

UCLA

Department of Statistics Papers

Title

Stochastic Dominance Under Bayesian Learning

Permalink

<https://escholarship.org/uc/item/0386c4f4>

Authors

Sushil Bikhchandani

Uzi Segal

Sunil Sharma

Publication Date

2011-10-24

STOCHASTIC DOMINANCE UNDER BAYESIAN LEARNING

SUSHIL BIKHCHANDANI

(Anderson School of Management, UCLA)

UZI SEGAL

(Department of Economics, University of Toronto)

SUNIL SHARMA

(Department of Economics, UCLA)

Working Paper #581
Department of Economics
University of California, Los Angeles
March 1990
Revised September, 1990

We are grateful to Ken Burdett, Seonghwan Oh, Lars Olson, Joe Ostroy, Zvi Safra and Lloyd Shapley for helpful comments.

1 Introduction

The concept of first-order stochastic dominance is usually defined on probability distributions over final outcomes (simple lotteries). Although this definition has been widely applied in economics, there are situations where it is inadequate. In this paper we analyze the concept of first-order stochastic dominance for probability distributions over probability distributions (compound lotteries) under Bayesian learning. We discuss conditions under which one person's beliefs dominate another person's beliefs by first-order stochastic dominance regardless of what they observe in common. We provide sufficient conditions on prior beliefs under which this is true. These conditions can be easily verified without taking any observations.

One may claim that it is unnecessary to define stochastic dominance relationships between updated beliefs. For instance, one may multiply probabilities in the compound lotteries and apply the usual definition of stochastic dominance to the resulting actuarially equivalent simple lotteries. However, the partial order on compound lotteries obtained by comparing actuarially equivalent simple lotteries may not be very useful from a normative standpoint. When a decision-maker faces a series of decisions, with some resolution of uncertainty between decisions, valuable information may be lost by multiplying probabilities in a distribution on distributions. In any setting where there is learning, such as in search models with learning or in the bandit problem, the usual concept of first-order stochastic dominance for distributions is inadequate. This is illustrated by the following example, adapted from Bikhchandani and Sharma (1989).

Let F_1 be a uniform distribution on $[0, 1]$ and let G_1, G_2 be uniform distributions on $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, respectively. F is the compound lottery which is degenerate at F_1 and G is the compound lottery which yields the simple lottery G_i with probability $\frac{1}{2}$, $i = 1, 2$. Consider a risk-neutral decision-maker searching sequentially (with or without recall) for the lowest price when the cost of each sample is $\frac{1}{4}$. It is well-known that when searching from a known distribution (i.e., a simple lottery) H , it is optimal to stop as soon as one observes a price less than some reservation price (see, for example, Lippman and McCall (1981)). The reservation price, r , is obtained by solving

$$c = \int_0^r H(x) dx \quad (1.1)$$

where c is the cost of each sample. Moreover, the minimum expected cost is equal to r . (The cost includes the price paid for the good as well as the sampling cost.) Thus, when searching from F , the minimum expected cost and the reservation price are both equal to $1/\sqrt{2}$. On the other hand, when searching from G , the decision-maker knows after exactly one observation whether he is searching from G_1 or G_2 . Using (1.1) it is easy to check that it is optimal to stop after the first observation from G , and the minimum expected cost is 0.75. Clearly, both F and G have the same actuarially equivalent simple lottery, and yet the expected cost is lower under F . The usual definition of first-order stochastic dominance, when

argue against this concept (for a discussion see Marschak (1975), and in particular de Finetti (1977)). We presume that they mean that when there is no learning probabilities of probabilities are equivalent to probabilities, but when there is learning the two are different. That, at least, is Totrep's² view in Kreps' (1988, p.150) splendid drama concerning this issue.

Our results are as follows. We show that when each underlying simple lottery is over the same two final outcomes, a compound lottery F dominates another compound lottery G by Bayes' first-order stochastic dominance if and only if F is a convex transformation of G (in a sense made precise later) if and only if the expected final outcome from any updated version of F is greater than the expected final outcome from an updated version of G using the same observations. This generalizes to the case where the underlying simple lotteries yield a finite number of final outcomes provided that the simple lotteries can be completely ordered by ordinary first-order stochastic dominance and satisfy an additional assumption. We also obtain sufficient conditions for Bayes' first-order stochastic dominance for the case when the underlying simple lotteries are not completely ordered by ordinary first-order stochastic dominance. These conditions require that the set of simple lotteries can be partitioned into disjoint subsets such that any two simple lotteries in two different subsets can be compared by ordinary first-order stochastic dominance. We obtain similar results for the alternative concept of stochastic dominance without the reduction axiom.

The paper is organized as follows. We give necessary and sufficient conditions for Bayes' first-order stochastic dominance in Section 2. An alternative concept of first-order stochastic dominance under learning, appropriate for decision-makers who do not subscribe to the reduction axiom, is discussed in Section 3. In both these sections it is assumed that the underlying simple lotteries can be ordered by ordinary first-order stochastic dominance. In Section 4 we present results for the case when the simple lotteries are not completely ordered. Applications to search theory, bandit problems and auction theory are discussed in Section 5. All proofs are in Appendix I.

2 Bayes' First-Order Stochastic Dominance

Let $L_s \equiv \{(P_1, P_2, \dots, P_M) : P_i > 0, \sum P_i = 1\}$ be the space of probability distributions on a finite set $\{x_1, x_2, \dots, x_M\}$, $x_1 < x_2 < \dots < x_M$. An element of L_s , $X_i = (P_{i1}, P_{i2}, \dots, P_{iM})$ represents a simple lottery yielding outcome x_j with probability P_{ij} .

Let $L_c \equiv \{(X_1, \alpha_1; X_2, \alpha_2; \dots; X_N, \alpha_N) : \alpha_i \geq 0, \sum \alpha_i = 1, X_i \in L_s\}$ be the space of probability distributions with finite support in L_s .³ Elements of L_c , called compound lotteries, are denoted by F, G , etc. The lottery $F = (X_1, \alpha_1; X_2, \alpha_2; \dots; X_N, \alpha_N)$ yields the simple lottery X_i with probability α_i . Since we will be interested in comparing two compound lotteries F and G , we can write without loss of generality

The expected final outcome under the posterior distribution $F(T)$ is

$$E[x|F(T)] = \sum_{j=1}^2 \sum_{i=1}^N x_j \alpha_i(T) P_{ij}$$

$E[x|G(T)]$ is similarly defined. For the case of two final outcomes it is easy to see that $F \succeq_1 G$ if and only if

$$E[x|F(T)] \geq E[x|G(T)], \quad \forall T \quad (2.3)$$

This equivalence has been shown by Berry and Fristedt (1985). However, (2.3) is an impractical condition for applications since it requires that the posterior distributions $F(T)$ and $G(T)$ be computed for all T . It is our aim here to find easily computable conditions equivalent to $F \succeq_1 G$.

Consider the following conditions on $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$, $G = (\beta_1, \beta_2, \dots, \beta_N) \in L_c$. Suppose there exist $a \geq 1$, $b \leq N$ such that

$$F = (0, \dots, 0, \alpha_a, \alpha_{a+1}, \dots, \alpha_N), \quad \alpha_i > 0, \quad \forall i \geq a \quad (2.4)$$

$$G = (\beta_1, \beta_2, \dots, \beta_b, 0, \dots, 0), \quad \beta_i > 0, \quad \forall i \leq b \quad (2.5)$$

Moreover, if $a < b$ then

$$\frac{\alpha_i}{\alpha_i + \alpha_{i+1}} \leq \frac{\beta_i}{\beta_i + \beta_{i+1}}, \quad a \leq i < b \quad (2.6)$$

or equivalently

$$\frac{\alpha_i}{\alpha_{i+1}} \leq \frac{\beta_i}{\beta_{i+1}}, \quad a \leq i < b$$

Conditions (2.4) and (2.5) imply that if for some i , $\alpha_i = 0$ and $\beta_i > 0$ then $\alpha_k = 0$, $\forall k < i$. Also, if for some i , $\alpha_i > 0$, $\beta_i = 0$ then $\beta_k = 0$, $\forall k > i$. Condition (2.6) implies that if $\alpha_i, \alpha_{i+1}, \beta_i, \beta_{i+1} > 0$ for some i then the conditional distribution of F on X_i, X_{i+1} dominates the conditional distribution of G on X_i, X_{i+1} by $FOSD_1$.

The main result of this section is stated below. The proof is omitted since Theorem 1 is a special case of Theorem 2 below.

