working paper
department
of economics

CHARACTERIZING PROPERTIES OF STOCHASTIC OBJECTIVE FUNCTIONS

Susan Athey

96-1 Oct. 1995

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

CHARACTERIZING PROPERTIES OF STOCHASTIC OBJECTIVE FUNCTIONS

Susan Athey

96-1 **Oct. 1995**

1 INTRODUCTION

This paper studies optimization problems where the objective function can be written in the form $V(\theta) \equiv \int_s \pi(s)dF(s;\theta)$, where π is a payoff function, F is a probability distribution, and θ and s are real vectors. For example, the payoff function π might represent an agent's utility or a firm's profits, the vector s might represent features of the current state of the world, and the elements of θ might represent an agent's investments, effort decisions, other agent's choices, or the nature of the exogenous uncertainty in the agent's environment.

The economic problem under consideration often determines some properties of the payoff function; for example, a utility function might be assumed to be nondecreasing and concave, while a multivariate profit function might have sign restrictions on cross-partial derivatives. These assumptions then determine a set Π of admissible payoff functions. We might then wish to answer questions such as: Is a set of investments θ is worthwhile? Are there decreasing returns to those investments? Does one investment increase the returns to another investment? To answer those questions, we need to know whether $V(\theta)$ satisfies the appropriate properties, i.e., nondecreasing, concave, or supermodular.[1]

The goal of this paper is to develop methods such that, given a particular property P (such as nondecreasing), a set of admissible payoff functions Π, and a parameterized probability distribution F, we can determine whether the following statement is true:

$$\int_s \pi(s)dF(s;\theta) \text{ satisfies property } P \text{ in } \theta \text{ for all } \pi \text{ in } \Pi. \tag{1.1}$$

That is, given P and Π, we wish to describe the set of probability distributions which satisfy (1.1).

This paper restricts attention to the case where the properties P and sets Π are "closed convex cones." We define a property P to be a "Closed Convex Cone" (CCC) property if the set of functions which satisfy P is closed (under an appropriate topology), positive combinations of functions in the set are also in the set, and constant functions are in the set. Important examples of CCC properties include nondecreasing, concave, supermodular, any property which is defined by placing a sign restriction on a partial derivative, and combinations of these properties. Thus, if the payoff function represents a consumer's utility, Π could be the set of univariate, nondecreasing, concave payoff functions; if the

[1]Intuitively, a function is supermodular if, given any two arguments of the function, increasing one increases the returns to increasing the other. When the function is differentiable, this amounts to positive cross-partial derivatives between every pair of arguments. Supermodularity is important in the context of comparative statics (see Topkis (1978), Milgrom and Roberts (1990, 1994), Milgrom and Shannon (1994)).

payoff function represents a firm's profits as a function of two complementary quality innovations, Π could be the set of bivariate, nondecreasing, supermodular payoff functions.

For sets Π and for properties P which are CCC, this paper develops a characterization of which probability distributions satisfy (1.1). Checking (1.1) directly is difficult because, in general, Π might be a very large set. Thus we ask, When is it possible to find a smaller set of payoff functions, denoted Γ, so that for any probability distribution F, statement (1.2) below will be true if and only if (1.1) is true?

$$\int_s \pi(s)dF(s;\theta) \text{ satisfies property } P \text{ in } \theta \text{ for all } \pi \text{ in } \Gamma \tag{1.2}$$

We can think of Γ as a "test set" for Π: ideally, Γ is a set of payoff functions which is smaller and easier to check than Π, but it can be used to test whether $\int_s \pi(s)dF(s;\theta)$ satisfies property P in θ, given only the information that π is in the larger set Π.

Using these ideas, we can restate the goal of the paper: we want a theory which helps us determine the best "test set" for a given Π. We proceed in two steps. First, we examine the case where the property P is "nondecreasing"; second, we study other CCC properties. The first case has been the focus of the literature on stochastic dominance, where different authors have studied different sets Π.[2] This paper unifies and extends the existing literature on stochastic dominance, providing an exact characterization of the mathematical structure underlying all stochastic dominance theorems. We further provide an algorithm for generating new stochastic dominance theorems which relaxes the ad hoc differentiability and continuity assumptions which are common in this literature.[3] In the second part of the paper we show that the methods from stochastic dominance can be applied to characterize when $V(\theta)$ satisfies other properties P, including concavity and supermodularity.

