THE FOERDER INSTITUTE FOR ECONOMIC RESEARCH

TEL-AVIV UNIVERSITY

RAMAT AVIV ISRAEL

מכון למחקר כלכלי ע״ש ד״ר ישעיהו פורדר ז״ל

ע״י אוניברסיטת תל־אביב

STOCHASTIC DOMINANCE IN MULTI SAMPLING ENVIRONMENTS

by

Arthur Fishman

Working Paper No.30-88

September, 1 9 8 8

ABSTRACT

Stochastic dominance rules are constructed for an environment in which the individual may sample sequentially according to a stopping rule. The dominance criterion used requires that all sampling take place in the dominant distribution with probability one even if the individual is permitted to switch during the course of sampling. The principal result is that the necessary dominance criterion is (conventional) first degree stochastic dominance even if the individual is known to be "very" risk averse. On the other hand, knowledge of an upper bound on risk aversion serves to rank random variables identically in both sampling and nonsampling environments.

This paper is based on ch.5 of my Ph.D thesis. I am grateful to my adviser Professor Eitan Sheshinski. The advice of an anonymous referee has helped to substantially improve the clarity and exposition of the paper. The standard caveat applies.

This research has been supported from funds granted to the Foerder Institute for Economic Research by the KEREN RAUCH FUND

FOERDER INSTITUTE FOR ECONOMICS RESEARCH
Faculty of Social Sciences,
Tel-Aviv University, Ramat Aviv, Israel

Introduction

Consider a set U of von Neumann-Morgenstern (VNM) utility functions and a pair of random variables, X and Y such that $u \in U$ must choose between them. The stochastic dominance literature addresses the following question: What relationship between the random variables must obtain so that each $u \in U$ maximizes expected utility by choosing X? As a stochastic dominance criterion induces only a partial ordering on the set of random variables, the focus of this literature has been to restrict U with the intent of obtaining a more complete ordering. With this objective in mind, several authors have imposed restrictions on the Arrow-Pratt risk aversion measure of U's elements. Thus Hadar and Russell (1969) and Hanoch and Levy (1969) construct dominance criteria for U all of whose elements are risk averse, while Whitmore (1970) focuses on utility functions with nonnegative third derivatives. Meyer (1977) presents a more general analysis, constructing dominance criteria for sets of utility functions whose Arrow-Pratt measure of risk aversion is bounded from below by an arbitrary constant.

An essential premise of this literature is that the individual's payoff is restricted to depend on a single random draw from the chosen distribution. In many economic contexts of interest, however, the individual may wish to sample repeatedly, rejecting unfavorable draws in anticipation of more desirable future outcomes. For example, in a world of imperfect job markets, a worker's wage in any particular profession is a random variable. She is not restricted, however, to accept the first job offered but may choose to search until more suitable employment is found.[1]

In this essay, therefore, I study a more general model in which up to n samples may be sequentially taken where n is any positive integer.[2] Specifically, following each draw, the individual may either terminate sampling and collect her payoff or incur a nonnegative sampling cost and draw a new sample from **either** of two fully specified probability distributions. Successive draws are i.i.d. and sampling must terminate after n draws have occurred. The payoff depends on the sampling "technology". If sampling is without recall, the payoff is the outcome of the last sample taken prior to termination. Sampling with recall describes a process in which the payoff is the **maximum** outcome of all samples taken prior to termination.[3]

The complexity involved in designing dominance rules for sampling environments stems from the fact that it is in general optimal to successively sample from **both** distributions as a simple example shows.

Example one: Sampling Without Recall

Suppose sampling is without recall, is costless and the sampler is risk neutral. Let F be the distribution which yields the payoffs -10 and +10 with probability 0.5 each and let G yield -21 and +20 with probability 0.5 each. If $n = 1$, F is preferred since its mean is larger. Supposes $n = 2$. The expected payoff from sampling sequentially from F **only** is calculated recursively as follows. The expected payoff from taking a second (and last) sample is simply F's expectation which is zero. Hence a second sample is taken if and only if the value of the first sample is $-10 < 0$, which occurs with probability 1/2. Thus with probability 0.5 the value of the first sample is 10 and no second sample is taken while with probability 0.5 the value of the first sample is -10 and a second sample whose expectation is zero is taken. The expected payoff from employing this optimal sequential sampling procedure on F is therefore:

$$0.5 \times 10 + 0.5 \times 0 = 5.$$

Similarly, a second sample is taken from G only if the value of the first sample is $-21 < -1/2$, G's expectation. Thus the value of using the sequential sampling procedure on G is:

$$0.5 \times 20 + 0.5 \times (-1/2) = 9.75 > 5.$$

There exists, however, a 'mixed' sampling rule which offers a higher payoff than either of the above procedures. Calculate the payoff from sampling first from G and, should it prove to be necessary, sampling a second time from F. Since a sample from F has an expectation of zero the first sample is accepted if and only if its value is 20 which occurs with probability 0.5. The expected payoff from this procedure is:

$$0.5 \times 20 + 0.5 \times 0 = 10 > 9.75.$$

In example 1, G is certainly the riskier of the two prospects and is appropriately rejected by a risk neutral individual if $n = 1$. Nevertheless, the same individual **strictly** prefers to first sample from G if $n \geq 2$. This example serves to identify the distinction between **underlying risk attitudes** (i.e. the concavity or convexity of the utility function) and **sampling risk attitudes** which describe preferences for risk in sampling problems. Even individuals whose underlying attitude is risk averse may display sampling attitudes which are risk loving.[4] Furthermore, while the underlying attitude is constant, sampling attitudes are determined by both the specification and the course of the sampling.[5] Thus, as example 1 suggests, and as is discussed more formally in section 3, the sampling attitude evolves towards increased risk aversion as n decreases if sampling is **without** recall. Intuitively, when a large number of samples remains available, the sampler can "afford" to assume risks she would otherwise avoid, being "cushioned" by the possibility of subsequently reverting to more conservative behavior. As the remaining available number of samples is depleted in the course of sampling, the sampling preference for risk steadily erodes, ultimately coinciding with the underlying risk attitude when only 1 sample remains. If sampling is **with** recall, however, the sampling attitude may also evolve towards **increasing** preference for risk: the individual's underlying and sampling attitude typically diverge even at the last stage, as the next example shows.

Example 2: Search with Recall

Suppose sampling is with recall. Let F be the distribution which assigns a probability of $1-p$ to zero and a probability $p > 0$ to $1 + \epsilon$, $\epsilon > 0$, let G be the uniform distribution with supp[0,1], let the sampling cost be zero and let $n = 2$. Every sampler whose underlying attitude is risk averse will initially sample from F if p and ϵ are sufficiently small. Since sampling is costless and the outcome of the first sample remains available, a second sample is taken with probability one. Since a second draw from G increases the expected payoff by $p\epsilon$, the last sample is optimally drawn from G if the realization of the first sample is sufficiently close to 1, an event which occurs with positive probability. The sampling preference for risk may therefore dramatically increase at the last draw if sampling is with recall.

The reason for this distinction between the two sampling environments is straightforward. Since sampling without recall is "memoryless" the sampling attitude towards risk is affected only by the remaining number of samples. By contrast, when sampling is with recall, the sampling attitude is determined by both history and remaining possibilities. The more favorable the history, the more the sampler is disposed towards the riskier prospect.

The preceding examples stress the importance of n in determining sampling attitudes towards risk. These are also affected by the size of the sampling cost as the desirability of switching to the less risky distribution at a later stage may be diminished when this cost is large.

To illustrate, consider example one with a sampling cost of 21. Since the increase in utility which derives from taking a second sample from either distribution is exceeded by the sampling cost whatever the outcome of the first draw, only one sample is ever observed. Under otherwise identical conditions, the same individual's sampling attitude displays greater risk aversion when the sampling cost is higher.

As sampling risk attitudes depend on the specifics of and location within a sampling problem, it is felt that meaningful dominance criteria applicable to general sampling environments should be independent of such arbitrary specifications. Accordingly we offer the following definition: F **sample dominates** G for a particular set of VNM utility functions if and only if each utility function in the set is maximized when every possible sample is optimally drawn from F, for every (integer) n and every nonnegative sampling cost.

We may state our main result. The knowledge that an individual's underlying attitude is risk averse conveys no useful information in the sampling context. More precisely, if the sampler's utility function is incompletely specified, knowledge of the extent to which her underlying attitude is risk averse cannot help determine her preferences in the sense of sampling dominance. On the other hand, knowledge of the extent to which the underlying attitude displays **preference** for risk is equally informative in both sampling and nonsampling contexts.

Our earlier discussion (of the distinction between underlying and sampling attitudes) suggests an intuitive explanation of the preceding result. Depending upon the specifics of the sampling problem, even inherently risk averse individuals may display sampling attitudes which are risk loving. Therefore an underlying attitude which is risk averse may be at odds with the individual's sampling attitude and barring a precise specification of the utility function it is not possible to determine which attitude "prevails". An underlying attitude which is risk loving, however, is complementary to the individual's sampling attitude and conveys useful information.