Theorem 1: Let $X_1, X_2, \dots, X_N \in L_c$ be simple lotteries with two final outcomes x_1 and x_2 . Let $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$ and $G = (\beta_1, \beta_2, \dots, \beta_N)$ be compound lotteries with outcomes in the set $\{X_1, X_2, \dots, X_N\}$. The following statements are equivalent:

- (i) $F \succeq_1 G$;
- (ii) $E[x|F(T)] \geq E[x|G(T)], \quad \forall T$;
- (iii) F, G satisfy (2.4), (2.5) and (2.6);
- (iv) F, G satisfy (2.4) and (2.5), and there exists a convex function, f , such that $F_i = f(G_i), \quad \forall i \in \{a, a+1, \dots, b\}$, where a and b are defined in (2.4) and (2.5), and $F_i = \sum_{r=1}^i \alpha_r, \quad G_i = \sum_{r=1}^i \beta_r$ are the cumulative distribution functions of F and G respectively. When F and G have the same support, that is, when

final outcomes X_1, X_2, \dots, X_N so that $P_{11} > P_{21} > \dots > P_{N1}$. Clearly X_1, X_2, \dots, X_N satisfy (2.8), (2.9) and (2.10). For the case of three final outcomes types 1, 2 and 3 are illustrated in Figures 1, 2 and 3 respectively. These triangle diagrams show simple lotteries in the $P_1 - P_3$ space.⁵ Choose any X_i and plot it on a triangle diagram as shown in Figure 1. All points (P_1, P_2, P_3) below the line AE satisfy $P_2/P_1 \geq P_{i2}/P_{i1}$, and all points above the line BD satisfy $P_3/P_2 \geq P_{i3}/P_{i2}$. Thus (2.8) implies that if X_1, X_2, \dots, X_N are of type 1 then X_{i+1} must lie in the region ACD, and X_{i-1} must lie in BCE.

A similar argument establishes that if X_1, X_2, \dots, X_N are of type 2 and X_i is as shown in Figure 2, then X_{i+1} must lie in the region AFCD, and X_{i-1} must lie in HCE. Also, if X_1, X_2, \dots, X_N are of type 3 and X_i is as in Figure 3, then X_{i+1} is in FCD and X_{i-1} is in HCEB.

In each of Figures 1, 2 and 3, all simple lotteries that dominate X_i (by $FOSD_1$) are north-west of X_i , and X_i dominates all simple lotteries to its south-east. Thus, at least for three outcomes, if X_1, X_2, \dots, X_N are of type 1, 2 or 3, then $X_N FOSD_1 X_{N-1} \dots FOSD_1 X_1$. The following lemma establishes this in general.

Lemma 2: *If X_1, X_2, \dots, X_N are of type 1, 2 or 3, then $X_N FOSD_1 X_{N-1} \dots FOSD_1 X_1$.*

The converse of Lemma 2 is not true. The simple lotteries X_1, X_2, X_3 , and X_4 in Figure 5 (see below) can be ordered by $FOSD_1$ but they are not of type 1, 2 or 3. Also, there exist X_1, X_2, \dots, X_N which are of types 1, 2 and 3. For example, we can choose X_1, X_2, \dots, X_N which satisfy (2.8) with $P_{ij} = P_{i+1,j}$, $j = 2, 3, \dots, M - 1$, $i = 1, 2, \dots, N - 1$. In the case of three outcomes, if X_1, X_2, \dots, X_N lie on the line DH in Figure 2 then they are of types 1, 2 and 3.

If the underlying simple lotteries are of type 1, 2 or 3, then for any i such that $\alpha_i, \alpha_{i+1}, \beta_i, \beta_{i+1} > 0$ it is possible to find a sequence of observations T such that (i) $F(T)$ and $G(T)$ assign most of their mass to X_i and X_{i+1} , and (ii) the relative mass assigned to X_i and X_{i+1} by $F(T)$ and $G(T)$ can be made arbitrarily close to the relative mass assigned to them by F and G , respectively. Therefore, $E[F(T)] FOSD_1 E[G(T)]$ only if $\alpha_i(T)/\alpha_{i+1}(T) \leq \beta_i(T)/\beta_{i+1}(T)$ only if $\alpha_i/\alpha_{i+1} \leq \beta_i/\beta_{i+1}$. Thus (2.6) is necessary for $F \succeq_1 G$. The main result of this section generalizes Theorem 1.

Theorem 2: *Let $X_1, X_2, \dots, X_N \in L$, be of type 1, 2 or 3. Let $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$ and $G = (\beta_1, \beta_2, \dots, \beta_N)$ be compound lotteries with outcomes in the set $\{X_1, X_2, \dots, X_N\}$. The following statements are equivalent:*

- (i) $F \succeq_1 G$;
- (ii) $E[x|F(T)] \geq E[x|G(T)], \forall T$;
- (iii) F, G satisfy (2.4), (2.5) and (2.6);
- (iv) F, G satisfy (2.4) and (2.5), and there exists a convex function, f , such that $F_i = f(G_i), \forall i \in \{a, a+1, \dots, b\}$, where a and b are defined in (2.4) and (2.5), and $F_i = \sum_{r=1}^i \alpha_r, G_i = \sum_{r=1}^i \beta_r$ are the cumulative distribution functions of F and G respectively.

- (i) if F $FOSD_2$ G then $E[F]$ $FOSD_1$ $E[G]$;
(ii) if $F \succeq_2 G$ then $F \succeq_1 G$.

The converse of Lemma 4(i) is not true (see Segal (1990)). For example, suppose that there are two final outcomes x_1 and x_2 and that $X_1 = (\frac{2}{3}, \frac{1}{3})$, $X_2 = (\frac{1}{2}, \frac{1}{2})$, $X_3 = (\frac{1}{3}, \frac{2}{3})$ are simple lotteries over x_1 and x_2 . Thus under X_1 the probability of x_1 is $\frac{2}{3}$ and so on. Let $F = (\alpha_1, \alpha_2, \alpha_3) = (0, 1, 0)$ and $G = (\beta_1, \beta_2, \beta_3) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$. Since $\sum_{i=1}^3 \alpha_i P_{i2} > \sum_{i=1}^3 \beta_i P_{i2}$, (2.2) implies that $E[F]$ $FOSD_1$ $E[G]$. But since $\alpha_1 + \alpha_2 > \beta_1 + \beta_2$, Lemma 3 implies that F $FOSD_2$ G does not hold.

Lemma 1 established that when the underlying simple lotteries can be ordered by $FOSD_1$, (2.4), (2.5) and (2.6) are sufficient for $F \succeq_1 G$. It turns out that they are also sufficient for $F \succeq_2 G$.

Lemma 5: Suppose that X_N $FOSD_1$ X_{N-1} ... $FOSD_1$ X_1 . Let $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$ and $G = (\beta_1, \beta_2, \dots, \beta_N)$ be compound lotteries with outcomes in the set $\{X_1, X_2, \dots, X_N\}$. If F and G satisfy (2.4), (2.5) and (2.6) then $F \succeq_2 G$.

Example A.2 of Appendix II shows that the converse of Lemma 4(ii) is not true. However, if the underlying simple lotteries are of type 1, 2 or 3, then \succeq_1 and \succeq_2 are equivalent.

Theorem 3: Let $X_1, X_2, \dots, X_N \in L_s$ be of type 1, 2 or 3. Let $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$ and $G = (\beta_1, \beta_2, \dots, \beta_N)$ be compound lotteries with outcomes in the set $\{X_1, X_2, \dots, X_N\}$. Then $F \succeq_1 G$ if and only if $F \succeq_2 G$.

Under the hypothesis of Theorem 3, $FOSD_1$ and $FOSD_2$ are not equivalent. This follows from the example after Lemma 4, where the simple lotteries are of type 1, 2 and 3.

4 Unordered Simple Lotteries

In this section we investigate the general case when X_1, X_2, \dots, X_N are not completely ordered by $FOSD_1$. First, we obtain sufficient conditions for \succeq_1 and \succeq_2 generalizing (2.4), (2.5) and (2.6). These sufficient conditions require that X_1, X_2, \dots, X_N admit a D-partition,⁷ that is, a partition such that all simple lotteries in any element of the partition are comparable by $FOSD_1$ to all simple lotteries in other elements of the partition. We close this section with a necessary condition for \succeq_2 .

Consider the following definition.

Definition 8: Let X_1, X_2, \dots, X_N be simple lotteries, and let $\{I_1, I_2, \dots, I_D\}$, $D \geq 2$, be a partition of $\{1, 2, \dots, N\}$. That is, for $i \neq k$, $I_i \cap I_k = \phi$, and $\cup_{i=1}^D I_i = \{1, 2, \dots, N\}$. Then $\{I_1, I_2, \dots, I_D\}$ is said to be a D-partition if $\forall i > k, \forall r \in I_i, \forall s \in I_k, X_r$ $FOSD_1$ X_s .

