Discussion Paper No. 26

ON RISK AVERSION AND OPTIMAL STOPPING

by

David C. Nachman

November 30, 1972

DEPARTMENT OF MANAGERIAL ECONOMICS AND DECISION SCIENCES
GRADUATE SCHOOL OF MANAGEMENT
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60201

1. INTRODUCTION

Risk aversion, the unwillingness of an individual to accept a gamble which is actuarilly fair, has played a significant role in the development of the economics of uncertainty. Indeed, as Arrow points out, the risk aversion hypothesis offers a satisfactory explanation of "...otherwise puzzling examples of economic behavior," including the buying of insurance and the holding of wealth in the form of non-interest-bearing cash balances. [1/] Such explanations constitute the qualitative results in the economics of uncertainty.

To obtain quantitative results in this area it is necessary to postulate how risk averse an individual is or how his aversion to risk changes with certain parameters of the problem under investigation. For example, risk aversion can tell us that an individual will hold cash balances, but to investigate the change in these holdings with say an increase in total wealth requires knowledge of how the individual's aversion to risk changes as he becomes wealthier. In 1964 and 1965, Pratt [26] and Arrow [1] introduced measures of absolute and relative risk aversion that permit comparisons of individuals, one of whom is "more risk averse" than the other, and comparison of the behavior of the same individual at varying levels of wealth. Since then these measures have been employed to derive quantitative properties of optimal policies in a variety of decision problems under uncertainty. [2/]

We continue the spirit of these investigations in this paper by studying a class of sequential decision problems known as optimal stopping problems. [3/] This class of problems is interesting for two reasons. The first and most obvious is that the problem of timing the implementation of a given action can be modeled as an optimal stopping problem. From optimality conditions for the stopping problem it is then possible to deduce the effect on optimal timing strategies of changes in the degree of risk aversion of the decision maker or changes in his initial wealth given his risk preferences.

The second reason is more fundamental. In all interesting decision problems, whether they are characterized by certainty or uncertainty, choice involves the foregoing of alternatives and hence the incurring of opportunity costs. The touchstone of analysis of decision problems under uncertainty is the opportunity cost involved in foregoing an alternative with a certain outcome for one with uncertain outcomes. This opportunity cost is a function of the decision maker's risk preferences and is fundamental to definitions of certainty equivalents and risk premiums. These latter concepts are in turn fundamental to the development and value of the Pratt-Arrow measures of risk aversion.

In sequential decision problems under uncertainty difficulties arise in the study of the trade-off between certainty and uncertainty. First, the risk premium of static analysis does not reflect information obtained in the course of making a sequence of decisions. Secondly, there is, in general, no temporal measure of risk aversion, i.e., there is no way of relating risk preferences in one period to those in the following or preceding periods. [4/]

In attempting to solve these two problems, one would like to study a sequence of decisions that in every period involves a binary choice between an alternative with a certain outcome and one with uncertain outcomes and whose reward and cost structure are sufficiently simple to lay bare these basic alternatives. The optimal stopping problem satisfies these criteria and from a study of this problem we have some candidates for solutions of the above problems. We report in this paper the first part of this study which deals with conditional risk premiums and the utility theoretic formulation of the optimal stopping problem under the assumption that risk preferences are homogeneous with respect to time. 5/

The remainder of the paper is organized as follows. We begin in section 2 with some notation and discuss the concept of absolute risk aversion. We define conditional risk premiums and prove some elementary results regarding them. Conditional risk premiums generalize the notion of static risk premium. They provide a guide for intuition in sequential decision problems and a tool for analyzing expectations of future rewards based on current information.

In section 3, we present the optimal stopping problem (OSP) from a utility theoretic point of view. There we prove a theorem relating absolute risk aversion and optimal timing policies. The theorem states that the more (absolute) risk averse individual will implement a given action sooner than will the less risk averse individual. We prove the theorem using the conditional risk premiums defined in section 2 and also using Pratt's [26, Theorem 1].

In section 4, we apply the results of section 3 to two important questions. The first deals with a class of simplified timing strategies we call one-period myopic rules. The result of section 3 is shown to hold for this class of rules and we give sufficient conditions for a rule in this class to be optimal. The second question deals with the conjecture that the individual who is less risk averse should in some way be rewarded for his greater willingness to bear risk. This conjecture is shown to be false in general.

Throughout sections 3 and 4, the results are interpreted for the problem of the timing of the sale of an asset. In addition, we discuss briefly models of the random walk type for stock prices and their implications for the timing of stock market transactions.

2. ABSOLUTE RISK AVERSION

2.1 Some Preliminaries.

Throughout this paper we will be dealing with random variables defined on a fixed probability space $(\Omega, \mathcal{F}, P)$, where Ω is a set of points, $\mathcal{F}$ a σ-algebra of events (subsets of Ω), and P a probability measure on $\mathcal{F}$. A property which holds except on a set of P-measure zero is said to hold almost surely or almost surely P and is abbreviated a.s. or a.s.P. All random variables will be assumed to be integrable with respect to P.[6/] If X is a random variable on Ω, we denote its expectation by EX. In denoting events such as $\{\omega : X(\omega) = \alpha\}$ we sometimes abbreviate to $\{\omega : X = \alpha\}$ or just $(X = \alpha)$.

If $\{X_\alpha ; \alpha \in A\}$ is a collection of random variables (indexed by the set A), we denote by $\mathcal{F}(X_\alpha ; \alpha \in A)$ the σ-algebra generated by the X_α. Similarly we denote by $\mathcal{F}(X_1, \ldots, X_n)$ (or $\mathcal{F}_n$ where no confusion arises) and by $\mathcal{F}(X)$, the σ-algebras generated by the random variables $X_1, \ldots, X_n$ and X respectively. We denote by $(\mathcal{R}^1, \mathcal{B}_1)$ the measurable space where $\mathcal{R}^1$ is the real line and $\mathcal{B}_1$ the Borel subsets of $\mathcal{R}^1$.

Let X be a random variable and Y a $\mathcal{F}$-measurable function taking values in some denumerable set $\{y_k\}_{k=1}^{\infty}$. The conditional expectation $E(X \mid Y=y_k)$ is defined constructively as

$$E(X \mid Y=y_k) = \frac{1}{P(Y=y_k)} \int_{(Y=y_k)} X \, dP$$

provided $P(Y=y_k) > 0$. Hence

$$\int_{(Y=y_k)} E(X \mid Y=y_k) \, dP = P(Y=y_k) \, E(X \mid Y=y_k) = \int_{(Y=y_k)} X \, dP \,. \qquad (2.1)$$

In many problems, however, we are interested in conditioning expectations on random variables or random vectors Y such that $P(Y=y) = 0$ for all y. 7/ Hence, in general, we result to nonconstructive definitions like (2.1) to handle these more difficult problems.

To be specific, let Y be a random vector taking values in the measurable space $(\mathcal{R}, \mathcal{B})$. We define the conditional expectations $E(X|Y=y)$ and $E(X|Y)$ as any $\mathcal{B}$-measurable, respectively $\mathcal{F}(Y)$-measurable, real valued functions satisfying

$$\int_B E(X|Y=y)\, dP_Y = \int_{(Y\in B)} X\, dP \ , \ \forall\ B \in \mathcal{B} \tag{2.2}$$

$$\int_A E(X|Y)\, dP = \int_A X\, dP \ , \ \forall\ A \in \mathcal{F}(Y) \tag{2.3}$$

where $P_Y(B) = P(Y\in B)$ for all $B\in\mathcal{B}$. Since in (2.3) the only restrictions are that A belong to $\mathcal{F}(Y)$ we may generalize to conditioning on an arbitrary sub-σ-algebra $\mathcal{D}$ of $\mathcal{F}$ by defining $E(X|\mathcal{D})$ as any $\mathcal{D}$-measurable random variable satisfying

$$\int_A E(X|\mathcal{D})\, dP = \int_A X\, dP \ , \ \forall\ A \in \mathcal{D} \tag{2.4}$$ 8/

Existence of the functions required in (2.2) - (2.4) is guaranteed by the Radon-Nikodym Theorem [5, p.398] and they are unique up to almost sure equivalence (a.s.P in (2.3) and (2.4) and a.s.P_Y in (2.2)). We will use these conditional expectations frequently in the sequel and we use the notation $E(X|Y)$, $E(X|\mathcal{D})$, $E(X|Y=y)$ to denote the equivalence class and any element thereof. Such elements are referred to as versions of the conditional expectation.

2.2 Risk Aversion and Risk Premiums.

By a utility function in this paper we mean any function $\mathcal{U}:\mathcal{R}^1 \to \mathcal{R}^1$ which is strictly increasing and continuous. It follows then that a utility function is $\mathcal{B}_1$ - measurable. Although we are going to use such functions in the sense of von Neumann-Morgenstern expected utility theory, we do not require them to be bounded as many axiom systems imply they must be. 9/ Since such functions are unique only up to a positive linear transformation, we will assume that $\mathcal{U}(0) = 0$ for any utility function.

Let $\mathcal{U}$ be a utility function, X a random variable, and $b \in \mathcal{R}^1$. Suppose $E \mid \mathcal{U}(X) \mid < \infty$. We may think of X as a random monetary reward and b as the initial wealth position. The risk premium $\pi(b,X)$ of X under P is defined by the equation

$$\mathcal{U}(b + EX - \pi(b,X)) = E\,\mathcal{U}(b + X) \qquad (2.5)$$

i.e., as the unique real number π that just makes $\mathcal{U}$ indifferent between receiving $EX - \pi$ for certain and receiving the random reward X. 10/ π is unique since $\mathcal{U}$ is strictly increasing and always exists (although it may be $\pm\infty$) since $\mathcal{U}$ is continuous. π depends on $\mathcal{U}$,b,X, and P, but, since we will not vary P in our analysis of gambles defined on Ω and since dependence of π on $\mathcal{U}$ will be noted with subscripts when more than one utility function is under consideration, the notation $\pi(b,X)$ seems appropriate.

In defining risk aversion and absolute risk aversion properties it is convenient to use the risk premiums defined on $\mathcal{R}^1$. Given $\mathcal{U}$, denote by $\mathcal{P}_{\mathcal{U}}$ the set of probability measures p on $(\mathcal{R}^1, \mathcal{B}_1)$ for which x and $\mathcal{U}(x)$ are integrable. Let $\pi(b,p)$ denote the risk premium defined by (2.5) (when $(\Omega, \mathcal{F}, P) = (\mathcal{R}^1, \mathcal{B}_1, p)$). The x is suppressed in this notation since the gamble is always the coordinate random variable on $(\mathcal{R}^1, \mathcal{B}_1)$. The varying of p in $\mathcal{P}_{\mathcal{U}}$ determines different gambles. Note that for degenerate p, $\pi(b,p) = 0$, $\forall\ b \in \mathcal{R}^1$.

The utility function $\mathcal{U}$ is said to exhibit risk aversion if $\pi(b,p) \geq 0$ for all $p \in \mathcal{P}_{\mathcal{U}}$ and all $b \in \mathcal{R}^1$. This property of $\mathcal{U}$ is easily shown to be equivalent to concavity of $\mathcal{U}$ and is the property of risk aversion mentioned in the introduction. From (2.5) we see that the risk averse individual is willing to accept something less than the actuarial value EX of the gamble for certain rather than face the gamble.

Although risk aversion is a reasonable hypothesis we will not need it for most of our analysis. Instead we will make use of comparisons of utility functions based on $\pi(b,p)$. Given two utility functions $\mathcal{U}_1$ and $\mathcal{U}_2$, we say that $\mathcal{U}_1$ is more (absolutely) risk averse than $\mathcal{U}_2$, and write $\mathcal{U}_1 \geq_r \mathcal{U}_2$, if $\pi_1(b,p) \geq \pi_2(b,p)$ for all $p \in \mathcal{P}_{\mathcal{U}_1} \cap \mathcal{P}_{\mathcal{U}_2}$ and all $b \in \mathcal{R}^1$. We say that the utility function $\mathcal{U}$ exhibits decreasing (increasing)(constant) absolute risk aversion, and write $\mathcal{U}$ is DARA (IARA)(CARA) if $b_1 > b_2$ implies $\pi(b_1,p) \leq (\geq) (=) \pi(b_2,p)$ for each $p \in \mathcal{P}_{\mathcal{U}}$. Again from (2.5) we see that these definitions accord well with the intuitive meaning of the properties inovlved.