The remainder of the paper is organized as follows. In section 2, the properties of the relevant optimal sequential sampling rule are reviewed. These properties are used to derive sampling dominance criteria for important sets of VNM utility functions. Section 3 presents corroborating evidence pertaining to the relaionship between n and sampling preference for risk. Brief concluding remarks close the paper.

Section 2

This section begins with a brief review of the optimal sequential sampling rule when the maximum number of samples is finite and the sampling cost is constant. Proofs and additional details may be found in e.g. Landsberger and Peled (1977). Throughout it is assumed that sampling is without recall. In the appendix it is shown that all the theorems of this section apply without change if sampling is with recall.

A random variable X is generated identically and independently at each sample by a cumulative probability distribution with bounded support assumed for convenience to be [0,1]. Whenever a new sample is drawn, a fixed cost $c \geq 0$ is incurred and all previously observed outcomes are forfeited. It is well known that the optimal sampling rule associates a unique 'cutoff' value with each sample such that the outcome is accepted if and only if it is not exceeded by this value.

Let $V_n(y)$ be the expected utilty from optimal sequential sampling when y is the outcome of the preceding sample (still available) and at most n samples remain, and let R_n be the expected utility from sampling at least once more and continuing optimally when at most n samples remain. The following relationship between $V_n(\cdot)$ and R_n obtains:

$$(1) \qquad V_n(y) = \max\{u(y), R_n\}$$

where $u(\cdot)$ is a continuous nondecreasing VNM utility function.

Since the outcome of the n^{th} from last sample is accepted if and only if it is not less than R_{n-1}, by (1), R_{n-1} is the cutoff point associated with n. The following recursive relationship is established:

$$(2) \qquad R_n = \int_{R_{n-1}}^{1} u(x)dF(x) + R_{n-1}F(R_{n-1}) - c.$$

where $F(\cdot)$ is the distribution function of X.

It is easy to prove that for each integer n, $R_n \geq R_{n-1}$. Intuitively, as the number of remaining sampling opportunities decreases, the individual is increasingly reluctant to reject offers which may not be followed by more attractive ones. Let $R_n(F,c)$ and $R_n(G,c)$ denote the value of R_n when the distribution is F and G respectively and the sampling cost is $c \geq 0$.

Definition 1:

Let U_i be a set of VNM utility functions. F is said to **sampling dominate** G for U_i ($F \underset{\sim}{>}_{SDi} G$) if for every (integer) n and every $c \geq 0$, each $u \in U_i$ maximizes expected utility by drawing every possible sample from F with probability one.

In the sequel, attention is restricted to twice differentiable utility functions.

Define:

$$U_0 = \{u(\cdot) \mid u'(\cdot) \geq 0\}$$

$$U_1 = \{u(\cdot) \mid u'(\cdot) \geq 0,\ u''(\cdot) = 0\}$$

$$U_2 = \{u(\cdot) \mid u'(\cdot) \geq 0, u''(\cdot) \leq 0\}$$

$$U_3 = \{u(\cdot) \mid u'(\cdot) \geq 0,\ u''(\cdot) \geq 0\}.$$

U_0 is the set of general nondecreasing utility unctions, while U_i, $i = 1,2,3$, are the sets of (inherently) risk neutral, risk-averse and risk loving utility functions respectively.

Theorem 1: $F \gtrsim_{SD1} G$ iff:

(3) $\forall t \in [0,1], \quad \int_t^1 [G(x) - F(x)]dx \geq 0.$

Proof:

Necessity: Define Δ_n as the payoff from sampling once in $G(\cdot)$ and continuing any further sampling in $F(\cdot)$ when at most n samples remain. Since the cutoff value at the n-th from last sample is $R_{n-1}(F)$,

$$\Delta_n = \int_{R_{n-1}(F)}^1 xdG + R_{n-1}(F) \cdot G(R_{n-1}(F)) - c$$

By definition $F \gtrsim_{SD1} G$ only if $R_n(F) \geq \Delta_n$ for each n and each $c \geq 0$. Integrate $R_n(F) - \Delta_n$ by parts to obtain the necessary condition:

(4) $$\int_{R_{n-1(F)}}^1 [G(x) - F(x)]dx \geq 0$$

The following facts about R_n follow directly from (2):

(a) $R_n(\cdot)$ is a continuous monotonic decreasing function of c for each n.

(b) $R_n\ (c = 1) < 0$ for each n.

(c) R_n is a nondecreasing sequence for each $c \geq 0$.

(d) $\lim_{n\to\infty} R_n(c = 0) = 1$.