In Sections 2 and 3, where we assumed X_N $FOSD_1$ X_{N-1} ... $FOSD_1$ X_1 , the sim-

Since $d_{n-1}(A) \subseteq d_n(A) \subseteq \mathcal{X}$, and \mathcal{X} is finite, there exists n^* such that $d_{n-1}(A) = d_n(A) = d_\infty(A)$, $\forall n > n^*$. For all n , $d_n(\{X_1\})$ is a set which contains X_1 and has the property that it cannot be partitioned into two disjoint subsets such that all simple lotteries in one subset can be compared by $FOSD_1$ to all simple lotteries in the other subset. The maximal set with this property is $d_\infty(\{X_1\})$. Thus if $d_\infty(\{X_1\}) \neq \mathcal{X}$, then any $X_k \in \mathcal{X} \setminus d_\infty(\{X_1\})$ (where \setminus stands for set difference), can be compared by $FOSD_1$ to any $X_i \in d_\infty(\{X_1\})$.

Lemma 6: *The set of simple lotteries, \mathcal{X} , admits a D-partition if and only if $d_\infty(\{X_1\}) \neq \mathcal{X}$.*

The proof of Lemma 6 is constructive and can be used to obtain a D-partition, when one exists. In fact the finest D-partition can be obtained by repeatedly applying the procedure in the proof.

Next we turn to an equivalent set of sufficient conditions for $F \succeq_1 G$ and $F \succeq_2 G$. Consider the following conditions

$$\text{if } X_i \text{ } FOSD_1 \text{ } X_k \text{ then } \alpha_i \beta_k - \beta_i \alpha_k \geq 0 \quad (4.5)$$

$$\text{if } X_i \text{ and } X_k \text{ are not comparable by } FOSD_1 \text{ then } \alpha_i \beta_k - \beta_i \alpha_k = 0 \quad (4.6)$$

Although (4.5) and (4.6) do not explicitly require the existence of a D-partition, they turn out to be equivalent to the hypothesis of Theorem 4.

Theorem 5: *Let $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$, $G = (\beta_1, \beta_2, \dots, \beta_N)$, $F \neq G$, be compound lotteries with outcomes in the set $\{X_1, X_2, \dots, X_N\}$. Then F and G satisfy (4.5) and (4.6) if and only if $\{X_1, X_2, \dots, X_N\}$ admits a D-partition and (4.1)–(4.4) are satisfied by F and G .*

When $\alpha_i, \beta_i > 0$, $\forall i$, Theorems 4 and 5 imply the following corollary.

Corollary 1: *Let $F, G \in L_c$, $\alpha_i, \beta_i > 0$, $\forall i$. Suppose that*

- (i) *if $X_i \text{ } FOSD_1 \text{ } X_k$ then $\frac{\alpha_k}{\alpha_k + \alpha_i} \leq \frac{\beta_k}{\beta_k + \beta_i}$;*
 - (ii) *if X_i and X_k are not comparable by $FOSD_1$ then $\frac{\alpha_k}{\alpha_k + \alpha_i} = \frac{\beta_k}{\beta_k + \beta_i}$.*
- Then $F \succeq_1 G$ and $F \succeq_2 G$.*

We now obtain necessary conditions for \succeq_2 . It is shown that if the simple lotteries X_1, X_2, \dots, X_N can be partitioned into two disjoint subsets such that the simple lotteries in the first subset cannot be compared by $FOSD_1$ to the simple lotteries in the second subset, then any two distinct compound lotteries F, G which yield outcomes in the set $\{X_1, X_2, \dots, X_N\}$ cannot be compared by \succeq_2 .

Theorem 6: *Let $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$, $G = (\beta_1, \beta_2, \dots, \beta_N)$ be distinct compound lotteries with outcomes in the set $\{X_1, X_2, \dots, X_N\}$. Suppose that (after relabeling the simple lotteries, if necessary),*

- (i) *there exists K such that $\forall i \leq K, \forall k > K, X_i$ and X_k are not comparable by $FOSD_1$;*
- (ii) *if $\forall i, i' \leq K, \forall k, k' > K, P_{ij} P_{kj} = P_{i'j} P_{k'j}, \forall j$, then $i = i', k = k'$.*

can search for the lowest price either from F or from G . Further, once he chooses F , say, and elicits a few price samples from F he is not allowed to search from G . Are there conditions on F and G such that, without actually computing the expected minimum prices, one can determine which one of the two distributions on distributions will be preferred by the individual?

We show that $FOSD_1$ cannot be used to choose between F and G . Let $X_1 = (1, 0)$, $X_2 = (\frac{1}{2}, \frac{1}{2})$ and $X_3 = (0, 1)$ be simple lotteries over two final outcomes x_1 and x_2 , $x_1 < x_2$. That is, X_1 yields x_1 with probability 1, etc. Let $F = (0, 1, 0)$, $G = (\frac{2}{3}, 0, \frac{1}{3})$, and $\hat{F} = (\frac{1}{2}, 0, \frac{1}{2})$ be distributions on the set $\{X_1, X_2, X_3\}$. First, consider the case when the individual is allowed to take two price samples from either F or G . Let $y_k \in \{x_1, x_2\}$, $k = 1, 2$ denote the k th sample observation. Since

$$E[\min(y_1, y_2)|F] = \frac{3}{4}x_1 + \frac{1}{4}x_2 < \frac{2}{3}x_1 + \frac{1}{3}x_2 = E[\min(y_1, y_2)|G]$$

F is preferred to G . If, instead, the choice is between \hat{F} and G , and again only two price samples are allowed, then

$$E[\min(y_1, y_2)|\hat{F}] = \frac{1}{2}x_1 + \frac{1}{2}x_2 > \frac{2}{3}x_1 + \frac{1}{3}x_2 = E[\min(y_1, y_2)|G]$$

implies that G is preferred to \hat{F} . Since $E[F] = E[\hat{F}] = (\frac{1}{2}, \frac{1}{2})$ and $E[G] = (\frac{2}{3}, \frac{1}{3})$, it follows that $E[F] \text{ } FOSD_1 \text{ } E[G]$ and $E[\hat{F}] \text{ } FOSD_1 \text{ } E[G]$. Hence, $FOSD_1$ is inadequate in this setting. As the next lemma shows, Bayes' first-order stochastic dominance can be used to choose between F and G .

Lemma 7: Let X_1, X_2, \dots, X_N be distributions on the set of prices, $\{x_1, x_2, \dots, x_M\}$, and let F and G be distributions on the set $\{X_1, X_2, \dots, X_N\}$. If X_1, X_2, \dots, X_N are of type 1 and $F \succeq_1 G$, then for search with recall over a finite horizon the expected minimum price under G is less than the expected minimum price under F .

In Bikhchandani and Sharma (1989) the problem of search from distributions on distributions is considered in an infinite horizon setting (with a strictly positive cost of obtaining a sample). It is shown that if $F \succeq_1 G$, and either F or G is increasing, then a risk-neutral decision-maker prefers G to F when searching with or without recall for the lowest price. Moreover, if the optimal stopping rules under both F and G are reservation price policies, then after any history of prices whenever it is optimal to stop under G , it is also optimal to stop under F .

Another application of Bayes' first-order stochastic dominance is to the bandit problem. Consider a decision-maker facing a sequential choice from a finite set of stochastic processes (also called arms, machines, etc.). At each stage the decision-maker chooses one arm and obtains a reward which depends on the observation made from the chosen arm. In the interesting case, the distribution of at most one arm is known. The decision-maker updates his prior distribution after each observation from an arm whose distribution is not known. At each stage a strategy

Appendix I

Proofs

We need the following result to prove Lemma 1.

Lemma A.1: Let F and G be as in (2.4) and (2.5) respectively. Then (2.6) implies

$$\sum_{i=1}^{\ell} \alpha_i \leq \sum_{i=1}^{\ell} \beta_i, \quad \forall \ell < N \quad (\text{A.1})$$

Proof of Lemma A.1: If $b \leq a$, (A.1) follows trivially. Suppose that $b > a$. If $\ell \in \{1, 2, \dots, a-1\}$ or $\ell \in \{b, b+1, \dots, N\}$ then again (A.1) follows trivially. Suppose that there exists $\ell \in \{a, a+1, \dots, b-1\}$ such that

$$\sum_{i=1}^{\ell} \alpha_i > \sum_{i=1}^{\ell} \beta_i, \quad (\text{A.2})$$

Without loss of generality we may assume that $\sum_{i=1}^s \alpha_i \leq \sum_{i=1}^s \beta_i$, $\forall s < \ell$. Therefore, $\alpha_\ell > \beta_\ell$. From (2.6) we know that $\alpha_i/\beta_i \leq \alpha_{i+1}/\beta_{i+1}$, $a \leq i < b$. Thus $\alpha_i > \beta_i$, $\ell \leq i \leq b$. This, together with (A.2), implies $\sum_{i=1}^b \alpha_i > \sum_{i=1}^b \beta_i = 1$, which contradicts the fact that $\sum_{i=1}^b \alpha_i \leq 1$. Thus (A.1) must be true. ■

Proof of Lemma 1: From Lemma A.1 we know that (2.4), (2.5) and (2.6) imply (A.1). Since $X_N \text{ FOSD}_1 X_{N-1} \dots \text{ FOSD}_1 X_1$, it is easily checked that (A.1) implies $E[F] \text{ FOSD}_1 E[G]$. Thus it is sufficient to show that after taking one sample the updated distributions satisfy (2.4), (2.5) and (2.6), and the lemma follows by repeated application. Let T_j be an M-vector with a 1 in the j^{th} place and 0 everywhere else. Thus T_j represents a sample of size one in which x_j was observed. We will show that $F(T_j)$ and $G(T_j)$ satisfy (2.4), (2.5) and (2.6).

