \times \mathbb{R}^\infty) = 0$; and otherwise $s' = s \wedge t_2$. Clearly $A(s') = A(s \wedge t_2)$ and, since $\mu(F_{k_0}) < \delta$ and $t_1 \geq k_0$, it follows that $\tilde{E}X_{s'} \leq \tilde{E}X_{s\wedge t_1} + \varepsilon/2$.

Applying Theorem 3.4, find t_3 and t_4 in $A(s \wedge t_2)$ which satisfy $EX_{t_3} = \tilde{E}X_{s'}$ and $EX_{t_4} = \tilde{E}X_{s\wedge t_2}$. Then

$$EX_{t_3} - \varepsilon < \tilde{E}X_s < EX_{t_4} + \varepsilon,$$

and since both t_3 and t_4 are in $A(s \wedge t_2) \subseteq \mathscr{T}$, the desired conclusion follows easily from Lemma 3.9. □

Remark. If $X_1, X_2, \cdots$ are independent, then the bounds "2" and "$(b - a)/4$" which were established for finite sequences $X_1, X_2, \cdots, X_n$ in Theorems 3.11 and 3.12, respectively, also hold for infinite sequences. Thus, even if $\tilde{P}$ does not preserve independence, non-measurable stopping functions do not yield "too much more" than measurable ones.

As the next example illustrates, the approximation of $\tilde{E}X_s$ by EX_t in Theorem 4.1 cannot be strengthened to obtain equality between $\tilde{E}X_s$ and EX_t.

Example 4.3. Let Ω be the interval [0, 1], $\mathscr{A}$ be the Borel sets, and μ be Lebesgue measure on [0, 1]. Let $X_1(\omega) = \omega$ for each ω in Ω, and for $n > 1$, let $X_n \equiv 1 - (1/n)$. Decompose Ω into disjoint $A_1, A_2, \cdots$ such that each A_n has Lebesgue inner measure 0 and outer measure 1. Extend P to $\tilde{P}$ in such a way that $\tilde{P}(\cup_{i=k}^\infty A_i) = 1$ for each k. Define s by letting $s(X_1(\omega), X_2(\omega), \cdots) = k$ if $X_1(\omega) \in A_k$. Then $\tilde{E}X_s = 1$, but $EX_t < 1$ for each measurable stopping function t.

5. Dependent sequences. The purpose of this section is to comment on the comparison between expected gains using measurable and non-measurable stopping functions in the case where $X_1, X_2 \cdots$ are arbitrarily-dependent integrable random variables.

The authors believe that, as in the independent case, the expected gain for every bounded non-measurable stopping function is equal to that for a measurable one, provided the extension $\tilde{P}$ satisfies certain properties analogous to, but considerably less simple or natural than "preserving independence", and that, under these conditions, the corresponding approximation for unbounded non-measurable stopping times by measurable ones also is valid.

By using outer integrals to evaluate non-measurable plans, Blackwell, Freedman, and Orkin (1974) have shown that, in a finite-stage dynamic programming framework allowing non-measurable transitions, measurable plans do just as well.

For special types of measures and extensions (strategic measures), results of Sudderth [(1971), Section 4] and Dubins and Sudderth [(1979), Corollary 4.1] imply that for the infinite-stage problem, one can do as well with measurable as with non-measurable stopping functions.

If $\tilde{P}$ is a completely arbitrary extension, though, it is possible that $\tilde{E}X_s > \sup\{EX_t : t \in \mathscr{T}_n\}$ for some $s \in \mathscr{S}_n$. However, the expected gain from a nonmeasurable stopping function is never more than n times the expected gain from an optimal measurable stopping function if the $\{X_i\}$ are nonnegative, as the next theorem shows.

Theorem 5.1. *If $n \in N$, and if $X_1, \cdots, X_n$ are non-negative, then*

$$\sup\{\tilde{E}X_s : s \in \mathscr{S}_n\} \leq n \sup\{EX_t : t \in \mathscr{T}_n\} \tag{8}$$

for all extensions $\tilde{P}$ of P, and this bound is best possible.