The following result due to Pratt [26; Theorem 1] will be used in sections 3 and 4 to derive some of our results. 11/ Let $\mathcal{U}_2^{-1}$, $\mathcal{U}_1^{-1}$ denote the inverses of $\mathcal{U}_2$ and $\mathcal{U}_1$ respectively.

Lemma 1. $\mathcal{U}_1 \geq_r \mathcal{U}_2$ if and only if $\mathcal{U}_1(\mathcal{U}_2^{-1}(t))$ is a concave function of t.

Proof: Note first that we are assuming $\mathcal{U}_1$ and $\mathcal{U}_2$ are defined on all of $\mathcal{R}^1$ so that $\mathcal{U}_1$ is defined on the range of $\mathcal{U}_2^{-1}$ which is all of $\mathcal{R}^1$. Pratt proves the implication $\mathcal{U}_1(\mathcal{U}_2^{-1}(t))$ concave in t implies $\mathcal{U}_1 \geq_r \mathcal{U}_2$ without any differentiability assumptions on $\mathcal{U}_1$ and $\mathcal{U}_2$. Suppose then that $\mathcal{U}_1 \geq_r \mathcal{U}_2$. Let t_1, t_2 be any two numbers in the range of $\mathcal{U}_2$ (the domain of $\mathcal{U}_2^{-1}$) and let x_1, x_2 be the unique numbers such that $\mathcal{U}_2(x_1) = t_1$ and $\mathcal{U}_2(x_2) = t_2$. Given $\lambda \in [0,1]$ denote by p_λ the probability which assigns probability λ to x_1 and $(1-\lambda)$ to x_2. By assumption,

$$\pi_1(0,p_\lambda) = \lambda x_1 + (1-\lambda)x_2 - \mathcal{U}_1^{-1}(\lambda\mathcal{U}_1(x_1) + (1-\lambda)\mathcal{U}_1(x_2))$$

$$\geq \pi_2(0,p_\lambda)$$

$$= \lambda x_1 + (1-\lambda)x_2 - \mathcal{U}_2^{-1}(\lambda\mathcal{U}_2(x_1) + (1-\lambda)\mathcal{U}_2(x_2))$$

and hence

$$\lambda\mathcal{U}_1(\mathcal{U}_2^{-1}(t_1)) + (1-\lambda)\mathcal{U}_1(\mathcal{U}_2^{-1}(t_2)) = \lambda\mathcal{U}_1(x_1) + (1-\lambda)\mathcal{U}_1(x_2)$$

$$\leq \mathcal{U}_1(\mathcal{U}_2^{-1}(\lambda\mathcal{U}_2(x_1) + (1-\lambda)\mathcal{U}_2(x_2)))$$

$$= \mathcal{U}_1(\mathcal{U}_2^{-1}(\lambda t_1 + (1-\lambda)t_2))$$

and we are done.

Although the entire analysis of this paper could be accomplished with no further considerations of risk aversion concepts, we will find it intuitively appealing to use the notion of conditional risk premium. Although Lemma 1 gives us the analytical tool to derive quantitative properties of optimal stopping rules, the condition $\mathcal{U}_1(\mathcal{U}_2^{-1}(t))$ concave in t provides little intuitive insight. We turn now to conditional risk premiums.

2.3 Conditional Risk Premiums

It is often convenient, as will be seen here and in section 3, to condition expectations and therefore risk premiums on $\mathcal{F}$-measurable functions (random variables and random vectors) and on sub-σ-algebras of $\mathcal{F}$. Let X be a random variable, Y a random vector taking values in the measurable space $(\mathcal{R},\mathcal{B})$, and $\mathcal{D}$ a sub-σ-algebra of $\mathcal{F}$. Assuming $E|\mathcal{U}(X)| < \infty$, analogous to (2.5) we define the conditional risk premiums $\pi(b,X \mid \mathcal{D})$ and $\pi(b,X \mid Y = y)$ by

$$\mathcal{U}(b + E(X \mid \mathcal{D}) - \pi(b,X \mid \mathcal{D})) = E(\mathcal{U}(b + X) \mid \mathcal{D}) \qquad (2.6)$$

$$\mathcal{U}(b + E(X \mid Y=y) - \pi(b,X \mid Y=y)) = E\mathcal{U}(b + X) \mid Y = y) \quad (2.7)$$ 12/

Note first that (2.6) defines $\pi(b, X \mid \mathcal{D})$ as the equivalence class (under the relation = a.s.P) of $\mathcal{D}$-measurable random variables Z satisfying

$$Z = b + E(X \mid \mathcal{D}) - \mathcal{U}^{-1}(E\mathcal{U}(b + X \mid \mathcal{D})) \quad (2.8)$$

and that (2.7) defines $\pi(b,X \mid Y = y)$ as the equivalence class (under = a.s.P_Y, where $P_Y(B) = P(Y \in B)$ for all $B \in \mathcal{B}$) of $\mathcal{B}$-measurable real valued functions Z' satisfying

$$Z' = b + E(X \mid Y = y) - \mathcal{U}^{-1}(E\mathcal{U}(b + X) \mid Y = y)) \quad (2.9)$$

We will use the notation $\pi(b,X \mid \mathcal{D})$ and $\pi(b,X \mid Y = y)$ to refer to both the equivalence class and any element thereof. Such an element will be referred to as a version of the conditional risk premium. We define $\pi(b, X \mid Y)$ as $\pi(b, X \mid Y) = \pi(b,X \mid Y = y) \circ Y$ a.s.P (where o denotes composition of functions).

Intuitively, conditioning on the σ-algebra $\mathcal{D}$ or the random vector Y conveys the idea of being able to obtain some information or observation prior to receiving the random reward X. The conditional risk premiums then play the same role as the risk premium of (2.5) given this information. The following lemmas provide further justification, both formal and intuitive, for considering conditional risk premiums.

<u>Lemma 2.</u> 13/ Let $\mathcal{U}$ be a utility function, $\mathcal{D}$ and $\mathcal{D}'$ sub-σ-algebras of $\mathcal{F}$, and X a random variable with $E \mid \mathcal{U}(b + X) \mid < \infty$, $b \in R^1$. Then

(i) if X is $\mathcal{D}$-measurable, then $\pi(b,X \mid \mathcal{D}) = 0$ a.s. for all $b \in R^1$;

(ii) if $\mathcal{F}(X)$ and $\mathcal{D}$ are independent, then $\pi(b,X \mid \mathcal{D}) = \pi(b,X)$ a.s.;

(iii) if $\mathcal{B}' \subset \mathcal{B}$, then $E(\pi(b,X \mid \mathcal{B}') \mid \mathcal{B}) = \pi(b, X \mid \mathcal{B}')$ a.s.;

(iv) if $\mathcal{U}$ is concave, then $\pi(b, X \mid \mathcal{B}) \geq 0$ a.s., and

$E\,\pi(b, X \mid \mathcal{B}) \leq \pi(b,X)$;

(v) if $\mathcal{U}$ is concave and $\mathcal{B}' \subset \mathcal{B}$, then $E(\pi(b, X \mid \mathcal{B}) \mid \mathcal{B}') \leq \pi(b,X \mid \mathcal{B}')$

a.s. and hence $0 \leq E\,\pi(b, X \mid \mathcal{B}) \leq E\,\pi(b, X \mid \mathcal{B}') \leq \pi(b,X)$.

Proof: (i) If X is $\mathcal{B}$-measurable, then $E(\mathcal{U}(b + X) \mid \mathcal{B}) = \mathcal{U}(b + X)$ a.s. and $E(X \mid \mathcal{B}) = X$ a.s. Hence from (2.6),

$$\mathcal{U}(b + E(X \mid \mathcal{B}) - \pi(b, X \mid \mathcal{B})) = \mathcal{U}(b + X) \text{ a.s.}$$

$$\Rightarrow b + X - \pi(b, X \mid \mathcal{B}) = b + X \text{ a.s.}$$

$$\Rightarrow \qquad \pi(b, X \mid \mathcal{B}) = 0 \text{ a.s.}$$

(ii) If $\mathcal{F}(X)$ and $\mathcal{B}$ are independent, then $E(\mathcal{U}(b + X) \mid \mathcal{B}) = E\,\mathcal{U}(b + X)$ a.s. and $E(X \mid \mathcal{B}) = EX$ a.s. Hence, from (2.5) and (2.6),

$$b + EX - \pi(b\ X \mid \mathcal{B}) = b + EX - \pi(b,X) \text{ a.s.}$$

and the conclusion follows.

(iii) If $\mathcal{B}' \subset \mathcal{B}$, then $\pi(b, X \mid \mathcal{B}')$ is $\mathcal{B}$-measurable and hence $E(\pi(b, X \mid \mathcal{B}') \mid \mathcal{B}) = \pi(b, X \mid \mathcal{B}')$ a.s.

(iv) If $\mathcal{U}$ is concave, then

$$\begin{aligned}\pi(b, X \mid \mathcal{B}) &= b + E(X \mid \mathcal{B}) - \mathcal{U}^{-1}(E(\mathcal{U}(b + X) \mid \mathcal{B}))\\ &\geq b + E(X \mid \mathcal{B}) - E(\mathcal{U}^{-1}(\mathcal{U}(b + X)) \mid \mathcal{B})\\ &= b + E(X \mid \mathcal{B}) - b - E(X \mid \mathcal{B}) = 0 \text{ a.s.,}\end{aligned}$$

the inequality following from Jensen's inequality for conditional expectations. 14/ Using Jensen's inequality for expectations gives

$$\begin{aligned}E\,\pi(b, X \mid \mathcal{B}) &= b + EX - E(\mathcal{U}^{-1}(E(\mathcal{U}(b + X) \mid \mathcal{B})))\\ &\leq b + EX - \mathcal{U}^{-1}(E(E(\mathcal{U}(b + X) \mid \mathcal{B})))\\ &= b + EX - \mathcal{U}^{-1}(E\mathcal{U}(b + X))\\ &= \pi(b,X)\end{aligned}$$

where we have used the fact that $E(E(X \mid \mathcal{B})) = EX$ and similarly for $\mathcal{U}(b + X)$.

(v) If $\mathcal{B}' \subset \mathcal{B}$ and $\mathcal{U}$ is concave, again using Jensen's inequality gives

$$\begin{aligned}
E(\pi(b,X \mid \mathcal{B}) \mid \mathcal{B}') &= E[b + E(X \mid \mathcal{B}) - \mathcal{U}^{-1}(E(\mathcal{U}(b + X) \mid \mathcal{B})) \mid \mathcal{B}'] \\
&= b + E(X \mid \mathcal{B}') - E(\mathcal{U}^{-1}(E(\mathcal{U}(b+X) \mid \mathcal{B})) \mid \mathcal{B}') \\
&\leq b + E(X \mid \mathcal{B}') - \mathcal{U}^{-1}(E(E(\mathcal{U}(b + X) \mid \mathcal{B}) \mid \mathcal{B}')) \\
&= b + E(X \mid \mathcal{B}') - \mathcal{U}^{-1}(E(\mathcal{U}(b + X) \mid \mathcal{B}')) \\
&= \pi(b,X \mid \mathcal{B}) \text{ a.s.}
\end{aligned}$$

where we have used the fact that $E(E(X \mid \mathcal{B}) \mid \mathcal{B}') = E(X \mid \mathcal{B}')$ a.s. and similarly for $\mathcal{U}(b + X)$. Hence from (iv) we get $0 \leq E\pi(b,X \mid \mathcal{B}) \leq E\pi(b,X \mid \mathcal{B}') \leq \pi(b,X)$.

To interpret these results, consider a gamble X and a random variable Y that may be observed before the value of X is known. Then $E(X \mid Y = y)$ is an estimate [15/] of the values of X that will obtain given that $Y = y$ has been observed. This estimate is based on knowledge of the event $\{\omega: Y(\omega) = y\}$ since this event is what can be inferred from the observation $Y = y$. Thus it is the events in the σ-algebra generated by Y, $\mathcal{F}(Y)$, that are important for the estimate $E(X \mid Y = y)$. We should expect then that the more events in $\mathcal{F}(Y)$, the better our estimate in some sense or at least the more information we have on which to base our estimate of X.