Since $\alpha_i(T_j) > 0$ if and only if $\alpha_i > 0$, (2.4) is automatically satisfied by $F(T_j)$. Similarly, (2.5) is satisfied by $G(T_j)$. Suppose that $b > a$. Choose i such that $a \leq i < b$. Since F, G satisfy (2.6), we know that $\alpha_i/\alpha_{i+1} \leq \beta_i/\beta_{i+1}$. Also, since $\alpha_i(T_j)/\alpha_{i+1}(T_j) = (P_{i,j}\alpha_i)/(P_{i+1,j}\alpha_{i+1})$, and $\beta_i(T_j)/\beta_{i+1}(T_j) = (P_{i,j}\beta_i)/(P_{i+1,j}\beta_{i+1})$ we have

$$\frac{\alpha_i(T_j)}{\alpha_{i+1}(T_j)} \leq \frac{\beta_i(T_j)}{\beta_{i+1}(T_j)} \quad (\text{A.3})$$

But (A.3) implies that $F(T_j)$ and $G(T_j)$ satisfy (2.6). ■

In the discussion on type 1 we claimed that (2.7) and (2.8) are equivalent. A proof is provided below.

Lemma A.2: The simple lotteries X_1, X_2, \dots, X_N are of type 1 if and only if

$$\frac{P_{i,j+1}}{P_{i,j}} \leq \frac{P_{i+1,j+1}}{P_{i+1,j}}, \quad \forall j < M, \forall i = 1, 2, \dots, N-1$$

Similarly, for all $\ell > \ell^*$,

$$\sum_{j=\ell}^M P_{ij} = \sum_{j=\ell}^M \lambda_j P_{i+1,j} \leq \lambda_\ell \sum_{j=\ell}^M P_{i+1,j} \leq \sum_{j=\ell}^M P_{i+1,j} \quad (\text{A.6})$$

Inequalities (A.5) and (A.6) imply X_{i+1} $FOSD_1$ X_i . ■

Lemmas A.3 and A.4 below are required to prove Theorem 2. First, some notation.

For $\theta_j \in [0, 1]$, $\sum \theta_j = 1$, define

$$C_i(\theta_1, \theta_2, \dots, \theta_M) \equiv \prod_{j=1}^M (P_{ij})^{\theta_j}, \quad 1 \leq i \leq N \quad (\text{A.7})$$

$$C_{ki}(\theta_1, \theta_2, \dots, \theta_M) \equiv \frac{C_k(\theta_1, \theta_2, \dots, \theta_M)}{C_i(\theta_1, \theta_2, \dots, \theta_M)}, \quad 1 \leq i, k \leq N \quad (\text{A.8})$$

We can rewrite (2.1) as

$$\alpha_i(T) = \frac{\alpha_i}{\sum_{k=1}^N \alpha_k \left(C_{ki} \left(\sum_{t_j}^{i_1}, \sum_{t_j}^{i_2}, \dots, \sum_{t_j}^{i_M} \right) \right)^{\sum t_j}}, \quad 1 \leq i \leq N \quad (\text{A.9})$$

Lemma A.3: Suppose that X_1, X_2, \dots, X_N are of type 1, 2 or 3. For any $k < i$ there exist $\theta_1^{ki}, \theta_2^{ki}, \dots, \theta_M^{ki}$ such that

$$C_{ri}(\theta_1^{ki}, \theta_2^{ki}, \dots, \theta_M^{ki}) = \begin{cases} 1, & \text{if } r \in \{k, i\} \\ < 1, & \text{if } r \in \{1, 2, \dots, k-1\} \cup \{i+1, \dots, N\} \end{cases}$$

Proof of Lemma A.3: Suppose that X_1, X_2, \dots, X_N are of type 1.

Let $S_{i\ell} = \sum_{j=1}^{\ell} P_{ij}$, $\forall \ell \leq M$. Then,

$$\begin{aligned} \ln C_i(\theta_1, \dots, \theta_M) &= \sum_{\ell=1}^{M-1} \theta_\ell \ln(S_{i\ell} - S_{i,\ell-1}) + (1 - \sum_{\ell=1}^{M-1} \theta_\ell) \ln(S_{iM} - S_{i,M-1}) \\ \Rightarrow \frac{\partial \ln C_i}{\partial S_{i\ell}} &= \frac{\theta_\ell}{S_{i\ell} - S_{i,\ell-1}} - \frac{\theta_{\ell+1}}{S_{i,\ell+1} - S_{i\ell}} = \frac{\theta_\ell}{P_{i\ell}} - \frac{\theta_{\ell+1}}{P_{i,\ell+1}}. \end{aligned}$$

Thus $\partial \ln C_i / \partial S_{i\ell} \geq 0$, $\forall \ell < M$ if and only if

$$\frac{\theta_\ell}{\theta_{\ell+1}} \geq \frac{P_{i\ell}}{P_{i,\ell+1}}, \quad \forall \ell < M \quad (\text{A.10})$$

By Lemma 2, $S_{r\ell} \leq S_{i\ell}$, $\forall r > i, \forall \ell < M$. Thus, if we choose $(\theta_1, \dots, \theta_M)$ satisfying the inequalities (A.10) then

Without loss of generality we assume that $\alpha_r = 0, \forall r \in \{k+1, k+2, \dots, i\}$. From the proof of Lemma A.3 it is clear that if X_1, X_2, \dots, X_N are of type t then $(\theta_1^{ki}, \dots, \theta_M^{ki})$ and $(\theta_1^{rs}, \dots, \theta_M^{rs}), r > k, s > i$, when considered as simple lotteries, are also of type t with $(\theta_1^{rs}, \dots, \theta_M^{rs})$ $FOSD_1$ $(\theta_1^{ki}, \dots, \theta_M^{ki})$. Thus we can find $(\theta_1^*, \dots, \theta_M^*)$ such that $(\theta_1^{ki}, \dots, \theta_M^{ki}), (\theta_1^*, \dots, \theta_M^*)$ and $(\theta_1^{i,i+1}, \dots, \theta_M^{i,i+1})$ are of type t with $(\theta_1^{i,i+1}, \dots, \theta_M^{i,i+1})$ $FOSD_1$ $(\theta_1^*, \dots, \theta_M^*)$ $FOSD_1$ $(\theta_1^{ki}, \dots, \theta_M^{ki})$, and

$$C_n(\theta_1^*, \dots, \theta_M^*) < 1, \quad \forall r \in \{1, 2, \dots, k-1\} \cup \{i+1, \dots, N\} \quad (\text{A.17})$$

$$C_{rk}(\theta_1^*, \dots, \theta_M^*) < 1, \quad \forall r \in \{1, 2, \dots, k-1\} \cup \{i+1, \dots, N\}$$

We can choose $(\theta_1^*, \dots, \theta_M^*)$ to be rational and $T = (t_1, t_2, \dots, t_M)$ such that $\frac{t_j}{\sum t_j} = \theta_j^*, \forall j$. Define H_r to be the compound lottery which yields X_r with probability 1. Substituting (A.17) in (A.9) we see that

$$\lim_{z \rightarrow \infty} F(zt_1, \dots, zt_M) = H_k$$

$$\lim_{z \rightarrow \infty} G(zt_1, \dots, zt_M) = \sum_{r=k+1}^i a_r H_r,$$

where $a_r \geq 0, \sum_{r=k+1}^i a_r = 1$. Thus for large enough z , $E[F(zt_1, \dots, zt_M)]$ does not dominate $E[G(zt_1, \dots, zt_M)]$ by $FOSD_1$ and hence $F \not\geq_1 G$.

A similar proof establishes that (2.5) is necessary for $F \succeq_1 G$.