PROOF. To establish (8), let $s \in \mathscr{S}_n$ and, for $1 \leq k \leq n$, let $B_k = \{\omega : s(X_1(\omega), \cdots, X_n(\omega)) = k\}$. Then

$$\tilde{E}X_s = \sum_{k=1}^n \tilde{E}(X_k \cdot I_{B_k}) \leq \sum_{k=1}^n EX_k \leq n \sup\{EX_t : t \in \mathscr{T}_n\}.$$

To show the bound is best possible, consider first the case $n = 3$ ($n = 1$ is trivial, and $n = 2$ is a consequence of Theorem 3.11). Define X_1, X_2, X_3 jointly by: (X_1, X_2, X_3) is uniform on $[1, 1 + \varepsilon] \times [0, \varepsilon] \times [0, \varepsilon]$ with probability $1 - \varepsilon$; uniform on $[1, 1 + \varepsilon] \times [1/\varepsilon, 1/\varepsilon + \varepsilon] \times [0, \varepsilon]$ with probability $\varepsilon - \varepsilon^2$; and is uniform on $[1, 1 + \varepsilon] \times [1/\varepsilon, 1/\varepsilon + \varepsilon] \times [1/\varepsilon^2, 1/\varepsilon^2 + \varepsilon]$ with probability ε^2. It is easy to see that $\sup\{EX_t : t \in \mathscr{T}_3\} \to 1$ *as* $\varepsilon \searrow 0$.

Next, let $S_1 \subset [1, 1 + \varepsilon]$ and $S_2 \subset [1/\varepsilon, 1/\varepsilon + \varepsilon]$ be non-(Lebesgue) measurable sets with inner measure 0 and outer measure ε. Let $\tilde{P}$ be any extension of P satisfying $\tilde{P}(S_1 \times [0, \varepsilon] \times [0, \varepsilon]) = 1 - \varepsilon$, $\tilde{P}(S_1^c \times S_2 \times [0, \varepsilon]) = \varepsilon - \varepsilon^2$ and $\tilde{P}(S_1^c \times S_2^c \times [1/\varepsilon^2, 1/\varepsilon^2 + \varepsilon]) = \varepsilon^2$. Let $s \in \mathscr{S}_3$ be given by $s = 1$ on $S_1 \times \mathbb{R}^2$, $s = 2$ on $S_1^c \times S_2 \times \mathbb{R}$, and $s = 3$ otherwise. It is easy to verify that $\tilde{E}X_s \to 3$ as $\varepsilon \searrow 0$.

The proof for general n is the analog of this case with values $1, 1/\varepsilon, 1/\varepsilon^2, \cdots, 1/\varepsilon^{n-1}$ and probabilities $1 - \varepsilon, \varepsilon - \varepsilon^2, \cdots, \varepsilon^{n-2} - \varepsilon^{n-1}, \varepsilon^{n-1}$ replacing the given ones. □

Acknowledgments. The authors are grateful to Michael Klass for a suggestion regarding the proof of Theorem 5.1, to N. Bouzar for pointing out an error in an earlier version of Theorem 3.4, and to the referee for various suggestions and corrections.

REFERENCES

[1] BLACKWELL, D. (1951). On a theorem of Lyapunov. *Ann. Math Statist.* **22** 112–114.

[2] BLACKWELL, D., FREEDMAN, D., and ORKIN, M. (1974). The optimal reward operator in dynamic programming. *Ann. Probability* **2** 926–941.

[3] CHOW, Y. S., ROBBINS, H., and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin, Boston.

[4] DIESTEL, J. and UHL, J. J. JR. (1977). Vector Measures. Mathematical Surveys, No. 15. American Mathematical Society, Providence.

[5] DUBINS, L. E. and SAVAGE, L. J. (1976). *Inequalities for Stochastic Processes (How to Gamble if You Must).* Dover, New York.

[6] DUBINS, L. E. and SUDDERTH, W. D. (1979). On stationary strategies for absolutely continuous houses. *Ann. Probability* **7** 461–476.

[7] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators, Part I.* Interscience, New York.

[8] HILL, T. P. and KERTZ, R. P. (1981a). Ratio comparisons of supremum and stop rule expectations. *Z. Wahrsch. verw. Gebiete* **56** 283–285.

[9] HILL, T. P. and KERTZ, R. P. (1981b). Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables. *Proc. Amer. Math. Soc.* **83** 582–585.

[10] KRENGEL, U. and SUCHESTON, L. (1978). On semiamarts, amarts, and processes with finite value. In: *Probability on Banach Spaces, Advances in Probability* **4** 197–266. Dekker, New York.

[11] PURVES, R. A. and SUDDERTH, W. D. (1976). Some finitely additive probability. *Ann. Probability* **4** 259–276.

[12] SUDDERTH, W. D. (1971). A gambling theorem and optimal stopping theory. *Ann. Math. Statist.* **42** 1697–1705.

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MIAMI
P.O. BOX 249085
CORAL GABLES, FLORIDA 33124