The results of Lemma 2 bear out this conjecture, but we must be careful to allow for the fact that information affects risk averters differently than say risk lovers (those whose risk premium $\pi(b,P)$ is non-positive and therefore whose utility function is convex). In (i) we have perfect information, i.e., all events that are relevant for the values of X are in $\mathcal{B}$ since $\mathcal{F}(X) \subset \mathcal{B}$.

To make this a little clearer, suppose Y is a random variable that may be observed before X is realized and that $\mathcal{F}(Y) = \mathcal{B}$. Then there is a $\mathcal{B}_1$-measurable function $\varphi: \mathcal{R}^1 \to \mathcal{R}^1$ such that $X = \varphi(Y)$. Thus knowledge of Y is equivalent to knowledge of X and we are in the certainty case. In this light the conclusion of (i) is not so startling.

In (ii) we have the case that $\mathcal{B}$ contains no useful information regarding X since $\mathcal{F}(X)$ and $\mathcal{B}$ are independent. Another example of the case when $\mathcal{B}$ has no relevant information is when $\mathcal{B} = \{\Omega, \emptyset\}$, when $\mathcal{B}$ consists of just Ω and the empty set $\emptyset$. In this case conditing on $\mathcal{B}$ is equivalent to no conditioning since for any random variable X, $E(X \mid \mathcal{B}) = EX$ a.s. This last case reinforces the idea that the size of the conditioning σ-algebra is directly related to the information provided by it (except of course in the independent case).

In case (iii) we interpret $E(\pi(b,X \mid \mathcal{B}') \mid \mathcal{B})$ as an estimate (estimator) of the random risk premium $\pi(b,X \mid \mathcal{B}')$ given information in $\mathcal{B}$, i.e., we begin with information in $\mathcal{B}'$ and then project forward on the basis of information in $\mathcal{B}$. The result in (iii) seems at first to fly in the face of our interpretation of information and the size of the conditioning σ-algebra. $\mathcal{B}$ is larger than $\mathcal{B}'$, but this additional information has no appreciable effect on the estimate. Notice, however, that what we are estimating is the random risk premium $\pi(b,X \mid \mathcal{B}')$ which is $\mathcal{B}'$ and hence $\mathcal{B}$-measurable. $\mathcal{B}$ already contains all events necessary to explain or predict the values of $\pi(b,X \mid \mathcal{B}')$ and hence the estimate is certain, i.e., $E(\pi(b,X \mid \mathcal{B}') \mid \mathcal{B}) - \pi(b,X \mid \mathcal{B}') = 0$ a.s.

The case of risk aversion is interesting in that it brings out the phenomenon that information reduces risk. From (iv) as expected $\pi(b,X \mid \mathcal{B}) \geq 0$, but we have that $E\pi(b,X \mid \mathcal{B}) \leq E\,\pi(b,X) = \pi(b,X)$ and from (v) for

$\mathcal{B}' \subset \mathcal{B}$, $0 \leq E\ \pi(b,X \mid \mathcal{B}) \leq E\ \pi(b,X \mid \mathcal{B}') \leq \pi(b,X)$. Thus the expected value of the non-negative conditional risk premium is bounded above by $\pi(b,X)$ and decreases with increasing information until $\mathcal{B} \supset \mathcal{J}(X)$ at which point from (i) $E\ \pi(b,X \mid \mathcal{B}) = 0$. For one who prefers risk ($\mathcal{U}$ convex) we get the opposite effect $\pi(b,X) \leq E\ \pi(b,X \mid \mathcal{B}') \leq E\ \pi\ (b,X \mid \mathcal{B}) \leq 0$. Since the risk premiums are non-positive for a risk lover, he would be willing to pay some positive (non-negative) amount to receive the random reward X. Since $\pi(b,X \mid \mathcal{B})$ is the maximum amount he would pay (given $\mathcal{B}$), the more information in $\mathcal{B}$, the less risk, and hence the less he would be willing to pay (in terms of expected value of $\pi(b,X \mid \mathcal{B})$) for the gamble X.

The relationship between conditional risk premiums and absolute risk aversion is straightforward.

<u>Lemma 3:</u>[16/] Let $b, X, \mathcal{B}, \mathcal{U}_1$, and $\mathcal{U}_2$ be given and suppose $E \mid \mathcal{U}_i(X) \mid < \infty$, $i = 1,2$. If $\mathcal{U}_1 \geq_r \mathcal{U}_2$, then $\pi_1(b, X \mid \mathcal{B}) \geq \pi_2(b,X \mid \mathcal{B})$ a.s.

Proof: The idea of the proof is to find risk premiums given by probability measures on $(\mathcal{R}^1, \mathcal{B}_1)$ that equal the conditional risk premiums. To this end, let $\hat{P}(d\ x \mid \mathcal{B})$ be the regular conditional distribution of X given $\mathcal{B}$. Then for each $\omega \in \Omega$ fixed, $\hat{P}\ (d\ x \mid \mathcal{B})\ (\omega)$ is a probability on $(\mathcal{R}^1, \mathcal{B}_1)$ and since X and $\mathcal{U}_i(X)$ are integrable, for almost all ω, $\hat{P}\ (d\ x \mid \mathcal{B})(\omega) \in \mathcal{P}_{\mathcal{U}_1} \cap \mathcal{P}_{\mathcal{U}_2}$. Define the risk premiums $\pi_i(b,\hat{P})$ in the usual way, i.e.,

$$\mathcal{U}_i(b + \int x\ \hat{P}(d\ x \mid \mathcal{B})\ (\omega) - \pi_i(b, \hat{P})(\omega))$$

$$= \int \mathcal{U}_i(b + x)\ \hat{P}(d\ x \mid \mathcal{B})(\omega),\ i = 1,2.$$

By assumption we have $\pi_1(b,\hat{P}) \geq \pi_2(b,\hat{P})$ a.s. But since

$$E(\mathcal{U}_i(b + x) \mid \mathcal{D}) = \int \mathcal{U}_i(b + x)\, \hat{P}(dx \mid \mathcal{D}) \text{ a.s., } i = 1,2$$

$$E(X \mid \mathcal{D}) = \int x\, \hat{P}(dx \mid \mathcal{D}) \text{ a.s.,}$$

we have that $\pi_i(b,\hat{P}) = \pi_i(b,X \mid \mathcal{D})$ a.s. $i = 1,2$. Hence a.s.

$$\pi_1(b,X \mid \mathcal{D}) = \pi_1(b,\hat{P}) \geq \pi_2(b,\hat{P}) = \pi_2(b,X \mid \mathcal{D}).$$

The converse of Lemma 3 is not true in general. The problem is that $(\Omega, \mathcal{J})$ may not be rich enough so that given a probability p on $(\mathcal{R}^1, \mathcal{B}_1)$ there is an X on $(\Omega, \mathcal{J})$ such that $P_X = p$, e.g., if Ω has only a finite number of points. If $(\Omega, \mathcal{J}) = (\mathcal{R}^1, \mathcal{B}_1)$, the converse is clearly true (if stated "for all $\mathcal{D} \subset \mathcal{B}_1$").

We conclude this section with the following result.

Lemma 4 : Let X and Y be two random variables and $\mathcal{U}_1$, $\mathcal{U}_2$ two utility functions with $E\,|\,\mathcal{U}_i(X)\,| < \infty$, $i = 1,2$, and $\mathcal{U}_1 \geq_r \mathcal{U}_2$. If $\mathcal{U}_2(y) \geq E(\mathcal{U}_2(X) \mid Y=y)$ a.s. P_Y, then $\mathcal{U}_1(y) \geq E(\mathcal{U}_1(X) \mid Y = y)$ a.s. P_Y. Similarly, if $\mathcal{U}_1(y) \leq E(\mathcal{U}_1(X) \mid Y = y)$ a.s. P_Y, then $\mathcal{U}_2(y) \leq E(\mathcal{U}_2(X) \mid Y = y)$ a.s. P_Y.

Proof: The proof follows directly from Lemma 3. We illustrate with the proof of the first statement.

$$\mathcal{U}_2(y) \geq E(\mathcal{U}_2(X) \mid Y = y) = \mathcal{U}_2(E(X \mid Y = y) - \pi_2(X \mid Y = y))$$
$$\Rightarrow \quad y \geq E(X \mid Y = y) - \pi_2(X \mid Y = y) \geq E(X \mid Y = y) - \pi_1(X \mid Y = y)$$
$$\Rightarrow \quad \mathcal{U}_1(y) \geq \mathcal{U}_1(E(X \mid Y = y) - \pi_1(X \mid Y = y)) = E(\mathcal{U}_1(X) \mid Y = y)$$

where $\pi_i(X \mid Y = y) = \pi_i(0,X \mid Y = y)$ $i = 1,2$ and all relations hold a.s. P_Y.

A temporal interpretation of this result is illuminating. We may consider

Y to be a current state variable say wealth and X to be the state variable some time in the future. Then $E(\mathcal{U}_i(X) \mid Y = y)$ is an estimate of future utility given the current state is y. Lemma 4 says that if the less risk averse individual views the future pessimistically, then so does the more risk averse individual. If the more risk averse individual views the future optimistically, then so does the less risk averse individual. Our results on optimal timing policies in the next two sections are applications of this principle. 17/

3. THE OPTIMAL STOPPING PROBLEM

In this section we introduce the optimal stopping problem and prove the theorem mentioned in the introduction. We begin by defining the problem in abstract form to facilitate references since it is in this form that the problem appears in most of the theoretical literature. It will be helpful to keep in mind, however, the interpretation of this problem as a model for optimal timing decisions. Such an interpretation is made throughout this section and section 4. For ease of exposition and to conserve space we consider only infinite horizon problems here. 18/

3.1 Introduction and Definitions.

A decision maker has the opportunity of observing in order a sequence of random variables $\{X_n;\ n \geq 1\}$ defined on the probability space $(\Omega, \mathcal{F}, P)$. A realization $X_n(\omega)$ for $\omega \in \Omega$ is interpreted as a monetary reward at stage n. At each stage, the decision maker may stop and accept X_n as his reward or continue and observe X_{n+1}. The decision maker is assumed to have a utility function $\mathcal{U}$ as described in section 2.2 and initial wealth $b \in \mathcal{R}^1$. He seeks to maximize the expected utility of terminal wealth. To do so he must have a strategy which tells him to stop observing and accept a reward.

Like most other optimization problems, the technical elements and interpretation of the problem structure impose restrictions on the strategies that can be considered. On the technical side, at the very least we want strategies to define terminal reward so that expected utility of terminal wealth has meaning. We should therefore require that a stopping strategy define terminal reward as a random variable.

On the interpretation side, there are two natural restrictions. Real world decision processes involve observation or gathering of information and implementation of actions based on that information. At the time an action is taken, observations are historical, and actions are eventually taken. It seems reasonable then to require that the decision to stop at stage k be made only on the basis of information obtained about the process up to and including stage k, i.e., on the observations $X_1,\ldots,X_k$, and that a stopping strategy should lead to action eventually, i.e., a stopping strategy should indeed stop (say a.s.).

With these restrictions in mind, the following definition of a stopping rule seems appropriate. Denote by $\mathcal{F}_n$ the σ-algebra generated by $X_1,\ldots,X_n$. A stopping rule for the process $\{X_n, \mathcal{F}_n, n \geq 1\}$ is a random variable t defined on Ω taking values in $\{1,2,\ldots,\infty\}$ and satisfying (a) $P(t < \infty) = 1$, and (b) $\{\omega : t(\omega) = k\} \in \mathcal{F}_k$, $k = 1,2,\ldots$ Condition (a) states that stopping must be an almost sure event and (b) that stopping at k be an event depending only on $X_1,\ldots,X_k$, i.e., the past and the present.