We now provide an overview of our results. In the first part of this paper, we show that for the property nondecreasing, if Π is a closed convex cone and contains constant functions, then the best Γ is the set of "extreme points" of Π. Just as a basis generates a linear space via linear combinations, so a set of extreme points generates a closed convex cone via *positive* linear combinations and limits. Thus, we show that if we know only that a payoff function π lies in the closed convex cone Π, and we want to know if $V(\theta)$ is nondecreasing, it is equivalent to check that $V(\theta)$ is nondecreasing on a smaller

---

[2]In particular, the univariate stochastic dominance problem has been studied by Rothschild and Stiglitz (1970, 1971) and Hadar and Russell (1971); notable contributions to the multivariate problem include Levy and Paroush (1974), Atkinson and Bourguignon (1982), and Meyer (1990). Shaked and Shanthikumar (1994) provide a reference book on the subject of stochastic orders and their applications in economics, biology, and statistics.

[3]Brumelle and Vickson (1975) take a first step towards relaxing these assumptions and identifying the mathematical structure behind stochastic dominance; in contrast to their work, which provides only sufficient conditions for a stochastic dominance relationship, this paper provides an exact characterization of all stochastic dominance theorems.

set of payoff functions, the extreme points of Π. It will often be much easier to verify monotonicity of $V(\theta)$ for the set of extreme points of Π than for Π itself. In this paper, the procedure of using the extreme points of Π as a test set for Π will be referred to as the "closed convex cone" method of proving stochastic dominance theorems.

Examples from the existing stochastic dominance literature, which are special cases of this result, include First Order Stochastic Dominance (FOSD), where Π is the set of univariate, nondecreasing payoff functions, and Second Order Stochastic Dominance (SOSD), where Π is the set of univariate, concave payoff functions. In the case of FOSD, the set of extreme points is the set of one-step functions which are zero up to some constant, and one thereafter. These are pictured in Figure 1. In the case of SOSD, the set of extreme points is the set of "angle," or "min" functions, pictured in Figure 2, where each function takes the minimum of its argument and some constant.

The closed convex cone method for stochastic dominance has been recognized and explored in the context of particular sets of payoff functions (Topkis, 1968; Brumelle and Vickson, 1974; Gollier and Kimball, 1995).[4] However, this paper goes beyond the existing literature in two respects. First, by developing appropriate abstract definitions to describe stochastic dominance theorems, we are able to make general statements about the entire class of stochastic dominance theorems, including those which have not yet been considered in the economics literature. Second and more important is the fact that we prove a new result about this class of theorems: we prove formally that the "closed convex cone" approach to stochastic dominance is exactly the right one. By this we mean that (1.1) and (1.2) return the same answer for every probability distribution F if and only if the closed convex cone of Γ (union the constant functions, if these are not in Γ) is equal to Π; no other Γ's will always make (1.1) and (1.2) equivalent. In this case, clearly the smallest set which generates Π is the best set to check.

The second part of this paper shows that the "closed convex cone" method can also be applied to characterize other properties of $V(\theta)$. We ask two questions: First, for what properties P is the closed convex cone approach valid? And second, for what properties P is the closed convex cone approach exactly the right one, as in the case of stochastic dominance? We first show that if P is a CCC property, then the closed convex cone approach can always be used to characterize when $V(\theta)$ satisfies P. We then find a subset of CCC properties, which we call "Linear Difference Properties," for which we can show that the closed convex cone approach is exactly the right one for checking whether $V(\theta)$ satisfies P. Examples of Linear Difference Properties include monotonicity, supermodularity, concavity, and properties which place sign restrictions on partial derivatives. Combinations of these properties, however, are not in general Linear Difference Properties, although such combinations are

[4]Independently, Gollier and Kimball (1995) argue for what they call the "basis approach" to stochastic dominance.

CCC properties. Table I summarizes properties which are CCC properties and Linear Difference Properties.

TABLE I

LINEAR DIFFERENCE PROPERTIES AND CLOSED CONVEX CONE PROPERTIES

Property	Closed Convex Cone	Linear Difference Property
Nondecreasing	Yes	Yes
Supermodular	Yes	Yes
Concave/Convex (Multivariate)	Yes	Yes
Sign Restriction on a Partial Derivative	Yes	Yes
Constant	Yes	No
Nondecreasing and Concave	Yes	No
Arbitrary Combinations of CCC Properties	Yes	No
Arbitrary Combinations of LDP Properties	Yes	No

For many sets of payoff functions, Π, which are commonly studied in economics, the existing literature on stochastic dominance has implicitly identified the extreme points of those sets. In such cases, the problem of characterizing a Linear Difference Property (such as supermodularity) becomes quite straightforward: simply look to the the existing literature to find the appropriate test set, Γ, for the set of payoff functions Π under consideration. Then, using the results of this paper, we know that (1.2) characterizes the set of parameterized probability distributions for which the stochastic objective function will be supermodular for all payoff functions in Π.