Next, suppose that $b > a$. Choose an increasing sequence of observations $T^\ell = (t_1^\ell, \dots, t_M^\ell)$ such that $(\frac{t_1^\ell}{\sum t_j^\ell}, \frac{t_2^\ell}{\sum t_j^\ell}, \dots, \frac{t_M^\ell}{\sum t_j^\ell})$ converges to $(\theta_1^{i,i+1}, \dots, \theta_M^{i,i+1})$, where $a \leq i < b$ and $(\theta_1^{i,i+1}, \dots, \theta_M^{i,i+1})$ is as defined in Lemma A.3. Thus Lemma A.3 and (A.9) imply that

$$\lim_{\ell \rightarrow \infty} \alpha_k(T^\ell) = \begin{cases} 0, & \text{if } k \notin \{i, i+1\} \\ \frac{\alpha_k}{\alpha_i + \alpha_{i+1}}, & \text{if } k \in \{i, i+1\} \end{cases}$$

$$\lim_{\ell \rightarrow \infty} \beta_k(T^\ell) = \begin{cases} 0, & \text{if } k \notin \{i, i+1\} \\ \frac{\beta_k}{\beta_i + \beta_{i+1}}, & \text{if } k \in \{i, i+1\} \end{cases}$$

Also, since $\alpha_i, \alpha_{i+1}, \beta_i, \beta_{i+1} > 0$, $\lim_{\ell \rightarrow \infty} F(T^\ell)$ and $\lim_{\ell \rightarrow \infty} G(T^\ell)$ are well-defined compound lotteries. Therefore

$$E[\lim_{\ell \rightarrow \infty} F(T^\ell)] \text{ } FOSD_1 \text{ } E[\lim_{\ell \rightarrow \infty} G(T^\ell)]$$

iff

$$\frac{\alpha_i}{\alpha_i + \alpha_{i+1}} \sum_{j=1}^k P_{ij} + \frac{\alpha_{i+1}}{\alpha_i + \alpha_{i+1}} \sum_{j=1}^k P_{i+1,j} \geq \frac{\beta_i}{\beta_i + \beta_{i+1}} \sum_{j=1}^k P_{ij} + \frac{\beta_{i+1}}{\beta_i + \beta_{i+1}} \sum_{j=1}^k P_{i+1,j}, \quad \forall k < M$$

iff

$$\frac{\alpha_i}{\alpha_i + \alpha_{i+1}} \leq \frac{\beta_i}{\beta_i + \beta_{i+1}}$$

1, 2, ..., N. The rest of the proof is similar to that of the second part of Lemma A.2.

■

The following lemma, which is a special case of a result in Kamae, Krengel, and O'Brien (1977), is required to prove Lemma 3.

Lemma A.5 (Kamae, Krengel, and O'Brien (1977)):

Let $F, G \in L_c$. F $FOSD_2$ G if and only if F and G can be written as $F = (X_1, q_1; X_2, q_2; \dots; X_L, q_L)$, $G = (Y_1, q_1; Y_2, q_2; \dots; Y_L, q_L)$ and X_i $FOSD_2$ Y_i , $\forall i = 1, 2, \dots, L$.

Proof of Lemma 3: The proof is a direct consequence of Lemma A.5 and is omitted. ■

Proof of Lemma 4: (i) Let $V_\ell : L_s \rightarrow R$ be defined as

$$V_\ell(X_i) = \sum_{j=\ell}^M P_{ij}, \quad \forall X_i \in L_s, \quad \ell = 2, 3, \dots, M.$$

Clearly V_ℓ is increasing with respect to $FOSD_1$. Therefore if F $FOSD_2$ G then

$$\begin{aligned} & \sum_{i=1}^N \alpha_i \sum_{j=\ell}^M P_{ij} \geq \sum_{i=1}^N \beta_i \sum_{j=\ell}^M P_{ij}, \quad \ell = 2, 3, \dots, M \\ \Rightarrow & \sum_{j=\ell}^M \sum_{i=1}^N \alpha_i P_{ij} \geq \sum_{j=\ell}^M \sum_{i=1}^N \beta_i P_{ij}, \quad \ell = 2, 3, \dots, M \\ \Rightarrow & \sum_{j=1}^{\ell-1} \sum_{i=1}^N \alpha_i P_{ij} \leq \sum_{j=1}^{\ell-1} \sum_{i=1}^N \beta_i P_{ij}, \quad \ell = 2, 3, \dots, M \\ \Rightarrow & E[F] \text{ } FOSD_1 \text{ } E[G]. \end{aligned}$$

(ii) Follows from (i) and Definitions 2 and 7. ■

Proof of Lemma 5: From Lemma A.1 we know that (2.4), (2.5) and (2.6) imply (A.1). Since X_N $FOSD_1$ $X_{N-1} \dots FOSD_1$ X_1 , Lemma 3 implies that F $FOSD_2$ G . Thus it is sufficient to show that after taking one sample the updated distributions satisfy (2.4), (2.5) and (2.6), and the lemma follows by repeated application. The rest of the proof is identical to the corresponding part of the proof of Lemma 1. ■

Proof of Theorem 3: By Lemma 4, $F \succeq_2 G \Rightarrow F \succeq_1 G$. By Theorem 2 and Lemma 5, $F \succeq_1 G \Rightarrow (2.4), (2.5) \text{ and } (2.6) \Rightarrow F \succeq_2 G$. ■

Proof of Theorem 4: In view of Lemma 4, we only need to show that (4.1)–(4.4) imply $F \succeq_2 G$. First, we establish that these conditions imply F $FOSD_2$ G . Let $|I_i|$ denote the cardinality of I_i . When $|I_i| = 1$, $\forall i$, (4.1)–(4.4) imply that we

Proof of Theorem 5: Suppose that $\mathcal{X} = \{X_1, X_2, \dots, X_N\}$ admits a D-partition. Clearly there exist distinct F and G which satisfy (4.1)–(4.4). It is then straightforward to check that F and G satisfy (4.5) and (4.6).

Suppose instead that there does not exist a D-partition. By Lemma 6, $d_\infty(\{X_1\}) = \mathcal{X}$. Thus there exists a finite sequence $i_1 = 1, i_2, i_3, \dots, i_L$ such that $\mathcal{X} = \{X_{i_1}, X_{i_2}, \dots, X_{i_L}\}$ and $X_{i_{s+1}} \in d_1(\{X_{i_s}\})$, $s = 1, 2, \dots, L - 1$. Suppose $F = (\alpha_1, \alpha_2, \dots, \alpha_N)$, $G = (\beta_1, \beta_2, \dots, \beta_N)$ satisfy (4.6). That is,

$$\alpha_i \beta_{i_{s+1}} - \alpha_{i_{s+1}} \beta_i = 0, \quad \forall s = 1, 2, \dots, L - 1. \quad (\text{A.18})$$

As established later if $\alpha_1 \beta_2 - \alpha_2 \beta_1 = 0$ and $\alpha_2 \beta_3 - \alpha_3 \beta_2 = 0$, then $\alpha_1 \beta_3 - \alpha_3 \beta_1 = 0$. Thus (A.18) is equivalent to

$$\alpha_i \beta_k - \alpha_k \beta_i = 0, \quad \forall i, k. \quad (\text{A.19})$$

Suppose there exists ℓ such that $\alpha_\ell > 0, \beta_\ell = 0$. Let k be such that $\beta_k > 0$. But then (A.19) is not satisfied for ℓ, k . Thus for all i , either $\alpha_i \beta_i > 0$ or $\alpha_i = \beta_i = 0$. Hence without loss of generality we may assume that $\alpha_i \beta_i > 0$, and (A.19) implies that for all i, k , $\frac{\alpha_i}{\alpha_k} = \frac{\beta_i}{\beta_k} = 0$, which implies $\alpha_i = \beta_i, \forall i$. Thus if there does not exist a D-partition, two distinct F and G cannot satisfy (4.6).

To complete the proof suppose that $\alpha_1 \beta_2 = \alpha_2 \beta_1$ and $\alpha_2 \beta_3 = \alpha_3 \beta_2$. If $\alpha_1 \beta_2 \neq 0$ then $\frac{\alpha_2 \beta_3}{\alpha_2 \beta_1} = \frac{\alpha_3 \beta_2}{\alpha_1 \beta_2}$ and thus $\alpha_1 \beta_3 = \alpha_3 \beta_1$. A similar proof holds if $\alpha_2 \beta_3 \neq 0$. If $\alpha_1 \beta_2 = \alpha_2 \beta_1 = \alpha_2 \beta_3 = \alpha_3 \beta_2 = 0$ then there are three possible cases: (a) $\alpha_1 = \alpha_2 = \alpha_3 = 0$, (b) $\beta_1 = \beta_2 = \beta_3 = 0$, (c) $\alpha_i = \beta_i$, for some $i = 1, 2, 3$. In (a) and (b) $\alpha_1 \beta_3 = \alpha_3 \beta_1$. In (c) we may ignore X_i . ■

The proof of Theorem 6 requires Lemmas A.6 and A.7 below. For each of these lemmas, it is assumed that the hypothesis of Theorem 6 holds.