Conditions (a) and (b) on a stopping rule formalize the restrictions on information and actions discussed above. The requirement that a stopping rule t be a random variable permits us to define the stopped variable X_t in the following very sensible way.

$$X_t(\omega) = X_k(\omega) \quad \text{if } t(\omega) = k$$
$$= \infty \quad \text{if } t(\omega) = \infty \tag{3.1}$$

The convention $X_t = \infty$ on $(t = \infty)$ just completes the definition of X_t but adds nothing to the problem since $P(t = \infty) = 0$. X_t has the obvious interpretation as the (random) reward received by the decision maker who employs a stopping rule strategy. That X_t is a random variable may be seen from the identity

$$\{\omega : X_t(\omega) \leq \alpha\} = \bigcup_{n=1}^{\infty} (\{\omega : t(\omega) = n\} \cap \{\omega : X_n(\omega) \leq \alpha\}) \in \mathcal{J}$$

for all $\alpha \in \mathcal{R}^1$. 19/

We can now state the optimal stopping problem (OSP) precisely. Let T denote the class of stopping rules t for which

$$E\,\mathcal{U}(b + X_t)^- < \infty \tag{3.2}$$

<u>OSP($\mathcal{U}$)</u>. Given $\{X_n, \mathcal{J}_n; n \geq 1\}$, $\mathcal{U}$, and b, find a stopping rule $t^* \in T$, if it exists, such that

$$E\,\mathcal{U}(b + X_{t^*}) \geq E\,\mathcal{U}(b + X_t) \;, \qquad \forall\; t \in T. \tag{3.3}$$

The rule t^*, if it exists, is called optimal for OSP($\mathcal{U}$) or a solution of OSP($\mathcal{U}$).

Before we discuss the characterization of optimal rules let us interpret the above specification of the OSP for optimal timing of a particular action, say the sale of an asset. For an individual who has decided to sell the asset, the decision that remains is when to sell. The X_n may be interpreted as random price offers. The individual may, after observing X_n, accept it or wait one more period for another offer X_{n+1} which, at that decision point, is random. We are assuming that what the seller can know about X_{n+1} is only what can be inferred from observing $X_1, \ldots, X_n$ and knowledge of P. Following a stopping rule strategy t yields a (random) selling price X_t. Terminal wealth then is $b + X_t$.[20]

3.2 Optimal Stopping Rules.

As mentioned in the notes to the introduction, the OSP has been studied thoroughly from a mathematical point of view. We present below in Theorems 8 and 9 the major results of that study applied to our utility theoretic formulation of OSP. Proofs of these results require considerable development and we omit them to conserve space. The interested reader may consult the sources given. All we have added in our statements of these results is the utility function $\mathcal{U}$. The preliminary work on further defining OSP($\mathcal{U}$) mathematically is given here.

In addition to the assumption that the X_n are integrable we make the following assumptions on the process $\{X_n, \mathcal{F}_n ; n \geq 1\}$ and any utility function $\mathcal{U}$ we may use in this section and the next. Some assumptions like these are necessary to make the problem tractable mathematically. Assumptions (A1) and (A3) are slightly stronger than necessary. (A1) is needed in section 3.3 and (A3) leads most directly to the optimality result we want to use.

(A1) $E(\sup_n | X_n |) < \infty$

(A2) $E | \mathcal{U}(b + X_n) | < \infty, \ n = 1,2,\ldots$

(A3) $E(\sup_n \mathcal{U}(b + X_n)^+) < \infty$

The value V of OSP($\mathcal{U}$) is defined as

$$V \equiv \sup_{t \in T} E\, \mathcal{U}(b + X_t) \tag{3.4}$$

Some immediate consequences of (A1) - (A3) are contained in the following lemma. For notational convenience let $Y_t = \mathcal{U}(b + X_t)$ for all stopping rules t.

Lemma 5. Under (A1) - (A3), T is nonempty, $E|X_t| < \infty$ and $E|Y_t| < \infty$ for all $t \in T$, and $|V| < \infty$.

Proof: Let $t \equiv n$. Then (A2) implies $t \in T$. Now for all $t \in T$, $EY_t^- < \infty$. From (A3),

$$EY_t^+ = E \sum_{n=1}^{\infty} \chi_{(t=n)} Y_n^+ \leq E \sum_{n=1}^{\infty} \chi_{(t=n)} (\sup_k Y_k^+)$$

$$= E(\sup_n Y_n^+) = E(\sup_n \mathcal{U}(b + X_n)^+) < \infty,$$

(where χ_A denotes the characteristic function of the set A) and hence $E|Y_t| = EY_t^+ + EY_t^- < \infty$ for all $t \in T$. The same argument with Y_t^+ replaced by $|X_t|$ shows, by (A1), that $E|X_t| < \infty$ for all $t \in T$.

Again from (A3)

$$V = \sup_{t \in T} EY_t = \sup_{t \in T} E \sum_{n=1}^{\infty} \chi_{(t=n)} Y_n$$

$$\leq \sup_{t \in T} E \sum_{n=1}^{\infty} \chi_{(t=n)} Y_n^+$$

$$\leq \sup_{t \in T} E \sum_{n=1}^{\infty} \chi_{(t=n)} (\sup_k Y_k^+)$$

$$= E(\sup_n Y_n^+) < \infty$$

Since $V \geq EY_n > -\infty$, the proof is complete.

The characterization of an optimal stopping rule involves the concept of the essential supremum of a collection of random variables. Let $\{Z_\alpha, \alpha \in A\}$ be such a collection (indexed by the set A). The $\operatorname{ess\,sup}_{\alpha \in A} Z_\alpha$ is a random variable Z such that (a) $Z_\alpha \le Z$ a.s. for all $\alpha \in A$, and (b) if W is a random variable satisfying (a), then $Z \le W$ a.s. Some properties of the ess sup are listed here in the form of a lemma for easy reference.

Lemma 6. Let $\{Z_\alpha : \alpha \in A\}$ be a collection of random variables. Then

(i) $Z = \operatorname{ess\,sup}_{\alpha \in A} Z_\alpha$ exists and is unique a.s.;

(ii) there exists a sequence $\{\alpha_k\}_{k=1}^{\infty} \subset A$ such that $Z = \sup_k Z_{\alpha_k}$ a.s.;

(iii) Z is $\mathcal{F}(Z_\alpha; \alpha \in A)$ - measurable;

(iv) if $A' \subset A$, then $\operatorname{ess\,sup}_{\alpha \in A'} Z_\alpha \le Z$ a.s.;

(v) if $F \in \mathcal{F}$ and W is a random variable such that $W \ge Z_\alpha$ a.s. on F for each $\alpha \in A$, then $W \ge Z$ a.s. on F.

Proof: [25; p. 44] and [11; Lemma 2.1]

Now let T_n denote the class of stopping rules $t \in T$ such that $t \ge n$ a.s. We associate with OSP($\mathcal{U}$) the process $\{S_n, \mathcal{F}_n; n \ge 1\}$ where the random variables S_n are defined by

$$S_n = \operatorname*{ess\,sup}_{t\in T_n} E(\mathcal{U}(b + X_t) \mid \mathcal{F}_n)$$
$$= \operatorname*{ess\,sup}_{t\in T_n} E(Y_t \mid \mathcal{F}_n) \tag{3.6}$$

The reason the ess sup is used in (3.6) instead of the supremum is that T_n may be uncountable. The expression $\sup_{t\in T_n} E(Y_t \mid \mathcal{F}_n)$ is then not well defined since it may not be measurable and any two versions (arising from different versions of the $E(Y_t \mid \mathcal{F}_n)$) may differ on a set of positive probability.

<u>Lemma 7.</u> S_n is $\mathcal{F}_n$- measurable and (A2) and (A3) imply $E|S_n| < \infty, n=1,2,\ldots$

Proof: By Lemma 6 (iii) S_n is measurable in the σ-algebra generated by the $E(Y_t \mid \mathcal{F}_n)$, $t \in T_n$. But each $E(Y_t \mid \mathcal{F}_n)$ is $\mathcal{F}_n$-measurable and hence so is S_n. Let $\{t_k\} \subset T_n$ be the sequence asserted in Lemma 6 (ii). Then by definition of S_n and (A3),

$$Y_n = E(Y_n \mid \mathcal{F}_n) \le S_n = \sup_k E(Y_{t_k} \mid \mathcal{F}_n)$$
$$\le \sup_k E(Y_{t_k}^+ \mid \mathcal{F}_n)$$
$$\le E(\sup_k Y_{t_k}^+ \mid \mathcal{F}_n)$$
$$\le E(\sup_m Y_m^+ \mid \mathcal{F}_n) \text{ a.s.}$$

Hence $EY_n \le ES_n \le E(E(\sup_m Y_m^+ \mid \mathcal{F}_n)) = E(\sup_m Y_m^+)$ and $E \mid S_n \mid < \infty$.

We may interpret S_n as the optimal gain to the decision maker who has not stopped before stage n. This interpretation is reinforced by the following fundamental recursion relation of optimal stopping.

Theorem 8. For each positive integer n,

(i) $S_n = \max [Y_n, \operatorname{ess\,sup}_{t \in T_{n+1}} E(Y_t \mid \mathcal{F}_n)]$ a.s.

(ii) $\operatorname{ess\,sup}_{t \in T_{n+1}} E(Y_t \mid \mathcal{F}_n) = E(S_{n+1} \mid \mathcal{F}_n)$ a.s.

(iii) $S_n = \max [Y_n, E(S_{n+1} \mid \mathcal{F}_n)]$ a.s.

Proof: (iii) follows from (i) and (ii). See [11, Lemma 3.1] and [7, Theorem 4.1] for proofs of (i) and (ii).

An examination of (iii) of Theorem 8 together with the above interpretation of S_n suggests the following rule as a candidate for an optimal rule. Define the rule τ as follows:

$$\tau(\omega) = \begin{cases} k \text{ if } S_j(\omega) > Y_j(\omega),\ j = 1,\dots,k-1 \\ \quad \text{and } S_k(\omega) = Y_k(\omega) \\ \infty \text{ if } S_k(\omega) > Y_k(\omega),\ k=1,2,\dots \end{cases} \tag{3.7}$$

τ says stop at k if the utility of the current reward is the first such reward that is at least as great as the conditional expectation of the optimal gain from continuing, i.e., $E(S_{k+1} \mid \mathcal{F}_k)$.

From (3.7) we get

$$(\tau = k) = \bigcap_{j=1}^{k-1} (S_j > Y_j) \cap (S_k = Y_k)$$

and hence $(\tau = k) \in \mathcal{F}_k$ ($\mathcal{F}_n \subset \mathcal{F}_{n+1}$, n=1,2,...). There is no claim, however, that $P(\tau < \infty) = 1$ and we have given no condition sufficient for this to be true. Also we have not shown that $EY_\tau^- < \infty$. That these results hold if an optimal stopping rule for OSP($\mathcal{U}$) exists is a consequence of the following optimality theorem.

Theorem 9. (i) If $P(\tau < \infty) = 1$, then τ is optimal for OSP($\mathcal{U}$) and $S_n = E(Y_\tau \mid \mathcal{F}_n)$ a.s. $(\tau \geq n)$, n=1,2,... (ii) if a solution t* exists for OSP($\mathcal{U}$), then τ is also a solution and $\tau \leq t^*$ a.s.

Proof: See [11, Theorem 3.1] and [7; Theorems 4.2 and 4.5].

Thus if solutions exist for OSP($\mathcal{U}$), τ is also a solution and is the minimal solution. Although there are many more results in the theory of optimal stopping that are interesting, including sufficient conditions for $P(\tau < \infty) = 1$ and methods of computing V and S_n by truncation, backward induction, and taking limits, Theorems 8 and 9 are sufficient for our purposes in this section. In section 4 we will make use of the following result.

Lemma 10: If t is any stopping rule such that

(a) EX_t exists

(b) $E(X_{n+1} \mid \mathcal{F}_n) \geq X_n$ a.s. $(t > n)$, n=1,2,...,and,

(c)

$$\underline{\lim}_n \int_{(t > n)} X_n^+ dP = 0,$$

then $E(X_t \mid \mathcal{F}_n) \geq X_n$ a.s. $(t > n)$.

Proof: [7; Lemma 3.3].

Note that under (A1), (a) and (c) of this lemma are satisfied. Condition (a) was proved in Lemma 5. (A1) gives

$$\lim_n \int_{(t > n)} |X_n| dP = 0$$

for every stopping rule t.

3.3 Optimal Stopping Rules and Absolute Risk Aversion.

We are now in a position to prove the basic result of this paper, Theorem 11. From the partial ordering $\geq_r$ on utility functions we get a corresponding ordering on minimal solutions to the associated OSP's. We consider optimal stopping problems for two utility functions $\mathcal{U}_1$ and $\mathcal{U}_2$ beginning with the same initial wealth $b \in R^1$. Denote by τ_1 and τ_2 and S_n^1 and S_n^2, $n=1,2,\ldots$, the random variables defined in (3.7) and (3.6) for OSP($\mathcal{U}_1$) and OSP($\mathcal{U}_2$) respectively.