These facts establish that to each $t \in [0,1]$ there corresponds an integer $\tilde{n}$ and $\tilde{c} \geq 0$ such that $R_{\tilde{n}}(\tilde{c}) = t$. Thus (4) obtains for each n and $c \geq 0$ only if (3) is satisfied.

Sufficiency: Let t be the expected utility from sampling optimally when $n - 1$ samples at most remain. Thus, if the n-th from last sample is drawn from F or G and sampling continues optimally thereafter, the expected utilities are respectively:

$$z_F \equiv \int_t^1 x dF + t \cdot F(t) - c,$$

and

$$z_G \equiv \int_t^1 x dG + t \cdot G(t) - c.$$

Integrating by parts, $z_F - z_G = \int_t^1 [G(x) - F(x)]dx \geq 0$ by (3). This proves that if (3) obtains, the n^{th} from last sample is optimally drawn from F. Since this argument applies to any $n \geq 1$, the proof is complete. □

The significance of condition (3) is due to the fact that when it obtains all individuals whose underlying attitude is risk loving (weakly) prefer F when $n = 1$[6] as shown by Meyer (1977, theorem 5). Theorem 1 thus establishes that an individual whose underlying attitude is risk neutral displays a (weak) sampling preference for F for each n and c only if the latter is (weakly) preferred in nonsampling environments by all individuals whose underlying attitude is risk loving. The discussion in the introduction which associates large n (and small c) with increased sampling preference for risk provides an interpretation of this result. Suppose to the contrary that there exists a nonempty set of individuals, say W, whose underlying attitude is risk loving and who strictly prefer G when $n = 1$. Sufficiently large n and small c could then induce an inherently risk neutral individual to display a sampling preference for risk which "mimics" the underlying preferences of $w \in W$. Hence the requirement, expressed by (3), that no such set exists.

Meyer (1977) has introduced a concept of dominance with respect to a specific utility function. For arbitrary $k(\cdot) \in U_0$ define:

$$U_{\bar{k}(\cdot)} = \{u(\cdot) \mid -u''(\cdot)/u'(\cdot) \geq -k''(\cdot)/k'(\cdot),\ u'(\cdot) \geq 0\}.$$

Here the function $k(\cdot)$ serves as a lower bound to underlying risk aversion for $u \in U_{\bar{k}(\cdot)}$ in the sense of Arrow-Pratt. Analogously define the sets of utility functions which are more risk **loving** than arbitrary $k(x) \in U_0$:

$$U_{\underline{k}(\cdot)} = \{u(\cdot) \mid -u''(\cdot)/u'(\cdot) \leq -k''(\cdot)/k'(\cdot), \quad u'(\cdot) \geq 0\}.$$

In particular, if $k(\cdot)$ is risk neutral, $U_{\bar{k}(\cdot)}$ and $U_{\underline{k}(\cdot)}$ are U_2 and U_3 respectively.

Theorem 2: For any $k(\cdot) \in U_0$, $F \underset{\sim}{>}_{SD\bar{k}(\cdot)} G$ iff:

$$\forall t \in [0,1], \quad F(t) \leq G(t).$$

Proof: Apply the proof of Theorem 1 to find that $F \underset{\sim}{>}_{SD\bar{k}(\cdot)} G$ iff:

$$(5) \qquad \forall t \in [0,1], \quad \forall u \in U_{\bar{k}(\cdot)}, \qquad \int_t^1 (F-G)du \leq 0.$$

For $t' \in [0,1]$ replace the lower limits of the integrals in the statement and proof of Meyer's (1977) theorem 2 by t' to obtain the equivalence of (5) and the condition:

$$(6) \qquad \forall\ y \geq t', \quad \int_{t'}^{y} (F-G)dk \leq 0.$$

Combining (5) and (6) proves the theorem. □

Theorem 2 is our main result. It establishes the irrelevance of information about a lower bound on an individual's underlying aversion to risk in sampling environments. In particular, the knowledge that the individual is (inherently) risk averse is uninformative. By contrast, the conventional stochastic dominance literature stresses the role of risk aversion. No matter how large n is when sampling begins, there always exists a positive probability that each possible opportunity to sample will be exploited if c is suficiently small. The sampling preference for risk, which may initially exceed the individual's underlying risk posture, erodes to its underlying level by the time the last possible sample is obtained. Therefore F is sampling dominant only if it dominates G in the conventional (nonsampling) sense for both the set of individuals whose underlying attitude is less risk averse than $k(\cdot)$ and the set whose underlying attitude is more risk averse than $k(\cdot)$. This is, of course, possible only if F is first-degree dominant with respect to G.