Lemma A.6: If $F \succeq_2 G$ then

- (i) $\sum_{i=1}^K \alpha_i(T) = \sum_{i=1}^K \beta_i(T), \forall T;$
- (ii) $\sum_{i=1}^K \sum_{k=K+1}^N (\alpha_i \beta_k - \alpha_k \beta_i) \prod_{j=1}^M (P_{ij} P_{kj})^{t_j} = 0.$

Proof of Lemma A.6: $F \succeq_2 G \Rightarrow F(T) \text{ FOSD}_2 G(T), \forall T$. In particular, $F \text{ FOSD}_2 G$. Thus, by Lemma A.5, F and G can be written as $F = (X_{i_1}, q_1; X_{i_2}, q_2; \dots; X_{i_L}, q_L)$, $G = (X_{k_1}, q_1; X_{k_2}, q_2; \dots; X_{k_L}, q_L)$, with $X_i, \text{ FOSD}_1 X_k,$

$\forall s = 1, 2, \dots, L$. By the definition of K it follows that $\sum_{i=1}^K \alpha_i = \sum_{i_s \leq K} q_s = \sum_{k_s \leq K} q_s = \sum_{i=1}^K \beta_i$. Similarly, $F(T) \text{ FOSD}_2 G(T)$ implies that $\sum_{i=1}^K \alpha_i(T) = \sum_{i=1}^K \beta_i(T), \forall T,$

Clearly, $a_I(zT^*) = a_I^*$. Now consider the distinct observation vectors $T^*, 2T^*, \dots, K(N-K)T^*$. For each observation vector we have to satisfy the equation specified in Lemma A.6(ii). We therefore obtain a system of $K(N-K)$ homogeneous equations in the "variables" $[\alpha_i\beta_k - \alpha_k\beta_i], i = 1, 2, \dots, K, k = K+1, \dots, N$. The coefficient matrix is given by

$$\begin{bmatrix} a_1 & a_2 & \dots & a_{K(N-K)} \\ (a_1)^2 & (a_2)^2 & \dots & (a_{K(N-K)})^2 \\ \vdots & \vdots & \ddots & \vdots \\ (a_1)^{K(N-K)} & (a_2)^{K(N-K)} & \dots & (a_{K(N-K)})^{K(N-K)} \end{bmatrix}$$

This is the van der Monde matrix and since $a_1, a_2, \dots, a_{K(N-K)}$ are distinct it has full rank equal to $K(N-K)$ (see Graybill 1983). It follows that the only solution to the system of homogeneous equations is the trivial solution. Hence

$$\begin{aligned} \alpha_i\beta_k - \alpha_k\beta_i &= 0, & \forall i = 1, 2, \dots, K, k = K+1, \dots, N \\ \Rightarrow \alpha_i &= \beta_i, & \forall i = 1, 2, \dots, N. \end{aligned}$$

Thus F and G cannot be distinct. ■

Proof of Lemma 7: Let $y_k \in \{x_1, x_2, \dots, x_M\}$, $k = 1, 2, \dots, L$ denote the k th price sample observation. We need to show that

$$E[\min(y_1, y_2, \dots, y_L)|G] \leq E[\min(y_1, y_2, \dots, y_L)|F] \quad (\text{A.20})$$

We will prove a stronger result: If X_1, X_2, \dots, X_N are of type 1, and $F \succeq_1 G$ then

$$\text{Prob}\{\min(y_1, y_2, \dots, y_L) \leq x_j|G\} \geq \text{Prob}\{\min(y_1, y_2, \dots, y_L) \leq x_j|F\}, \quad \forall j < M \quad (\text{A.21})$$

Since $F \succeq_1 G$ implies $E[F] \text{ FOSD}_1 E[G]$, (A.21) is true for $L = 1$. That is,

$$\text{Prob}\{y_1 \leq x_j|G\} \geq \text{Prob}\{y_1 \leq x_j|F\}, \quad \forall j \leq M \quad (\text{A.22})$$

Suppose that (A.21) is true for $L = k$. It is enough to show that (A.21) is true for $L = k+1$. Let $F(x_\ell) [G(x_\ell)]$ denote the posterior distribution of $F [G]$ if $F [G]$ is selected and the first observation is x_ℓ . Clearly, $F(x_\ell) \succeq_1 G(x_\ell)$, $\forall \ell \leq M$. Thus, since (A.21) is true for $L = k$, we have

$$\begin{aligned} \text{Prob}\{\min(y_2, y_3, \dots, y_{k+1}) \leq x_j|G(x_\ell)\} &\geq \text{Prob}\{\min(y_2, y_3, \dots, y_{k+1}) \leq x_j|F(x_\ell)\}, \\ &\forall j \leq M, \forall \ell \leq M \end{aligned} \quad (\text{A.23})$$

Since X_1, X_2, \dots, X_N are of type 1, both F and G are increasing. Therefore $G(x_{\ell+1}) \succeq_1 G(x_\ell)$, and the fact that (A.21) is true for $L = k$ implies that

$$\begin{aligned} \text{Prob}\{\min(y_2, y_3, \dots, y_{k+1}) \leq x_j|G(x_\ell)\} &\geq \text{Prob}\{\min(y_2, y_3, \dots, y_{k+1}) \leq x_j|G(x_{\ell+1})\}, \\ &\forall j \leq M, \forall \ell < M \end{aligned} \quad (\text{A.24})$$

$$\frac{\alpha_2 \hat{C}_3(T) + \alpha_4 \hat{C}_4(T)}{\alpha_2 \hat{C}_3(T) + \alpha_3 \hat{C}_3(T) + \alpha_4 \hat{C}_4(T)} \geq \frac{\beta_3 \hat{C}_3(T)}{\beta_1 \hat{C}_1(T) + \beta_2 \hat{C}_2(T) + \beta_3 \hat{C}_3(T)}$$

which after substituting α_i, β_i becomes

$$15\hat{C}_1(T)\hat{C}_3(T) + 120\hat{C}_1(T)\hat{C}_4(T) + 16\hat{C}_2(T)\hat{C}_4(T) \geq \hat{C}_2(T)\hat{C}_3(T) \quad (\text{A.27})$$

Dividing across by $\hat{C}_2(T), \hat{C}_3(T)$ we have

$$15\hat{C}_{12}(T) + 120\hat{C}_{12}(T)\hat{C}_{43}(T) + 16\hat{C}_{43}(T) \geq 1$$

Since, either $C_{12}(\theta_1, \theta_2, \theta_3) \geq 1$ or $C_{43}(\theta_1, \theta_2, \theta_3) \geq 1$, for all $(\theta_1, \theta_2, \theta_3)$, from (A.26) we know that either $\hat{C}_{12}(T) \geq 1$ or $\hat{C}_{43}(T) \geq 1$, for all T . Thus (A.27) is satisfied and $F(T)$ $FOSD_2$ $G(T)$, for all T .

Example A.2: In this example we show that in general \succeq_1 and \succeq_2 are not equivalent. Thus the converse of Lemma 4(ii) is false. Consider the following simple lotteries over three final outcomes: $X_1 = (0.3, 0.4, 0.3)$, $X_2 = (0.5, 0.05, 0.45)$, $X_3 = (0.1, 0.5, 0.4)$ and $X_4 = (0.3, 0.1, 0.6)$. These are plotted in $P_{i1} - P_{i3}$ space in Figure 6. Let $F = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0, 0, 0.6, 0.4)$ and $G = (\beta_1, \beta_2, \beta_3, \beta_4) = (0.5, 0.5, 0, 0)$. Note that X_4 $FOSD_1$ X_2 , and X_2 cannot be compared with X_1 and X_3 by $FOSD_1$. Thus, since $\beta_2 + \beta_4 > \alpha_2 + \alpha_4$, Lemma A.5 implies that F does not dominate G by $FOSD_2$. To show that $F(T)$ $FOSD_1$ $G(T)$ for all $T = (t_1, t_2, t_3)$ we need to verify that

$$\alpha_3(T)P_{31} + \alpha_4(T)P_{41} \leq \beta_1(T)P_{11} + \beta_2(T)P_{21} \quad (\text{A.28})$$

$$\alpha_3(T)P_{33} + \alpha_4(T)P_{43} \leq \beta_1(T)P_{13} + \beta_2(T)P_{23} \quad (\text{A.29})$$

Since $\min\{P_{11}, P_{21}\} = \max\{P_{31}, P_{41}\}$ and $\alpha_3(T) + \alpha_4(T) = \beta_1(T) + \beta_2(T) = 1$, we know that (A.28) is true. Substituting $P_{i3}, i = 1, \dots, 4$ in (A.29) we have

$$4\alpha_3(T) - 3\beta_1(T) \leq 3 \quad (\text{A.30})$$

After simplification (A.30) yields

$$2\hat{C}_{12}(T) + 4\hat{C}_{12}(T)\hat{C}_{43}(T) + 2\hat{C}_{43}(T) \geq 1 \quad (\text{A.31})$$

where $\hat{C}_{i,k}(T)$ is as in (A.26). Direct computation shows that the $C_{12}(\theta_1, \theta_2, \theta_3) = 1$ line is $\theta_3 = 0.84 - 1.04\theta_1$, and the $C_{43}(\theta_1, \theta_2, \theta_3) = 1$ line is $\theta_3 = 0.8 - 1.34\theta_1$. Plotting this on Figure 6 we see that for all $(\theta_1, \theta_2, \theta_3)$, either $C_{12}(\theta_1, \theta_2, \theta_3) > 1$ or $C_{43}(\theta_1, \theta_2, \theta_3) > 1$. Therefore for all T either $\hat{C}_{12}(T) > 1$ or $\hat{C}_{43}(T) > 1$, and (A.31) holds. Thus $F \succeq_1 G$.