<u>Theorem 11.</u> If $\mathcal{U}_1 \geq_r \mathcal{U}_2$, then $\tau_2 \geq \tau_1$ a.s.

Proof: It suffices to show that $\{\omega:\tau_2(\omega) = n\} \subset \{\omega:\tau_1(\omega) \leq n\}$ a.s. We will prove this first using conditional risk premiums and then using Lemma 1.

First Proof: By definition of τ_2, $\chi_{(\tau_2=n)}\, \mathcal{U}_2(b + X_n) = \chi_{(\tau_2=n)}\, S_n^2$. Hence for each $t \in T_n$,

$$\chi_{(\tau_2=n)}\, \mathcal{U}_2(b + X_n) \geq \chi_{(\tau_2=n)}\, E(\mathcal{U}_2(b + X_t) \mid \mathcal{F}_n) \text{ a.s.}$$

$$= \chi_{(\tau_2=n)}\mathcal{U}_2(b + E(X_t \mid \mathcal{F}_n) - \pi_2(b, X_t \mid \mathcal{F}_n)) \text{ a.s.} \quad (3.8)$$

By monotonicity of $\mathcal{U}_2$ then for each $t \in T_n$,

$$b + X_n \geq b + E(X_t \mid \mathcal{F}_n) - \pi_2(b,X_t \mid \mathcal{F}_n) \quad (3.9)$$

a.s. $(\tau_2 = n)$. From Lemma 3, by assumption $\pi_1(b,X_t \mid \mathcal{F}_n) \geq \pi_2(b,X_t \mid \mathcal{F}_n)$ a.s. for each $t \in T_n$. Hence from (3.9), for each $t \in T_n$

$$b + X_n \geq b + E(X_t \mid \mathcal{F}_n) - \pi_1(b,X_t \mid \mathcal{F}_n) \quad (3.10)$$

a.s. $(\tau_2 = n)$. By monotonicity of $\mathcal{U}_1$, then for each $t \in T_n$,

$$\mathcal{U}_1(b + X_n) \geq \mathcal{U}_1(b + E(X_t \mid \mathcal{F}_n) - \pi_1(b, X_t \mid \mathcal{F}_n))$$
$$= E(\mathcal{U}_1(b + X_t) \mid \mathcal{F}_n) \tag{3.11}$$

where the inequality holds a.s. $(\tau_2 = n)$ and the equality a.s. From (3.11) and Lemma 6 (v) we get $\mathcal{U}_1(b + X_n) \geq S_n^1$ a.s. $(\tau_2 = n)$, i.e., $\mathcal{U}_1(b + X_n) = S_n^1$ a.s. $(\tau_2 = n)$. By definition of τ_1 then $\tau_1 \leq n$ a.s. $(\tau_2 = n)$ and we are done.

Second Proof: From (3.8), for each $t \in T_n$,

$$b + X_n \geq \mathcal{U}_2^{-1}(E(\mathcal{U}_2(b + X_t) \mid \mathcal{F}_n)) \quad \text{a.s. } (\tau_2 = n) \tag{3.12}$$

and hence

$$\mathcal{U}_1(b + X_n) \geq \mathcal{U}_1(\mathcal{U}_2^{-1}(E(\mathcal{U}_2(b + X_t) \mid \mathcal{F}_n))) \quad \text{a.s. } (\tau_2 = n) \tag{3.13}$$

By Lemma 1 and Jensen's inequality, we get from (3.13) that for each $t \in T_n$,

$$\mathcal{U}_1(b + X_n) \geq E(\mathcal{U}_1(\mathcal{U}_2^{-1}(\mathcal{U}_2(b + X_t))) \mid \mathcal{F}_n) \quad \text{a.s. } (\tau_2 = n)$$
$$= E(\mathcal{U}_1(b + X_t) \mid \mathcal{F}_n) \tag{3.14}$$

which is (3.11).

The following two corollaries are immediate.

Corollary 12. If $\mathcal{U}_1 \geq_r \mathcal{U}_2$ and solutions to OSP($\mathcal{U}_1$) and OSP($\mathcal{U}_2$) exist, then solutions t_1 and t_2 exist such that t_i solves OSP ($\mathcal{U}_i$), $i = 1,2$, and $t_1 \leq t_2$ a.s. If t_2 solves OSP ($\mathcal{U}_2$), then there is a solution t_1 to OSP($\mathcal{U}_1$) such that $t_1 \leq t_2$ a.s.

Proof: Follows from Theorems 9 and 11.

Corollary 13. Assume $\mathcal{U}$ is DARA and consider τ defined in (3.7) for OSP($\mathcal{U}$) as a function of b. Then $\tau(b)$ is a.s. nondecreasing in b.

Proof: Let $\mathcal{U}_2(x) = \mathcal{U}(x+k)$ for $k > 0$ and $\mathcal{U}_1(x) = \mathcal{U}(x)$ and apply Theorem 11.

To complete the work of this section, we need a partial converse of the last statement in Corollary 12 and a sharpening of the inequality $\tau_1 \leq \tau_2$ a.s.

in Theorem 11. We have found no conditions where the existence of a solution to OSP($\mathcal{U}_1$) implies the existence of one for OSP($\mathcal{U}_2$) save those that guarantee existence for both problems. Given the results of Theorems 9 and 11 we should not expect to do so, but we defer any further discussion for future research.

We can sharpen the inequality of Theorem 11, by defining the relation $>_r$ in the obvious way. Given two utility functions $\mathcal{U}_1$ and $\mathcal{U}_2$ we say that $\mathcal{U}_1$ is strictly more risk averse than $\mathcal{U}_2$, and write $\mathcal{U}_1 >_r \mathcal{U}_2$, if $\mathcal{U}_1 \geq_r \mathcal{U}_2$ and $\pi_1(b,p) > \pi_2(b,p)$ for all non-degenerate $p \in \mathcal{P}_{\mathcal{U}_1} \cap \mathcal{P}_{\mathcal{U}_2}$ and all $b \in \mathcal{R}^1$. By the same methods used to prove Lemma 1 we have $\mathcal{U}_1 >_r \mathcal{U}_2$ if and only if $\mathcal{U}_1(\mathcal{U}_2^{-1}(t))$ is strictly concave in t. To conserve space we will prove the desired result using the strict concavity of $\mathcal{U}_1(\mathcal{U}_2^{-1}(\cdot))$.

<u>Theorem 14</u>. Suppose $\mathcal{U}_1 >_r \mathcal{U}_2$ and a solution to OSP($\mathcal{U}_1$) exists. If there exists an $n \geq 1$ with $P(\tau_1 = n) > 0$ and

$$P((\tau_1 \geq n) \cap (E(\mathcal{U}_1(b + X_{\tau_1}) \mid \mathcal{F}_n) \neq \mathcal{U}_1(b + X_{\tau_1}))) > 0 \qquad (*)$$

then $P(\tau_1 < \tau_2) > 0$.

Proof: Let $\hat{\tau}_1 = \max(n,\tau_1)$. Then $\hat{\tau}_1$ is a stopping rule since τ_1 is by Theorem 9 (ii) and we have $E(\mathcal{U}_1(b + X_{\hat{\tau}_1}) \mid \mathcal{F}_n) = E(\mathcal{U}_1(b + X_{\tau_1}) \mid \mathcal{F}_n)$ a.s. $(\tau_1 \geq n)$. By Theorem 9 (i), then a.s. $(\tau_1 \geq n)$

$$\begin{aligned}
\mathcal{U}_1(b + X_n) \leq S_n^1 &= E(\mathcal{U}_1(b + X_{\tau_1}) \mid \mathcal{F}_n) \\
&= E(\mathcal{U}_1(b + X_{\hat{\tau}_1}) \mid \mathcal{F}_n) \\
&\Rightarrow b + X_n \leq \mathcal{U}_1^{-1}(E(\mathcal{U}_1(b + X_{\hat{\tau}_1}) \mid \mathcal{F}_n)) \\
&\Rightarrow \mathcal{U}_2(b + X_\tau) \leq \mathcal{U}_2(\mathcal{U}_1^{-1}(E(\mathcal{U}_1(b + X_{\hat{\tau}_1}) \mid \mathcal{F}_n)))
\end{aligned}$$

$$< E\ (\mathcal{U}_2(\mathcal{U}_1^{-1}(\mathcal{U}_1(b + X_{\hat{\tau}_1}))) \mid \mathcal{F}_n)$$

$$= E(\mathcal{U}_2(b + X_{\hat{\tau}_1}) \mid \mathcal{F}_n) \leq S_n^2 \qquad (**)$$

where the strict inequality follows from strict convexity of $\mathcal{U}_2(\mathcal{U}_1^{-1}(\cdot))$ and (*), and the last inequality from the fact that $\hat{\tau}_1 \in T_n$. From Theorem 11, $(\tau_1 \geq n) \subset (\tau_2 \geq n)$ and from (**) we get $(\tau_1 \geq n) \subset (\tau_2 > n)$ which gives $(\tau_1 = n) \subset (\tau_2 > n)$. $P(\tau_1 = n) > 0$ then gives the conclusion.

Note that the hypotheses of Theorem 14 rule out the case where the X_n are independent and identically distributed. Indeed, we have the following result. Denote by OSP(X) the optimal stopping problem when the utility function is linear and by $\bar{\tau}$ and $\bar{S}_n$ the random variables associated with OSP(X) by (3.7) and (3.6).

Corollary 15. If the X_n are independent and identically distributed and $\bar{\tau}$ solves OSP(X), then $\bar{\tau}$ solves OSP($\mathcal{U}$) for any utility function $\mathcal{U}$ for which a solution to OSP($\mathcal{U}$) exists.

Proof: For convenience we assume b = 0. For any utility function $\mathcal{U}$ (including the linear utility function X), $\mathcal{U}(X_t)$ is independent of $X_1,\ldots,X_n$ for each $t \in T_{n+1}$. Hence $E(\mathcal{U}(X_t) \mid \mathcal{F}_n) = E\,\mathcal{U}(X_t)$ a.s. and by Theorem 8 (ii),

$$E(S_{n+1} \mid \mathcal{F}_n) = \operatorname*{ess\,sup}_{t\in T_{n+1}} E(\mathcal{U}(X_t) \mid \mathcal{F}_n)$$

$$= \sup_{t\in T_{n+1}} E\,\mathcal{U}(X_t) = \sup_{t\in T} E\,\mathcal{U}(X_t) = V,$$

since the X_n are identically distributed. Hence $\bar{\tau}(\omega)$ = first positive integer k such that $X_k(\omega) \geq \bar{V}$. Let τ be the rule given by (3.7) for OSP($\mathcal{U}$) and $W = \mathcal{U}^{-1}(V)$.

By assumption and Theorem 9 (ii), $\tau \in T$ and τ solves OSP($\mathcal{U}$). Hence $\overline{V} \geq EX_\tau \geq W$ and thus $X_{\overline{\tau}} \geq X_\tau$ a.s. Then $\mathcal{U}(X_{\overline{\tau}}) \geq \mathcal{U}(X_\tau)$ a.s. and $E\mathcal{U}(X_{\overline{\tau}}) \geq E\,\mathcal{U}(X_\tau)$. Since $\overline{\tau} \in T$, $E\,\mathcal{U}(X_{\overline{\tau}}) = E\,\mathcal{U}(X_\tau) = V$ and $\overline{\tau}$ solves OSP($\mathcal{U}$).

3.4 An Example.

We close this section with an example. Let $\Omega = \{\omega_1,\omega_2,\omega_3,\omega_4\}$, $\mathcal{F} = 2^\Omega$ (the power set of Ω) and $P = (p_1,p_2,p_3,p_4)$ where $P(\omega_i) = p_i > 0$, $i = 1,\ldots,4$ and $\sum_{i=1}^{4} p_i = 1$. Let x_1,x_2,x_3,x_4,x_5 be any real numbers with $0 < x_1 < x_2 < x_3 < x_4 < x_5$. We define a random variable on Ω by giving a 4-tuple of these numbers that indicate the values the random variable takes on $\omega_1,\ldots,\omega_4$. Let

$$X_1 \equiv (x_2,x_2,x_4,x_4)$$

$$X_2 \equiv (x_1,x_1,x_2,x_1)$$

$$X_3 \equiv (x_3,x_3,x_2,x_2)$$

$$X_4 \equiv (x_1,x_5,x_4,x_3)$$

$$X_n \equiv 0,\ n \geq 5$$

Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be two utility functions with $\mathcal{U}_1 >_r \mathcal{U}_2$. Tables 1 and 2 give the values of S_n^1 and S_n^2 respectively.