Theorem 3: For any $k(\cdot) \in U_0$, $F \underset{\sim}{>}_{SD\underline{k}(\cdot)} G$ iff

$$\int_0^1 u(x)dF(x) \geq \int_0^1 u(x)dG(x)$$

for each $u \in U_{\underline{k}(\cdot)}$.

Proof: The proof is an application of Meyer's (1977) theorem 5 analogous to the proof of (our) theorem 2.

□

Theorem 3 establishes the equivalence of conventional and sampling dominance in the presence of an upper bound on underlying risk aversion. Intuitively, corresponding to any $\tilde{u} \in U_{\underline{k}(\cdot)}$, $\tilde{n} \geq 1$ and $\tilde{c} \geq 0$ there exists $\tilde{\tilde{u}} \in U_{\underline{k}(\cdot)}$ such that $\tilde{\tilde{u}}$'s underlying preference for risk is greater or equal to that of $\tilde{u}$ and such that $\tilde{\tilde{u}}$'s preferences when $n = 1$ are equivalent to those of $\tilde{u}$ when $n = \tilde{n}$ and $c = \tilde{c}$. Therefore, sampling and conventional dominance criteria coincide in this case.

Section Three

The perception offered in the preceding analysis that an increased number of samples corresponds to an increased affinity for risk suggests the following. Suppose F is riskier than G in the sense that every **sufficiently** risk loving individual (weakly) prefers F to G when $n = 1$. One would then expect that **any** individual known to (weakly) prefer F to G (in either one of the senses presented below) for some $\bar{n} < \infty$ continues to do so if the sample size is increased. For search without recall this is shown to be true for an important class of distributions. This includes pairs of cumulative distribution functions F and G which cross once at the most. We argue below that this "single crossing property" provides a natural formalization of the notion that "F is preferred by all sufficiently risk loving individuals".[7]

Definition 2: The distribution function $F(\cdot)$ said to cross the distribution function $G(\cdot)$ once at the most from above if there exists $0 \le y^* < 1$ such that

$$F(x) \ge G(x) \iff x \le y^*$$

Lemma 1: Suppose $F(x)$ crosses $G(x)$ once at the most from above. For any $k(\cdot) \in U_0$, $\int_0^1 k(x)dF(x) \ge \int_0^1 k(x)dG(x) \Rightarrow \int_0^1 u(x)dF(x) \ge \int_0^1 u(x)dG(x)$ for each $u \in U_{\bar{k}(\cdot)}$.

Proof: By the single crossing property and the nonnegativity of $dk(x)$, $\int_0^1 k(x)[dF(x) - dG(x)] = \int_0^1 [G(x) - F(x)]dk(x) \ge 0 \rightarrow \forall t \in [0,1]$, $\int_t^1 [G(x) - F(x)]dk(x) \ge 0$. By Theorem 5 of Meyer (1977) this implies that $\forall\ u(x) \in U_{\bar{k}(\cdot)}$, $\int_0^1 u(x)[dF(x) - dG(x)] \ge 0$. □

Note that the lemma does not obtain if F and G cross more than once. For example consider F and G as illustrated in figure 1.

[Figure One Here]

Assume $\int_0^1 (F-G)dx > 0$ and define:

$$v(x) = \begin{cases} x & \text{if} \quad x < y_1 \\ y_1 & \text{if} \quad x \geq y_1. \end{cases}$$

Clearly $\int_0^1 v(x)dF(x) > \int_0^1 v(x)dG(x)$. It is untrue, however, that F is preferred by all individuals who are more risk loving than $v(x)$. For example, risk neutral individuals strictly prefer G because it has the greater mean.

The following theorem applies to an environment in which the individual is restricted to sample from only one distribution. For any $\tilde{u} \in U_0$ let $R_{n,\tilde{u}}(F,c)$ and $R_{n,\tilde{u}}(G,c)$ respectively represent $R_n(F,c)$ and $R_n(G,c)$ for $\tilde{u} \in U_0$. Thus $R_{n,\tilde{u}}(F,c)$ $(R_{n,\tilde{u}}(G,c))$ refers to the expected value of sequentially sampling in F(G) **only**.

Theorem 4: Suppose $F(\cdot)$ and $G(\cdot)$ cross once at the most from above. For any $\tilde{u} \in U_0$, any $\bar{n}$ and any $\bar{c} \geq 0$, $R_{\bar{n},\tilde{u}}(F,\bar{c}) \geq R_{\bar{n},\tilde{u}}(G,\bar{c}) \Rightarrow$ $R_{n,\tilde{u}}(F,\bar{c}) \geq R_{n,\tilde{u}}(G,\bar{c})$ for any $n > \bar{n}$.