Example A.3: In Theorem 6 we require that if $\forall i, i' \leq K, \forall k, k' > K, P_{ij}P_{kj} = P_{i'j}P_{k'j}, \forall j$, then $i = i', k = k'$. This condition is generically satisfied by X_1, X_2, \dots, X_N . We show that when this condition is not satisfied, we can have $F \neq G, F \succeq_2 G$.

REFERENCES

- BERRY, D.A., AND B. FRISTEDT (1985), *Bandit Problems: Sequential Allocation of Experiments*, Chapman and Hall: London.
- BIKHCHANDANI, S., AND S. SHARMA (1989), "Optimal Search with Learning," working paper, UCLA.
- BIKHCHANDANI, S., U. SEGAL, AND S. SHARMA (1989), "Stochastic Dominance under Learning," working paper, UCLA.
- BLACKWELL, D., (1953), "Equivalent Comparison of Experiments," *Annals of Mathematics and Statistics*, 24, 265-272.
- BOHNENBLUST, H.F., L.S. SHAPLEY, AND S. SHERMAN (1949), "Reconnaissance in Game Theory," working paper, RM-208, RAND Corporation.
- DE FINETTI, B. (1977), "Probabilities of Probabilities: A Real Problem or a Misunderstanding?" in *New Developments in the Applications of Bayesian Methods*, edited by A. Aykac and C. Brumet, North Holland: Amsterdam.
- DEKEL, E., Z. SAFRA, AND U. SEGAL (1989), "Existence of Nash Equilibrium with Non-Expected Utility Preferences", forthcoming, *Journal of Economic Theory*.
- EASLEY, D., AND N.M. KIEFER (1988), "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045-64.
- FERGUSON, T.S., (1967), *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press: New York.
- GRAYBILL, F. (1983), *Introduction to Matrices with Applications in Statistics*, 2nd edition, Wadsworth: Belmont, CA.
- KAHNEMAN, D., AND A. TVERSKY (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263-291.
- KAMAE, T., U. KRENGEL, AND G.L. O'BRIEN (1977), "Stochastic Inequalities on Partially Ordered Spaces," *Annals of Probability*, 5, 899-912.
- KELLER, L.R. (1985), "Testing of the 'Reduction of Compound Alternatives' Principle," *Omega*, 13, 349-358.
- KREPS, D. (1988), *Notes on the Theory of Choice*, Westview Press: Boulder, CO.
- LIPPMAN, S.A., AND J.J. MCCALL (1981), "The Economics of Uncertainty: Selected Topics and Probabilistic Methods," in *Handbook of Mathematical Economics*, vol. I, edited by K.J. Arrow and M.D. Intrilligator, North Holland: Amsterdam.

Footnotes

¹ The question of ordering distributions on distributions has been addressed before by Bohnenblust, Shapley and Sherman (1949), and Blackwell (1953). Bohnenblust, Shapley and Sherman (1949) obtain a partial ordering on information systems, that is on distributions on distributions, based on their value to the decision-maker. Blackwell (1953) shows that this partial ordering is the same as the one obtained from the statistical notion of sufficiency. This notion compares two information systems based on the criterion of second-order stochastic dominance of the expected posterior distributions after one observation.

² Trade-Off Talking Rational Economic Person.

³ In the Bayesian statistics literature, a distribution over distributions is sometimes called a *random distribution*.

⁴ We assume $P_{ij} > 0$ for simplicity. Our results can be extended to allow $P_{ij} = 0$ for some i and j .

⁵ Such diagrams were reintroduced in the literature by Mark Machina. He attributes these diagrams to Jacob Marschak (see Machina (1987)).

⁶ A stronger definition of Bayes' first-order stochastic dominance would be to require that stochastic dominance be maintained after any observation T from F and T' from G . Thus corresponding to \succeq_1 and \succeq_2 we may define

$$\begin{aligned} F \succeq_1^* G & \text{ iff } E[F(T)] \text{ FOSD}_1 E[G(T')], \quad \forall T, T' \\ F \succeq_2^* G & \text{ iff } F(T) \text{ FOSD}_2 G(T'), \quad \forall T, T' \end{aligned}$$

Let $F = (X_1, p_1; X_2, p_2; \dots; X_N, p_N)$, $p_i > 0$, and $G = (Y_1, q_1; Y_2, q_2; \dots; Y_L, q_L)$, $q_\ell > 0$. It is easily verified that $F \succeq_1^* G$ if and only if $F \succeq_2^* G$ if and only if $X_i \text{ FOSD}_1 Y_\ell, \forall i, \ell$. First note that for all T the support of $F(T)$ is (X_1, X_2, \dots, X_N) , and for all T' the support of $G(T')$ is (Y_1, Y_2, \dots, Y_L) . Therefore a sufficient condition for $F \succeq_1^* G$ and $F \succeq_2^* G$ is that $X_i \text{ FOSD}_1 Y_\ell, \forall i, \ell$. From (2.1) it is clear that for any i and ℓ one can find T and T' such that $F(T)$ places most of its mass on X_i , and $G(T')$ on Y_ℓ . Thus a necessary condition for $F \succeq_1^* G$ is that $X_i \text{ FOSD}_1 Y_\ell, \forall i, \ell$. This condition is also necessary for $F \succeq_2^* G$ since Lemma 4(i) implies that if $F \succeq_2^* G$ then $F \succeq_1^* G$.

⁷ The 'D' in D-partition stands for dominance.

⁸ That is, if each X_i is an independent draw from a uniform probability distribution on the simplex in R^M , then with probability one X_1, X_2, \dots, X_N satisfy condition (ii).

⁹ Related assumptions are *affiliation* (Milgrom and Weber (1981)) and *conditional stochastic dominance* (see Riley (1988)).