TABLE 1

	ω_1	ω_2	ω_3	ω_4
s_1^1	$\mathcal{U}_1(x_3)$	$\mathcal{U}_1(x_3)$	$\mathcal{U}_1(x_4)$	$\mathcal{U}_1(x_4)$
s_2^1	"	"	"	$\mathcal{U}_1(x_3)$
s_3^1	"	"	"	"
s_4^1	$\mathcal{U}_1(x_1)$	$\mathcal{U}_1(x_5)$	"	"
s_n^1, $n \geq 5$	0	0	0	0

TABLE 2

	ω_1	ω_2	ω_3	ω_4
s_1^2	A_2	A_2	$\mathcal{U}_2(x_4)$	$\mathcal{U}_2(x_4)$
s_2^2	"	"	"	$\mathcal{U}_2(x_3)$
s_3^2	"	"	"	"
s_4^2	$\mathcal{U}_2(x_1)$	$\mathcal{U}_2(x_5)$	"	"
s_n^2, $n \geq 5$	0	0	0	0

These tables were construcred by backward induction using Theorem 8 (iii) and the assumption that

$$\frac{p_1}{p_2} = \frac{\mathcal{U}_1(x_5) - \mathcal{U}_1(x_3)}{\mathcal{U}_1(x_3) - \mathcal{U}_1(x_1)} \equiv B_1$$

which implies that

$$A_1 \equiv \frac{\mathcal{U}_1(x_1)\ p_1 + \mathcal{U}(x_5)\ p_2}{p_1 + p_2} = E(S_4^1 \mid \mathcal{T}_3)(\omega_1)$$

$$= E(S_4^1 \mid \mathcal{T}_3)(\omega_2) = \mathcal{U}_1(x_3).$$

This can clearly be done since $B_1 > 0$. From Pratt [26, Theorem 1(e)] we have that

$$B_2 \equiv \frac{\mathcal{U}_2(x_5) - \mathcal{U}_2(x_3)}{\mathcal{U}_2(x_3) - \mathcal{U}_2(x_1)} > B_1$$

and hence

$$A_2 \equiv \frac{\mathcal{U}_2(x_1)p_1 + \mathcal{U}_2(x_5)p_2}{p_1 + p_2} = E(S_4^2 \mid \mathcal{T}_3)\ (\omega_1)$$

$$= E(S_4^2 \mid \mathcal{T}_3)(\omega_2) > \mathcal{U}_2(x_3).$$

Computing τ_1 and τ_2 from Tables 1 and 2 we have

TABLE 3

	ω_1	ω_2	ω_3	ω_4
τ_1	3	3	1	1
τ_2	4	4	1	1

We will use this example again in Section 4.2.

4. SOME APPLICATIONS.

In terms of the interpretation of the OSP for optimal timing decisions, the results of section 3.3 say simply that the more risk averse individual will implement a given action at least as soon as will the less risk averse individual. Not much more can be said about the effect of risk aversion on optimal timing strategies. There are, however, two questions related to timing strategies that need attention. The first concerns the effect of risk preferences on a class of simplified, perhaps non-optimal timing strategies. The second concerns the effect of risk preferences on the values of the stopped reward. We consider these questions in this section. For convenience we assume b = 0 throughout.

4.1 One-period Myopic Rules.

The optimality properties of the rule τ of (3.7) are intuitive, but it is unlikely, except in very special cases, that τ could actually be computed for infinite horizon problems. A rule similar to τ is optimal for finite horizon problems and can be computed by backward induction, [21/] but for most problems only at great expense. It is natural then to ask if the results of section 3.3 hold for any class of simpler, perhaps non-optimal rules. The answer is yes for at least one class of such rules worth considering, which we call one-period myopic rules.

It may be that an individual (a firm, consumer, investor, etc.) does not have sufficient resources including time and computational capacity to compute the S_n involved in some timing decision. For example, after observing $X_1,\ldots,X_n$ it may be that the only estimate of future gain he can come up with is $E(\mathcal{U}(X_{n+1}) \mid \mathcal{F}_n)$ or $E(\mathcal{U}(X_{n+1}) \mid X_n)$, i.e., the expected utility of the next reward given knowledge of $X_1,\ldots,X_n$ or X_n respectively. We define the class

M of one-period myopic rules to be the collection of rules t of the form

$$t(\omega) = \begin{cases} k \text{ if } \mathcal{U}(X_j) \leq E(\mathcal{U}(X_{j+1}) \mid \mathcal{B}_j),\ j = 1,\ldots,k-1 \\ \quad \text{and } \mathcal{U}(X_k) \geq E(\mathcal{U}(X_{k+1}) \mid \mathcal{B}_k) \\ \\ \infty \text{ if } \mathcal{U}(X_k) < E(\mathcal{U}(X_{k+1}) \mid \mathcal{B}_k),\ k = 1,2,\ldots \end{cases} \tag{4.1}$$

for some sequence $\{\mathcal{B}_n;\ n \geq 1\}$ where $\mathcal{B}_n$ is a sub-σ-algebra of $\mathcal{F}_n$ for each n.

Clearly if $t \in M$ and $P(t < \infty) = 1$, then t is a stopping rule. If we denote by $\bar{M}$ the subset of M of rules that are stopping rules, then under (A2) $\bar{M} \subset T$. A rule in M bases the decision to stop or continue one more period on the utility of the current period reward and the expected utility of the following period reward given some subset of the information available in the current period. There are conditions under which an element of $\bar{M}$ solves OSP($\mathcal{U}$) and we give one such result in Proposition 17. First, however, we prove an easy corollary to Theorem 11.

Let $\{\mathcal{B}_n;\ n \geq 1\}$ be a sequence of sub-σ-algebras of $\mathcal{F}, \mathcal{B}_n \subset \mathcal{F}_n, n=1,2,\ldots,$ and define the rule λ by

$$\lambda(\omega) = \begin{cases} k \text{ if } \mathcal{U}(X_j) < E(\mathcal{U}(X_{j+1}) \mid \mathcal{B}_j),\ j = 1,\ldots,k-1 \\ \quad \text{and } \mathcal{U}(X_k) \geq E(\mathcal{U}(X_{k+1}) \mid \mathcal{B}_k) \\ \infty \text{ if } \mathcal{U}(X_k) < E(\mathcal{U}(X_{k+1}) \mid \mathcal{B}_k),\ k = 1,2,\ldots \end{cases} \tag{4.2}$$

λ simply says stop at k if k is the first positive integer such that utility of current reward X_k is at least as great as expected utility of next period's reward X_{k+1} given knowledge of events in $\mathcal{B}_k$. Consider λ_1 and λ_2 given by (4.2) for utility functions $\mathcal{U}_1$ and $\mathcal{U}_2$.

<u>Theorem 16:</u> If $\mathcal{U}_1 \geq_r \mathcal{U}_2$, then $\lambda_1 \leq \lambda_2$ a.s.

Proof: We prove this result using Lemma 1, but the first proof of Theorem 11 is also applicable (see also Lemma 4). The method of the second proof is used to conserve space. On $(\lambda_2 = n)$,

$$\mathcal{U}_2(X_n) \geq E(\mathcal{U}_2(X_{n+1}) \mid \mathcal{B}_n)$$

$$\Rightarrow \quad X_n \geq \mathcal{U}_2^{-1}(E(\mathcal{U}_2(X_{n+1}) \mid \mathcal{B}_n))$$

$$\Rightarrow \quad \mathcal{U}_1(X_n) \geq \mathcal{U}_1(\mathcal{U}_2^{-1}(E(\mathcal{U}_2(X_{n+1}) \mid \mathcal{B}_n)))$$

$$\geq E(\mathcal{U}_1(\mathcal{U}_2^{-1}(\mathcal{U}_2(X_{n+1}))) \mid \mathcal{B}_n)$$

$$= E(\mathcal{U}_1(X_{n+1}) \mid \mathcal{B}_n)$$

$$\Rightarrow \quad (\lambda_1 \leq n) \subset (\lambda_2 = n) \text{ a.s.}$$

Now consider the case where λ is given by (4.2) for $\mathcal{B}_n = \mathcal{F}_n$, $n = 1,2,\ldots$, i.e., λ uses all the information available up to time n, but projects only to period n + 1. Let $A_n = \{\omega : \mathcal{U}(X_n) \geq E(\mathcal{U}(X_{n+1}) \mid \mathcal{F}_n)\}$. In addition to (A2) and (A3) we assume the following:

(A4) $X_n \geq 0$ a.s., $n = 1,2,\ldots$, and $\lim_n X_n = 0$ a.s.;

(A5) $A_n \subset A_{n+1}$, $n = 1,2,\ldots$, and $\Omega = \bigcup_{n=1}^{\infty} A_n$ a.s.

Proposition 17. Under (A2) - (A5), λ solves OSP($\mathcal{U}$).

Proof: This is a special case of [7; Theorem 3.3].

Condition (A5) is the key to the optimality of λ and we could not expect one-period myopia to yield an optimal strategy without such an assumption. (A4) is a little stronger than necessary, but we use it because of its interpretation for the optimal timing of the sale of an asset. In such a problem the random price offers are non-negative. The second half of (A4) is a uniform integrability condition that makes proving optimality a little easier, but may be interpreted as indicating that the asset eventually becomes valueless.

The market prices for assets like capital equipment subject to physical deterioration or technical obsolescence, commodities like perishable farm produce, and certain financial assets and options with a fixed horizon maturity, would satisfy assumption (A4). The market prices of common stocks that follow a postulated random walk model, will not in general satisfy (A4). These models, however, have more significance for the second question of this section and we postpone a discussion of them until we have taken up this question.

4.2 Risk Preferences and Stopped Variables

One is tempted to conjecture that because the less risk averse individual is more willing to bear risk, his reward should in some sense be greater than that of the more risk averse individual. Such a conjecture is relatively easy to formulate, but quite difficult to prove in general and we give only partial answers here. First we must settle on what we mean by "in some sense greater".

An immediate candidate is $X_{\tau_2} \geq X_{\tau_1}$ a.s. where τ_1 and τ_2 are given by (3.7) for OSP($\mathcal{U}_1$) and OSP($\mathcal{U}_2$) respectively. But this condition makes no sense since $X_{\tau_2} \geq X_{\tau_1}$ a.s. $\Rightarrow \mathcal{U}_1(X_{\tau_2}) \geq \mathcal{U}_1(X_{\tau_1}) \Rightarrow E\,\mathcal{U}_1(X_{\tau_2}) \geq E\,\mathcal{U}_1(X_{\tau_1})$ and τ_2 is as good a strateyy for $\mathcal{U}_1$ as τ_1. Indeed, if $X_{\tau_2} \geq X_{\tau_1}$ a.s. and $P(X_{\tau_1} \neq X_{\tau_2}) > 0$

then τ_1 is not optimal for $\mathcal{U}_1$ (providing $\tau_2 \in T$). The next most likely candidate, $EX_{\tau_2} \geq EX_{\tau_1}$, does have some reasonable implications, but is not true in general. We give below conditions on $\mathcal{U}_1, \mathcal{U}_2$ and $\{X_n, \mathcal{F}_n; n \geq 1\}$ for $EX_{\tau_2} \geq EX_{\tau_1}$, and $EX_{\tau_1} \geq EX_{\tau_2}$, and an example when $EX_{\tau_1} > EX_{\tau_2}$.

Lemma 18. If $\mathcal{U}_1 \geq_r \mathcal{U}_2$, $P(\tau_2 < \infty) = 1$, and if

$$E(X_{\tau_2} \mid \mathcal{F}_n) \geq X_n \text{ a.s. } (\tau_1 = n),\ n = 1,2,\ldots, \tag{4.3}$$

then $EX_{\tau_2} \geq EX_{\tau_1}$.