Proof: For any $\tilde{u} \in U_0$,

$$\forall y < 1, \quad \int_0^1 (G-F)d\tilde{u} \geq 0 \Rightarrow \int_y^1 (G-F)d\tilde{u} \geq 0$$

by the single crossing property. In this case F sample dominates G for $\tilde{u}$ and the theorem follows trivially. If $\int_0^1 (G-F)d\tilde{u} < 0$, $\exists \bar{y} \le y^*$ such that $\int_y^1 (G-F)du \ge 0$ as $y \underset{(\le)}{\ge} \bar{y}$ by the single crossing property and the nonnegativity of $d\tilde{u}$.

By assumption $R_{1,\tilde{u}}(G,\bar{c}) > R_{1,\tilde{u}}(F,\bar{c})$ and $R_{\bar{n},\tilde{u}}(F,\bar{c}) \le R_{\bar{n},\tilde{u}}(G,\bar{c})$, $\bar{n} > 1$.

Thus there exists $1 < m \le \bar{n}$ such that:

$$R_{m,\tilde{u}}(F,\bar{c}) \ge R_{m,\tilde{u}}(G,\bar{c})$$

$$R_{m-1,\tilde{u}}(F,\bar{c}) < R_{m-1,\tilde{u}}(G,\bar{c})$$

Since

$$\int_{R_{m-1,\tilde{u}}(G,\bar{c})}^1 (G-F)du > \int_{R_{m-1,\tilde{u}}(G,\bar{c})}^1 Gd\tilde{u} - \int_{R_{m-1,\tilde{u}}(F,\bar{c})}^1 Fd\tilde{u}$$

$$= R_{m,\tilde{u}}(F,\bar{c}) - R_{m,\tilde{u}}(G,\bar{c}) \ge 0,$$

it follows by the single crossing property that $R_{m-1,\tilde{u}}(G,\bar{c}) \ge \bar{y}$

$$\Rightarrow R_{m,\tilde{u}}(G,\bar{c}) \ge \bar{y}, \quad R_{m,\tilde{u}}(F,\bar{c}) \ge \bar{y}$$

$$\Rightarrow \forall n \ge m:$$

$$\min \{R_{n,\tilde{u}}(G,\bar{c}), \; R_{n,\tilde{u}}(F,\bar{c})\} \ge \bar{y}.$$

Now

$$R_{\bar{n}+1,\tilde{u}}(F,\bar{c}) - R_{\bar{n}+1,\tilde{u}}(G,\bar{c}) =$$

$$\int_{R_{\bar{n},\tilde{u}}(G,\bar{c})}^{1} G d\tilde{u} - \int_{R_{\bar{n},\tilde{u}}(F,\bar{c})}^{1} F d\tilde{u}$$

$$\geq \int_{R_{\bar{n},\tilde{u}}(F,\bar{c})}^{1} (G-F) d\tilde{u} \geq 0.$$

By the same argument one shows that for $\forall\, n \geq \bar{n} + 1$, $R_{n-1,\tilde{u}}(F,\bar{c}) \geq R_{n-1,\tilde{u}}(G,\bar{c}) \Rightarrow R_{n,\tilde{u}}(F,\bar{c}) \geq R_{n,\tilde{u}}(G,\bar{c})$ and the theorem is proved by induction. □

The next theorem observes that if sampling is not restricted to one distribution, the number of samples drawn from the riskier distribution increases as the sample size increases. Of course, the single crossing property is not sufficient to ensure that all samples are taken from F since this condition is in general not sufficient for sampling dominance.

Theorem 5: Suppose F crosses G once at the most from above. Any $\tilde{u} \in U_0$ which maximizes expected utility by drawing the first sample from F when the sample size is $\bar{n}$ and the sampling cost is $\bar{c} \geq 0$ also maximizes utility by drawing the first k+1 samples from F when the sample size is $\bar{n}+k$, $k \geq 1$, and the sampling cost is $\bar{c}$.

Proof: As the proof is similar to that of theorem 6, it is omitted.

Concluding Remarks

A concept of stochastic dominance appropriate to a sequential sampling environment in which the maximum number of samples is finite and sampling costs are constant has been defined and applied to important sets of utility functions. The main observation is that knowledge of a lower bound on risk aversion is uninformative in the sampling environment while knowledge of an upper bound is equally informative in both sampling and nonsampling environments. These results are a consequence of the assumption that n is finite which allows for sampling attitudes to evolve in the course of sampling. If n is infinite the sampling attitude cannot evolve; whichever distribution is initially preferred continues to be preferred for ever.

An alternative sampling environment in which sampling attitudes evolve is one in which n is infinite but search costs are increasing. I expect that very similar results to those presented are obtainable in this framework. The investigation of this conjecture is the subject of further research.