FIGURE 1

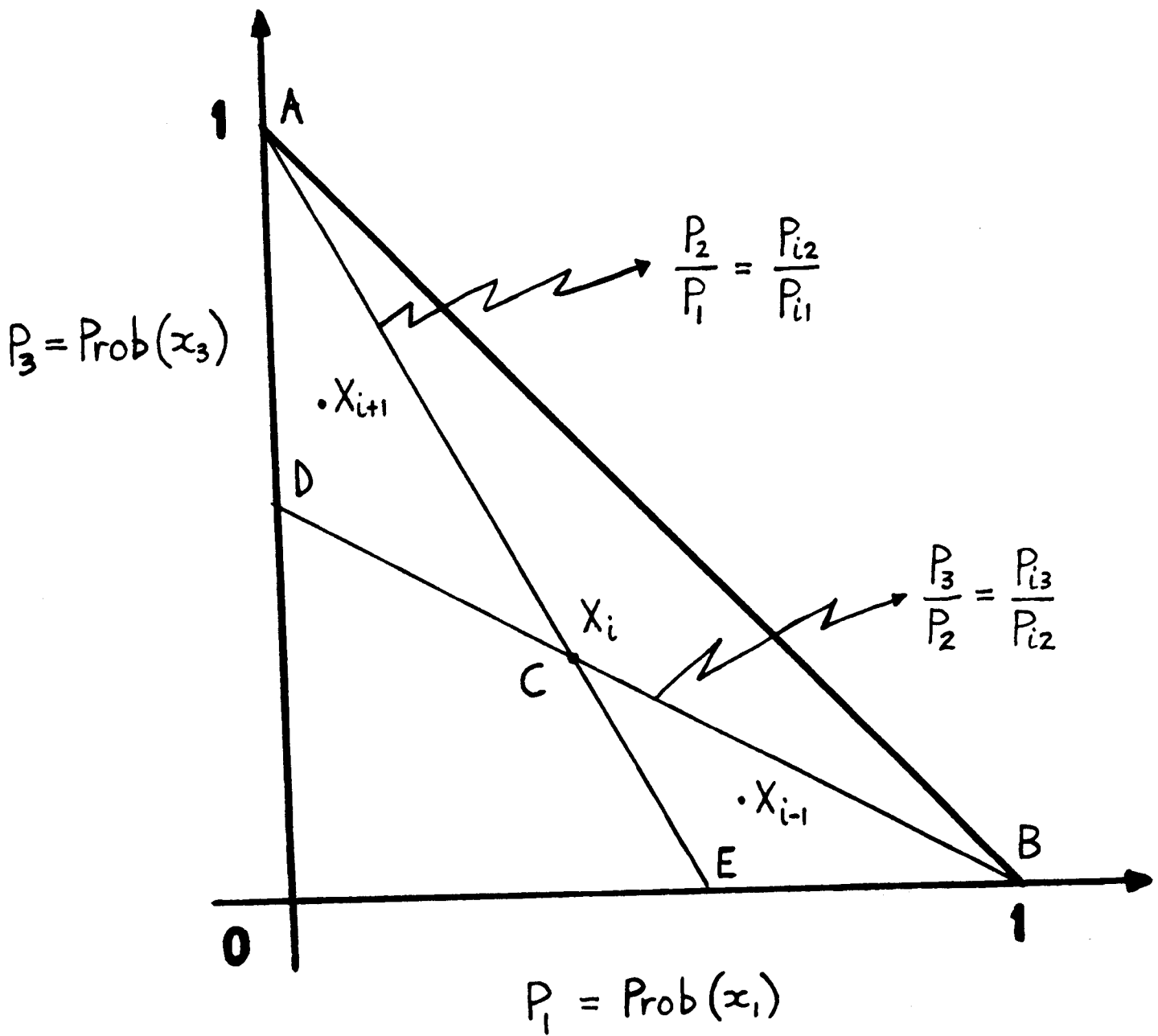


FIGURE 3

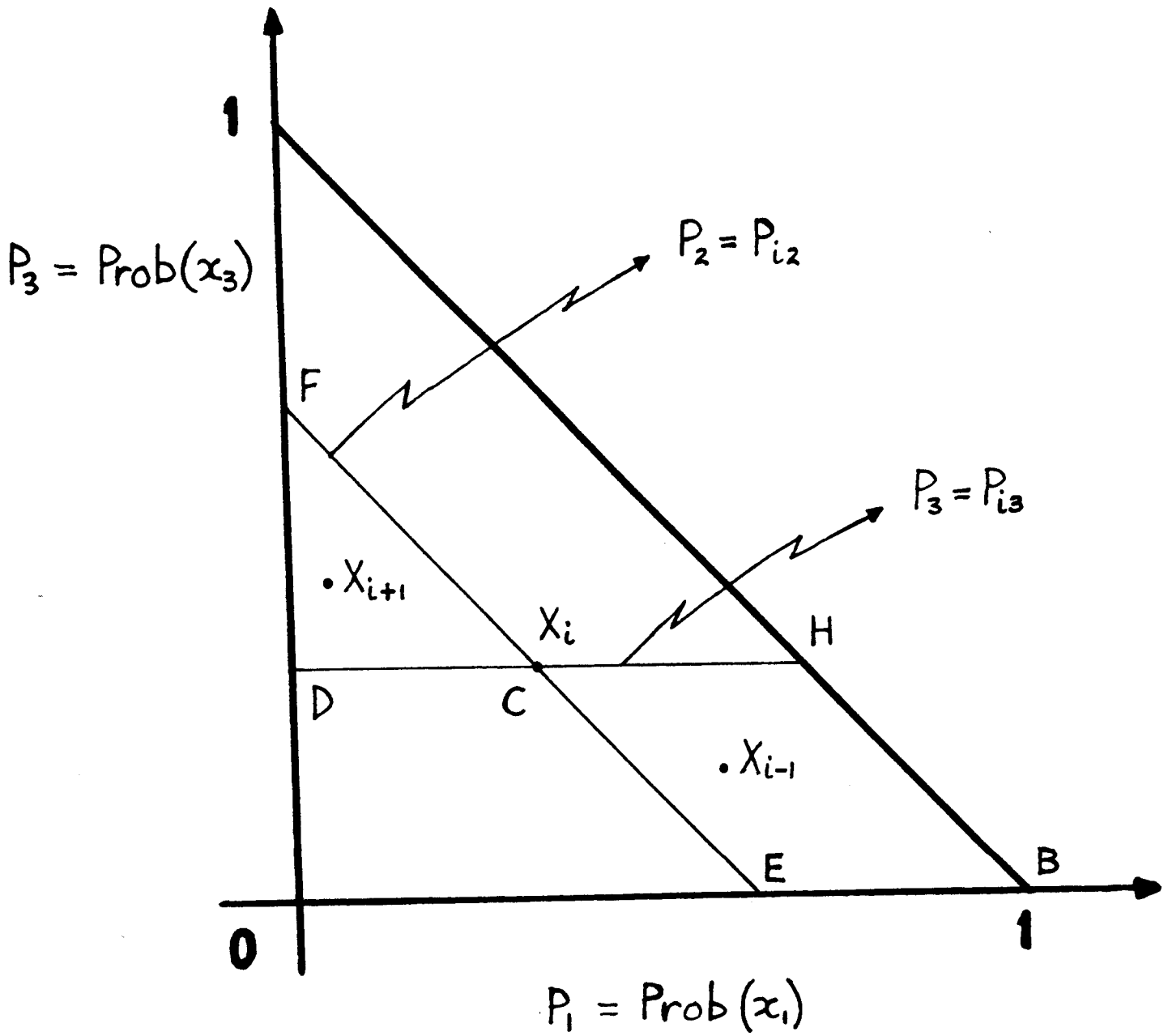


FIGURE 5

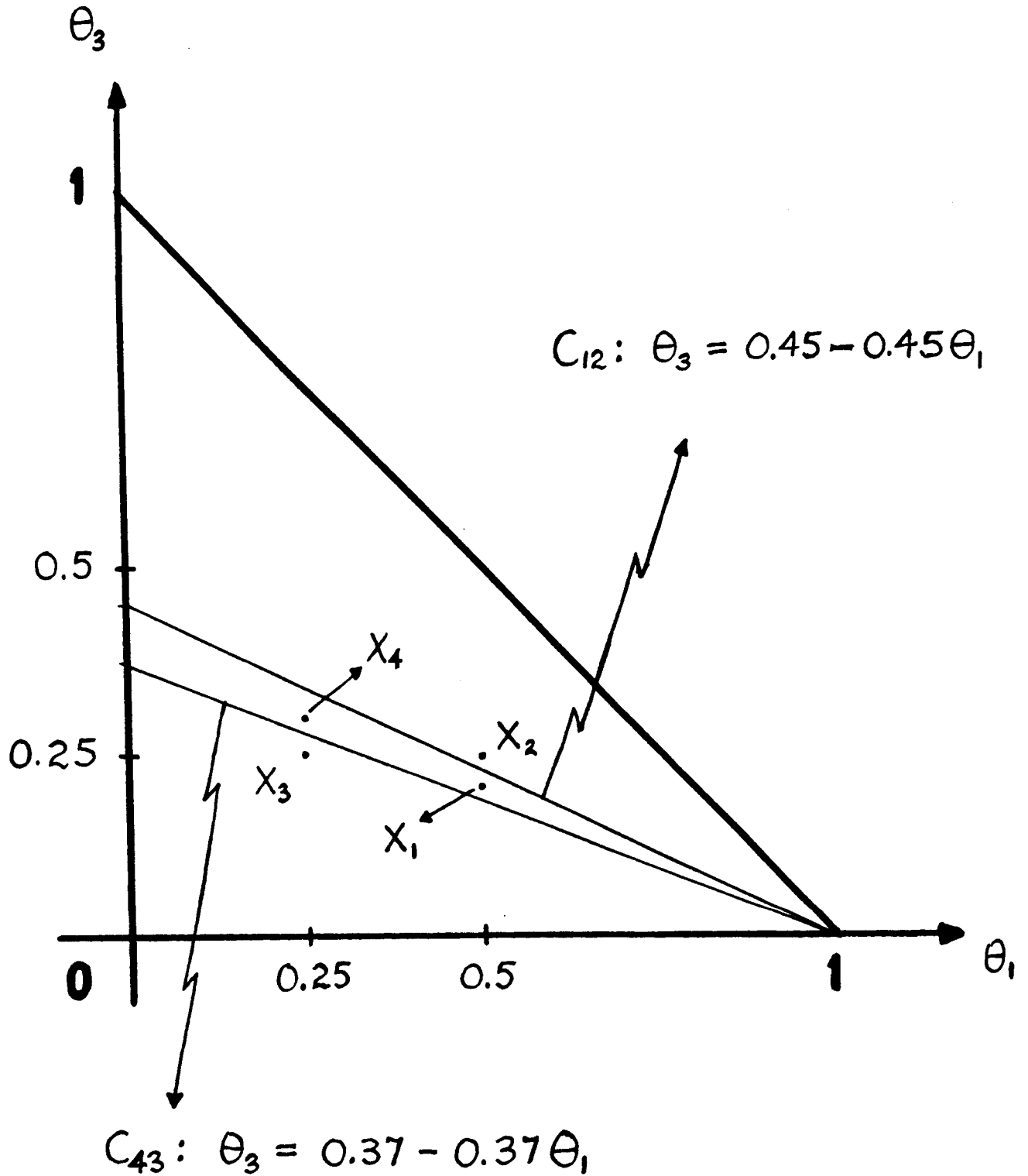


FIGURE 7