Proof: From Lemma 5 , $E \mid X_{\tau_2} \mid < \infty$, and from Theorem 11, $\tau_1 \leq \tau_2$ a.s. Hence

$$EX_{\tau_2} = \int_{(\tau_2 < \tau_1)} X_{\tau_2} dP + \int_{(\tau_1 \leq \tau_2)} X_{\tau_2} dP$$

$$= \int_{(\tau_1 \leq \tau_2)} X_{\tau_2} dP = \sum_{n=1}^{\infty} \int_{(\tau_1 = n) \cap (\tau_2 \geq n)} X_{\tau_2} dP$$

$$= \sum_{n=1}^{\infty} \int_{(\tau_1 = n)} X_{\tau_2} dP = \sum_{n=1}^{\infty} \int_{(\tau_1 = n)} E(X_{\tau_2} \mid \mathcal{F}_n) dP$$

$$\geq \sum_{n=1}^{\infty} \int_{(\tau_1 = n)} X_n dP = \int_{\Omega} X_{\tau_1} dP = EX_{\tau_1},$$

and the proof is complete. [22/]

The condition (4.3) of the lemma is a special case of a condition called admissability of a stopping rule. [23/] A stopping rule t is said to be admissable for $\{X_n, \mathcal{F}_n;\ n \geq 1\}$ if EX_t exists and

$$E(X_t \mid \mathcal{F}_n) \geq X_n \text{ a.s. } (t > n),\ n = 1,2,\ldots, \tag{4.4}$$

Since (4.4) holds with equality on $(t = n)$, and since $\tau_1 \leq \tau_2$ a.s., if τ_2 is admissable for $\{X_n, \mathcal{F}_n;\ n \geq 1\}$, then (4.3) holds. Whether or not τ_2 is admissable for $\{X_n, \mathcal{F}_n;\ n \geq 1\}$ (or satisfies (4.3)) does not depend on the relation $\mathcal{U}_1 \geq_r \mathcal{U}_2$. We give now two conditions which guarantee the admissability of τ_2 The first depends on the risk aversion of $\mathcal{U}_2$ itself and the second on the "fairness" of the process $\{X_n, \mathcal{F}_n\}$. We then show that if the process is "unfair", we get $EX_{\tau_1} \geq EX_{\tau_2}$.

<u>Proposition 19.</u> If $\mathcal{U}_2$ is concave and $P(\tau_2 < \infty) = 1$, then τ_2 is admissable for $\{X_n, \mathcal{F}_n;\ n \geq 1\}$.

Proof: By Theorem 9(i), τ_2 is admissable for $\{\mathcal{U}_2(X_n), \mathcal{F}_n;\ n \geq 1\}$, i.e., for $n \geq 1$

$$E(\mathcal{U}_2(X_{\tau_2}) \mid \mathcal{F}_n) \geq \mathcal{U}_2(X_n) \text{ a.s. } (\tau_2 > n) \tag{4.5}$$

Hence a.s. $(\tau_2 > n)$

$$\begin{aligned} X_n &\leq \mathcal{U}_2^{-1}\ (E(\mathcal{U}_2(X_{\tau_2}) \mid \mathcal{F}_n)) \\ &\leq E(X_{\tau_2} \mid \mathcal{F}_n) \end{aligned}$$

since $\mathcal{U}_2^{-1}$ is convex.

Lemma 10 gives sufficient conditions for a stopping rule to be admissable and the next result follows from it directly.

Proposition 20. If $P(\tau_2 < \infty) = 1$ and $\{X_n, \mathcal{F}_n;\ n \geq 1\}$ is a submartingale, then τ_2 is admissable for $\{X_n, \mathcal{F}_n;\ n \geq 1\}$.

We have, however,

Proposition 21. If $\mathcal{U}_1 \geq_r \mathcal{U}_2$ and $\{X_n, \mathcal{F}_n;\ n \geq 1\}$ is a supermartingale, then X_{τ_1} and X_{τ_2} are integrable and $EX_{\tau_1} \geq EX_{\tau_2}$.

Proof: By Theorem 11, $\tau_1 \leq \tau_2$ a.s. Also

$$X_n \geq -X_n^- \geq \inf_k (-X_k^-) = -\sup_k X_k^- \geq -\sup_k |X_k|$$

and by (A1), X_n is bounded below by an integrable random variable. These are the hypotheses of [20; Theorem 28, p. 90]. The conclusion is that X_{τ_1}, X_{τ_2} are integrable and $X_{\tau_1} \geq E(X_{\tau_2} \mid \mathcal{F}_{\tau_1})$ a.s. where $\mathcal{F}_{\tau_1}$ is the σ-algebra of events A such that $A \cap (\tau_1 = n) \in \mathcal{F}_n$ for each $n \geq 1$. Taking expectations we get $EX_{\tau_1} \geq E(E(X_{\tau_2} \mid \mathcal{F}_{\tau_1})) = EX_{\tau_2}$.

Corollary 22. If $\mathcal{U}_1 \geq_r \mathcal{U}_2$, $P(\tau_2 < \infty) = 1$, and if $\{X_n, \mathcal{F}_n;\ n \geq 1\}$ is a martingale, then $EX_{\tau_2} = EX_{\tau_1}$.

Proof: Follows directly from Propositions 20 and 21 and Lemma 18.

To show that we may indeed have $EX_{\tau_1} > EX_{\tau_2}$, consider the example of section 3.4. If $\mathcal{U}_1 >_r \mathcal{U}_2$ we have, again from Pratt [26; Theorem 1 (e)], that

$$B_2 = \frac{\mathcal{U}_2(x_5) - \mathcal{U}_2(x_3)}{\mathcal{U}_2(x_3) - \mathcal{U}_2(x_1)} > \frac{\mathcal{U}_1(x_5) - \mathcal{U}_1(x_3)}{\mathcal{U}_1(x_3) - \mathcal{U}_1(x_1)} = B_1$$

It is possible then, say if $\mathcal{U}_2$ is strictly convex and $\mathcal{U}_1$ concave, to pick p_1 and p_2 so that

$$B_2 > \frac{p_1}{p_2} > \frac{x_5 - x_3}{x_3 - x_1} \geq B_1 \tag{4.6}$$

In this case Tables 1, 2, and 3 remain unchanged and we have

$$\begin{aligned} EX_{T_1} - EX_{T_2} &= (p_1 + p_2)x_3 + (p_3 + p_4)\, x_4 \\ &\quad - (p_1x_1 + p_2x_5) - (p_3 + p_4)\, x_4 \\ &= (p_1 + p_2)x_3 - (p_1x_1 + p_2x_5) > 0 \end{aligned}$$

by (4.6).

4.3 Stock Market Models

We propose in this section to draw some implications of our work for stock market transactions under the assumption of an efficient market in common stocks and in particular under the assumptions that successive price changes or successive percentage price changes are independent, i.e., the random walk models of stock price behavior. [24/] Some applications of stopping rules have been made to problems of when to exercise options to buy and sell common stocks. [25/] We will concentrate on the problem of when to sell a given fixed number of shares of a particular common stock. We wish to examine the expected selling price under the relation $\geq_r$.

The concrete form of the efficient market model postulates that

$$E(X_{n+1} \mid \mathcal{B}_n) = [1 + E(Z_{n+1} \mid \mathcal{B}_n)]\, X_n \quad \underline{26/} \tag{4.7}$$

where X_n is the price of the stock at time n, $Z_{n+1} = (X_{n+1} - X_n)/X_n$ and $\mathcal{B}_n$ is a symbol which represents whatever information is available at time n. To make (4.7) meaningful we assume the X_n are, as before, random variables defined on $(\Omega, \mathcal{F}, P)$ and that the $\mathcal{B}_n$ are an increasing sequence of sub-σ-algebra of $\mathcal{F}$ such that $\mathcal{F}_n \subset \mathcal{B}_n$, n = 1,2,... (By increasing we mean $\mathcal{B}_n \subset \mathcal{B}_{n+1}$, n = 1,2,...). All our results up to this point hold for a process $\{X_n, \mathcal{B}_n;\ n \geq 1\}$ where $\mathcal{F}_n \subset \mathcal{B}_n$ and the $\mathcal{B}_n$ are increasing. In defining stopping rules for this process we would then require $(t=n) \in \mathcal{B}_n$. Since we are dealing with a fixed number of shares, we may assume X_n represents the total selling price for this fixed number.

What we would like to conclude is that a "favorable" market in common stocks rewards those more willing to bear risk relatively, and that an "unfair" market rewards those less willing to bear risk relatively. Therefore we confine ourselves to the cases when

(C1) $E(Z_{n+1} \mid \mathcal{B}_n) \geq 0$ a.s.

(C2) $E(Z_{n+1} \mid \mathcal{B}_n) = 0$ a.s.

(C3) $E(Z_{n+1} \mid \mathcal{B}_n) \leq 0$ a.s.

These cases specify $\{X_n, \mathcal{B}_n;\ n \geq 1\}$ as (C1) a submartingale, (C2) a martingale, and (C3) a supermartingale.

Even though we limit ourselves to (C1) - (C3), we cannot apply Proposition 20 and Lemma 18, Proposition 21, and Corollary 22 to these cases without some qualifications. There are two technical problems involving the assumption (A1). This assumption is quite strong, implying among other things that the process $\{X_n, \mathcal{B}_n; n \geq 1\}$ is uniformly integrable, a property not possessed by many processes satisfying (4.7) and either (C1), (C2), or (C3).

The first problem is that Lemma 18 and Proposition 21 are based on the conclusion of Theorem 11 that $\tau_1 \leq \tau_2$ a.s. The proofs we have given for Theorem 11 are in turn based not only on (A2) and (A3) which seem reasonable in order to define OSP($\mathcal{U}$) in a tractable manner, but in a very crucial way on (A1). Recalling the definitions of the risk premium (2.5) and conditional risk premiums (2.6) - (2.9), it was specifically assumed that the random variable involved had a finite expectation. In using conditional risk premiums (or Jensen's inequality) in the proof of Theorem 11, it was necessary then to have conditions which insured that each stopped variable had a finite expectation. The proof of Lemma 5 uses (A1) to insure this.

The second problem is that the proofs of Propositions 20 and 21 depend, as we have stated them, directly on (A1). Proposition 20 depends on (a) and (c) of Lemma 10 being satisfied, which they are under (A1) as we have noted. Proposition 21 requires that the X_n be bounded below by an integrable random variable which requires an assumption like (A1).

The first problem concerning the conclusion of Theorem 11 can be repaired completely using certainty equivalents. The stumbling block in using risk premiums is that we need EX to be finite in order to define $\pi(X) = EX - \mathcal{U}^{-1}(E\,\mathcal{U}(X))$ since if $E\,|\,\mathcal{U}(X)\,| < \infty$ and $EX = \infty$, then $\pi(X) = \infty$ and $EX - \pi(X)$ is undefined. The quantity $EX - \pi(X)$ is a certainty equivalent of the gamble X. Define $C(b,X) = \mathcal{U}^{-1}(E\,\mathcal{U}(b + X)) - b$ when $E\,\mathcal{U}(X)$ exists. Using the certainty equivalents $C(b, X \mid \mathcal{B}) = \mathcal{U}^{-1}(E(\mathcal{U}(b + X) \mid \mathcal{B})) - b$, Theorem 11 remains true if we redefine

$\geq_r$ as follows. Let $C(b,p)$ denote the certainty equivalent of the coordinate random variable on $\mathcal{R}^1$ whose probability is the measure p and initial wealth is $b \in \mathcal{R}^1$. Denote by $\mathcal{P}^*_{\mathcal{U}}$ the collection of probability measures p on $(\mathcal{R}^1, \mathcal{B}_1)$ such that $\int \mathcal{U}(x)dp$ exists. We say $\mathcal{U}_1 \geq_{r^*} \mathcal{U}_2$ if $C_2(b,p) \geq C_1(b,p)$ for all $p \in \mathcal{P}^*_{\mathcal{U}_1} \cap \mathcal{P}^*_{\mathcal{U}_2}$ and all $b \in \mathcal{R}^1$. The first proof of Theorem 11 now goes through with $\geq_r$ replaced by $\geq_{r^*}$ and $E(X_t \mid \mathcal{J}_n) - \pi_i(b, X_t \mid \mathcal{J}_n)$ replaced by $C_i(b,X_t \mid \mathcal{J}_n)$, $i = 1,2$. Throughout the rest of this section, therefore, we assume $\tau_1 \leq \tau_2$ a.s.

The second problem can be handled in several ways. Certain decompositions of submartingales yield sufficient conditions for τ_2 to be admissable, but these conditions boil down to those of Lemma 10. Conditions similar to those of Lemma 10 can be used to handle the case (C3). The interested reader is referred to [7; Chapter 2] and [5; Chapter 5].