FOOTNOTES

1 The analysis is not related to the theory of order statistics because, due to the assumption that sampling is sequential the actual number of samples taken is a random variable whose distribution is determined by the sampling rule. This distinguishes the present essay from the literature on multivariate stochastic dominance (Paroush and Levy (1974).

2 A finite sample size may represent a financial constraint equal to n search costs. Alternatively one can think of environments in which the sampling must terminate prior to some deadline. An example is the academic job market.

3 When an infinite number of samples may be observed, it is well known that the two sampling environments are equivalent. This is not the case if the maximum possible number of samples is finite. See, for example, Landsberger and Peled (1977).

4 Thus Kohn and Shavell (1974) show that the effect of a mean presserving spread on a risk averse individual's expected utility from search is ambiguous.

5 If risk aversion is not constant, preferences may change as payments of sampling costs decrease the individual's wealth. This effect is ignored in the analysis.

6 Since it is assumed throughout that sampling is without recall, it is unambiguous to identify nonsampling environments with the

statement "n = 1". Such is **not** the case when sampling is with recall, as example 2 demonstrates. In that case, n = 1 identifies a nonsampling environment only if n = 1 before any sampling takes place.

7 Intuitively, the argument also applies if F is a mean preserving spread of G. Repeated efforts to prove this have been unsuccessful.

APPENDIX

In this appendix it is shown that all the theorems of section 2 apply to sampling with recall as well.

Let $W_n(y)$ be the expected utility from optimal sequential sampling with recall when the sample size is n and y is the highest sample observed to date (which remains available by the definition of sampling with recall) and let $Z_n(y)$ be the expected utility from sampling at least once more and continuing optimally with n and y as defined above. Then:

$$W_n(y) = \max\{u(y), Z_n(y)\}$$

There is associated with each n a cutoff value $\bar{x}_n$ such that $W_n(y) = Z_n(y)$ iff $y < \bar{x}_n$ (Landsberger and Peled (1977)). The recursive relationship between $z_n(y)$ and $Z_{n-1}(y)$ is then:

$$\text{(A.1)} \qquad Z_n(y) = \begin{cases} u(y)F(y) + \int_y^1 u(x)dF - c & \text{if } y \geq \bar{x}_{n-1} \\ z_{n-1}(y) \cdot F(y) + \int_y^{\bar{x}_{n-1}} z_{n-1}(x)dF & \\ + \int_{\bar{x}_{n-1}}^1 u(x)dF - c & \text{if } y < \bar{x}_{n-1}. \end{cases}$$

It can be shown (Landsberger and Peled (1977)) that $Z_n(y)$ is nondecreasing in y for each n. Let $Z_n(y, F, c)$ and $Z_n(y, G, c)$ denote $Z_n(y)$ for the sampling cost c and the distributions F and G respectively and similarly for $\bar{x}_n(F, c)$ and $\bar{x}_n(G, c)$.

For the case of risk neutrality, Landsberger and Peled (1977) show that $\forall\, n \geq 1,\ \bar{x}_n = \bar{x}$. This fact is used to derive the following claims.

Claim A.1: Suppose $u(x) = x$. Then

$$dZ_n(y)/dy = F(y) \qquad \text{if} \quad y \geq \bar{x}.$$

Proof: If $y \geq \bar{x}$, by (A.1)

$$Z_n(y) = yF(y) + \int_y^1 x dF - c$$

$$= 1 - \int_y^1 F dx - c.$$

Differentiation of the last expression proves the claim. □

Claim A.2: If $u(x) = x$ and $y < \bar{x}$,

$$dZ_n(y)/dy = (F(y))^n \rightarrow dZ_{n+1}(y)/dy = (F(y))^{n+1}$$

Proof: By (A.1) if $y < \bar{x}$,

$$Z_{n+1}(y) = F(y)Z_n(y) + \int_{\bar{x}}^{1} x dF + \int_{y}^{\bar{x}} Z_n(x) dF$$

$$= 1 - \int_{y}^{\bar{x}} F(x) dZ_n(x)$$

$$= 1 - \int_{y}^{\bar{x}} F(x) \cdot (F(x))^n dx.$$

The second equality is derived integrating by parts and simplifying, the third by assumption. Differentiation of the last expression proves the claim. □

Claim A.3: $Z_n(y)$ is everywhere quasi convex.

Proof: The claim follows directly from claims A.1 and A.2 by induction. □

Theorem A.1: Theorem 1 applies to search with recall.