If we confine ourselves further to random walk models of stock prices we can obtain some interesting results. The two models we have in mind are the absolute and relative random walks. The absolute random walk specifies that

$$\begin{aligned} X_{n+1} &= X_n + Y_{n+1}, \quad n = 1,2,\ldots \\ X_1 &= Y_1 \end{aligned} \qquad (4.9)$$

where $Y_1, Y_2, \ldots$ are independent random variables. (4.7) in this case becomes

$$E(X_{n+1} \mid \mathcal{B}_n) = X_n + EY_{n+1}$$

and $\{X_n, \mathcal{B}_n;\ n \geq 1\}$ is in case (C1), (C2), or (C3) according as $EY_k \geq, =, \leq 0$, $k = 2,3,\ldots$ Although this model allows for negative prices, it offers an interesting application of Wald's lemma [7; Lemma 3.1].

Suppose $Y_1, Y_2, \ldots$ are also identically distributed with common mean μ, $|\mu| < \infty$. Wald's lemma then says that for any stopping rule t such that EX_t exists,

$$EX_t = \mu Et \tag{4.10}$$

provided $\mu = 0$ and $Et = \infty$ do not both hold. We then have

<u>Corollary 23.</u> If EX_{τ_1} and EX_{τ_2} exist, then

$$EX_{\tau_1} = \mu E\tau_1 \geq (>) \ (\leq) \ (<) \ \mu E\tau_2 = EX_{\tau_2}$$

provided $\mu < 0$ ($\mu < 0$, $E\tau_1 < \infty$, and $P(\tau_1 \neq \tau_2) > 0$) ($\mu > 0$)($\mu > 0$, $E\tau_1 < \infty$, and $P(\tau_1 \neq \tau_2) > 0$). If $\mu = 0$ and $E\tau_2 < \infty$, then $EX_{\tau_1} = EX_{\tau_2} = 0$.

Proof: Follows directly from (4.10) and $\tau_1 \leq \tau_2$ a.s. ($\Rightarrow E\tau_1 \leq E\tau_2$).

The relative random walk specifies that

$$\begin{aligned} X_{n+1} &= (1 + Z_{n+1})\, X_n, \ n = 1,2,\ldots \\ X_1 &\geq 0 \text{ given} \end{aligned} \tag{4.11}$$

where $X_1, Z_2, Z_3, \ldots$ are independent with $Z_k \geq -1$, $k = 2,3,\ldots$ (4.7) in this case becomes

$$E(X_{n+1} \mid \mathcal{B}_n) = (1 + EZ_{n+1})\, X_n$$

and $\{X_n, \mathcal{B}_n;\ n \geq 1\}$ is in case (C1), (C2), or (C3) according as $EZ_k \geq, =, \leq 0$, $k = 2,3,\ldots$ This model yields no specifically useful structure, but does give non-negative prices. When prices are non-negative, the hypotheses of [20; Theorem 28, p. 90] are satisfied and hence (C3) $\Rightarrow EX_{\tau_1} \geq EX_{\tau_2}$.

In handling cases (C1) and (C2) for non-negative prices one must either appeal to hypotheses like those of Lemma 10 or to economic restrictions on the selling problem that imply these hypotheses. For example, it may be that the individual who is selling the shares is required to sell if the price goes above some fixed constant or before some fixed date. Restrictions like these yield reward sequences that satisfy the hypotheses of Lemma 10.

FOOTNOTES

1. [1 : p.29]

2. See for example [2], [3], [18], [19], [21], [22], [24], [27], [30], [31], [32], and [34].

3. The theoretical and applied literature on optimal stopping problems is quite extensive. For the latter see [4], [6], [8], [13], [14], [15], [16], [17], and [33]. Our references to theoretical literature will be confined, for the most part, to [7] and [11], the former being a recent thorough treatment of the entire theory including an extensive bibliography.

4. Examining risk aversion properties of derived utility functions works in certain special cases, e.g., [24].

5. The entire study, including temporal risk aversion relations, constitutes the heart of my thesis [23].

6. By integrable we mean $\int |X| \, dP < \infty$. When we say that the expectation of X exists we mean that either X^+ is integrable or X^- is integrable (where $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$).

7. A random vector is a function, say Y, mapping Ω into some measureable space say $(\mathcal{R}, \mathcal{B})$, that is $\mathcal{F}$- measurable (again, $\mathcal{F}(Y)$ denotes the σ-algebra generated by Y). Thus a random variable is a real valued random vector.

8. See [5 : Chapter 4] for these definitions and further motivation. We will make free use of the concepts and results of this chapter of Brieman throughout the rest of the paper.

9. See [1] and [10] for axiom systems which imply boundedness of the utility function. For an axiom system that does not see [8; Chapter 7].

10. Notice that $E \mid \mathcal{U}(X) \mid < \infty$ implies $E\,\mathcal{U}(b + X)$ exists for all $b \in \mathcal{R}'$, i.e., either $E\,\mathcal{U}(b + X)^+ < \infty$ or $E\,\mathcal{U}(b + X)^- < \infty$. For $b > 0$, $\mathcal{U}(b + X)^- \leq \mathcal{U}(X)^-$ and hence $E\,\mathcal{U}(b + X)^- \leq E\,\mathcal{U}(X)^- < \infty$ and for $b < 0$, $\mathcal{U}(b + X)^+ \leq \mathcal{U}(X)^+$ and hence $E\,\mathcal{U}(b + X)^+ \leq E\,\mathcal{U}(X)^+ < \infty$.

11. Proofs may also be found in [12; Theorems 85 and 92].

12. Again we allow for the conditional risk premiums to take on the values $\pm \infty$.

13. Similar results obtain for the conditional risk premium $\pi(b, X \mid Y = y)$.

14. See [20; p. 29] for a proof.

15. Actually $E(X \mid Y = y)$ is an <u>estimator</u> of X since it is a function of the observation y. Strictly speaking, all references in the text should be to $E(X \mid Y = y)$ as an estimator of X, but the qualification of $E(X \mid Y = y)$ as a function of y as an <u>estimator</u> and for a particular y as an <u>estimate</u> seems cumbersome.

16. Similar results hold for $\pi(b, X \mid Y = y)$.

17. This principle is also important in defining temporal risk aversion relations in [23].

18. For the finite horizon case see [7] and [23].

19. X_t is a random variable defined on Ω, but we have given no conditions which ensure that X_t is integrable. This is a temporary departure from our assumption that all random variables are integrable. To state the O S P as we have below, and indeed to characterize solutions to the O S P we do not need that X_t be integrable. The proof of our basic results in sections 3.3, 4.1, and 4.2 are simplified greatly, however, if $E \mid X_t \mid < \infty$ for each $t \in T$. This condition is ensured in section 3.2 and relaxed somewhat in section 4.3.

20. Clearly we can incorporate a (possibly random) observation cost C_n with rewards X'_n. Then $X_n = X'_n - \sum_{k=1}^{n} C_k$.

21. [7; Chapter 3].

22. Notice that we get $EX_{\tau_2} > EX_{\tau_1}$ if (4.3) holds for each n and with strict inequality for some n with $P(\tau_1 = n) > 0$.

23. [7; p. 64].

24. See [9] for an excellent survey of both the theoretical and empirical literature on efficient market models of stock price behavior.

25. [6], [28], [29], and [33].

26. [9; p. 384].

BIBLIOGRAPHY!

1. Arrow, Kenneth J., Aspects of the Theory of Risk Bearing (Helsinki: Yrjö Jahnssonin Säätiö, 1965).

2. Baron, David P., "Price Uncertainty, Utility, and Industry Equilibrium in Pure Competition", International Economic Review, XI (October, 1970), 463-481.

3. ———, "Demand Uncertainty in Imperfect Competition", International Economic Review, XII (June, 1971), 196-208.

4. Brieman, Leo, "Stopping Rule Problems", Chapter 10 in Applied Combinatorial Mathematics, ed. Edwin E. Beckenbach (New York: John Wiley and Sons, 1964).

5. ———, Probability (Reading, Mass: Addison Wesley, 1968).

6. Chen, Andrew H. Y., "A Model of Warrant Pricing in a Dynamic Market", Journal of Finance, XXV (December, 1970), 1041-1059.

7. Chow, Y. S., Robbins, H., and Siegmund, D., Great Expectations: The The Theory of Optimal Stopping (Boston: Houghton Mifflin, 1971).

8. De Groot, Morris H., Optimal Statistical Decisions (New York: McGraw-Hill, 1970).

9. Fama, Eugene F., "Efficient Capital Markets: A Review of Theory and Empirical Work", Journal of Finance, XXV (May, 1970), 383-417.

10. Fishburn, Peter C., Utility Theory for Decision Making (New York: John Wiley and Sons, 1970).

11. Haggstrom, Gus W., "Optimal Stopping and Experimental Design", Annals of Mathematical Statistics, XXXVII (1966), 7-29.

12. Hardy, G. H., Littlewood, J. E., and Pólya G., Inequalities (Cambridge: Cambridge University Press, 1967).

13. Kaplan, Robert S.,"Stochastic Growth Models", Management Science, XVIII (January, Part 1, 1972), 249-264.

14. Karlin, Samuel, "Stochastic Models and Optimal Policies for Selling an Asset", Chapter 9 in Studies in Applied Probability and Management Science, K. J. Arrow, S. Karlin, and H. Scarf, eds., (Stanford University Press, 1962).

15. McCall, John J., "The Economics of Information and Optimal Stopping Rules", Journal of Business, XXXVIII (July, 1965), 300-317.

16. ———, "Economics of Information and Job Search", Quarterly Journal of Economics, LXXXIV(February, 1970), 113-126.

17. ———, "Probabilistic Microeconomics", Bell Journal of Economics and Management Science, II (Autumn, 1971), 403-433.

18. Menezes, C. F. and Hanson, D. L., "Risk Aversion and Bidding Theory", in Papers in Quantitative Economics, J. P. Quirk and A. M. Farley, eds. (Lawrence, Kansas: University of Kansas Press, 1968).

19. ———, "On the Theory of Risk Aversion", International Economic Review, XI (October, 1970), 481-487.

20. Meyer, Paul A., Probability and Potentials, (Waltham, Mass: Blaisdell, 1966).

21. Mossin, Jan.,"Optimal Multiperiod Portfolio Policies", Journal of Business, XLI (April, 1968), 215-229.

22. ———, "Aspects of Rational Insurance Purchasing", Journal of Political Economy, LXXVI (July-August, 1968), 553-568.

23. Nachman, David C., "On Risk Aversion and Sequential Decisions", unpublished manuscript, Northwestern University, Evanston, Illinois, (1972).

24. Neave, Edwin H., "Multiperiod Consumption - Investment Decisions and Risk Preference", Journal of Economic Theory, III (March, 1971), 40-53.

25. Neveu, J., Mathematical Foundations of the Calculus of Probability (San Francisco: Holden-Day, 1965).

26. Pratt, John W., "Risk Aversion in the Small and in the Large", Econometrica, XXXII (January-April, 1964), 122-136.

27. Rothschild, M., and Stiglitz, J. E., "Increasing Risk II: Its Economic Consequences", Journal of Economic Theory, III (March, 1971), 66-84.

28. Samuelson, Paul A., "Rational Theory of Warrant Pricing", Industrial Management Review, VI (Spring, 1965), 13-39.

29. ————, and Merton, Robert C., "A Complete Model of Warrant Pricing That maximizes Utility", Industrial Management Review, X (Winter, 1969), 17-46.

30. Sandmo, A., "Portfolio Choice in a Theory of Saving", Swedish Journal of Economics, LXX (1966), 106-122.

31. ———, "Capital Risk, Consumption and Portfolio Choice", Econometrica, XXXVII (October, 1969), 586-599.

32. Stiglitz, J. E., "The Effects of Income, Wealth and Capital Gains Taxation on Risk-Taking", Quarterly Journal of Economics, LXXXIII (May, 1969), 263-283.

33. Taylor, Howard M., "Evaluating a Call Option and Optimal Timing Strategy in the Stock Market", Management Science, XIV (September, 1967), 111-120.

34. Yaari, M. E., "Some Remarks on Measures of Risk Aversion and on Their Uses", Journal of Economic Theory, I(1969), 315-329.