Proof:

Necessity: If $c = 0$ it is always optimal to sample once more if $n = 1$ and $y < 1$. The value of taking the last sample from F and G respectively are $Z_1(y, F, 0)$ and $Z_1(y, G, 0)$ where $Z_1(y, F, 0) = yF(y) + \int_y^1 xdF - c$ and $Z_1(y, G, 0) = yG(y) + \int_y^1 xdG - c$. Integrating by parts,

$$Z_1(y, F, 0) - Z_1(y, G, 0) = \int_y^1 (G-F)dx.$$

If there exists $y' \in [0,1]$ such that $\int_{y'}^1 (G-F)dx < 0$, the last sample is optimally drawn from G if y' is the highest outcome available. Thus (3) is a necessary condition.

Sufficiency: Let $\Delta_n(G, y)$ and $\Delta_n(F, y)$ denote the expected utility from sampling once in G and F respectively and continuing optimally afterwards when n samples remain. (The argument c is suppressed if no ambiguity results.) Suppose for each y it is optimal to draw the $n\text{-}i^{th}$ from last sample from F with probability one, $i = 0,1,2,\ldots,n-1$. We now show that it is then optimal to draw the $n+1^{st}$ from last sample from F as well if condition (3) obtains. Define

$$\lambda_n^F(y) = \max\{y, Z_n(F, y)\}$$

Note that

$$d\lambda_n^F(y)/dy = \begin{cases} 1 & \text{if} \quad x > \bar{x} \\ (F(x))^n & \text{if} \quad x < \bar{x} \end{cases}$$

so that $d\lambda_n^F(y)/dy$ is nondecreasing. By assumption and using (A.1):

$$\Delta_{n+1}(y, F) =$$

$$Z_{n+1}(y,F) = \begin{cases} yF(y) + \int_y^1 x dF - c & \text{if} \quad y \geq \bar{x}(F) \\ Z_n(y)F(y) + \int_y^1 \lambda_n(x) dF - C & \text{if} \; y < \bar{x}(F) \end{cases}$$

$$\Delta_{n+1}(y, G) = \begin{cases} yG(y) + \int_y^1 x dG - c & \text{if} \quad y \geq \bar{x}(F) \\ Z_n(y, F)G(y) + \int_y^1 \lambda_n(X,F) dG - c & \text{if} \quad y < \bar{x}(F) \end{cases}$$

Integrating by parts,

$$\Delta_{n+1}(y,F) - \Delta_{n+1}(y,G) = \begin{cases} \int_y^1 [G(x) - F(x)]dx & \text{if} \quad y \geq \bar{x}(F) \\ \int_y^1 [G(x)-F(x)]d\lambda_n(x) & \text{if} \quad y < \bar{x}(F). \end{cases}$$

Clearly $\Delta_{n+1}(y,F) - \Delta_{n+1}(y,G) \geq 0$ for $y \geq \bar{x}(F)$ if (3) obtains. For $y < \bar{x}(F)$ it can be shown that $\Delta_{n+1}(y, F) - \Delta_{n+1}(y,G) \geq 0$ by (3), the fact that $d\lambda_n^F(x)/dx$ is nondecreasing, and using arguments analogous to those used by Hadar and Russel (1969) and Hanoch and Levy (1969) in deriving second degree stochastic dominance for concave functions. Thus the $n+1^{th}$ sample is optimally drawn from F if all subsequent samples are drawn from F and (3) obtains. Since by (3), $R_1(y,F) - R_1(y,G) = \int_y^1 (G-F)dx \geq 0$, the theorem is proved by induction. □

It is now easy to derive theorem 2 for search with recall. Clearly first-degree stochastic dominance is a sufficient condition. Necessity is proved analogously to the necessity proof presented in theorem A.1. The derivation of theorem 3 for search with recall is analogous to the proof in the text.

REFERENCES

Hanoch, G. and Levy, H. (1969), "Efficiency Analysis of Choices Involving Risk", **Review of Economic Studies**.

Hadar, J. and Russel, W. (1969), "Rules for Ordering Uncertain Prospects", **American Economic Review**, LIX.

Kohn, M.G. and Shavell, S. (1974), "The Theory of Search", **Journal of Economic Theory**, 9.

Landsberger, M. and Peleg, D. (1977), "Duration of Offers, Price Structure and the Gain from Search", **Journal of Economic Theory**, 16.

Meyer, J. (1977), "Second Degree Stochastic Dominance with Respect to a Function", **International Economic Review**, 18.

Paroush, J. and Levy, H. (1974), "Towards Multi-Variate Efficiency Criteria", **Journal of Economic Theory**, 7.

Whitmore, J. (1970), "Third Degree Stochastic Dominance", **American Economic Review**, Vol.60, 3.

Figure 1