We illustrate this technique by showing how the closed convex cone approach can be used to characterize the property supermodularity for several important classes of payoff functions. These results can in turn provide sufficient (and sometimes necessary) conditions for comparative statics conclusions. In particular, we examine applications in principal-agent theory, welfare economics, and the study of coordination problems in firms.

This paper proceeds as follows. In Section 2, we introduce a motivating example, the problem of a risk-averse agent's choice of effort. In this problem, we characterize the properties nondecreasing, concave, and supermodular for the agent's expected utility function. Our general analysis of stochastic dominance takes place in Section 3. We provide an exact characterization of stochastic dominance theorems, highlighting the important role played by linearity of the integral. We further extend our result to incorporate "conditional stochastic dominance." Section 4 develops characterizations of other properties of $\int_s \pi(s)dF(s;\theta)$ and provides applications of the property supermodularity. In Section 5, we analyze conditions under which functions of the form $\int_s \pi(x,s)dF(s;\theta)$ are supermodular or concave in (x,θ), showing how to apply stochastic monotonicity results to this problem as well. Section 6 concludes.

2 MOTIVATING EXAMPLE: A RISK-AVERSE AGENT'S CHOICE OF EFFORT

In this section, we present a motivating example, where we analyze how a risk-averse agent's choice of effort affects her expected payoffs, and how that choice of effort interacts with exogenous parameters which describe the probability distribution. Formally, we characterize the properties nondecreasing, concave, and supermodular for the agent's expected utility function. These results are specific examples -- examples where the payoff function is nondecreasing and concave -- of a stochastic dominance theorem, a "stochastic concavity theorem," and a "stochastic supermodularity theorem." These examples illustrate the parallel structure underlying the three classes of theorems.

Consider a risk-averse agent whose utility (π) depends on the output of a stochastic production technology, where the output is denoted s. Suppose that the agent's effort (e) affects the probability distribution of s. Further, consider a parameter t which represents exogenous changes in the stochastic production technology. For example, a change in t might represent a worker moving from one job to another. Then, we can write the agent's problem as follows:

$$\max_e \int_s \pi(s) \cdot dF(s;e,t) - c(e)$$

First observe that it is not trivial to verify that $V(e,t) = \int_s \pi(s) \cdot dF(s;e,t)$ is nondecreasing in effort: an increase in effort which is productive on average might not increase expected utility if the effort also increases the riskiness of the distribution. Second, note that if $V(e,t)$ - $c(e)$ is concave in effort, then the first order conditions (FOCs) characterize the optimum, a property which is useful in the analysis some economic problems, such as principal agent problems. Further, if $V(e,t)$ fails to be concave in e, then there exists some linear cost function $c(e) = a \cdot e$ such that the FOCs fail to characterize the optimum. Finally, observe that if $V(e,t)$ is supermodular, then the optimal choice of effort, $e^*(t)$, is nondecreasing in t. If $V(e,t)$ fails to be supermodular, then even if V is concave, there exists some linear cost function $c(e) = a \cdot e$ such that $e^*(t)$ fails to be nondecreasing in t (this follows from Milgrom and Shannon (1994); see Theorem A.1 in the Appendix). Thus, the requirement that $V(e,t)$ satisfies the property supermodular (or concave, respectively) is sufficient for the desired economic conclusion, and further it cannot be relaxed as long as we require that the conclusion holds for all linear cost functions.

Let us first identify conditions under which $V(e,t)$ is nondecreasing in e. The following well-known result is adapted from the Rothschild and Stiglitz (1970, 1971) work on stochastic dominance (where we suppress t in the notation):

Proposition 2.1 *The following two conditions are equivalent for all probability distributions* $F(\cdot;e)$:

(i) For all π *nondecreasing and concave,* $\int_s \pi(s) \cdot dF(s;e)$ *is nondecreasing in e.*

(ii) The following are satisfied:

(a) $\int_{-\infty}^{\infty} s \cdot dF(s;e)$ *is nondecreasing in e.*

(b) *For all a,* $-\int_{-\infty}^{a} F(s;e)$ *is nondecreasing in e.*

Intuitively, for a risk-averse agent who likes income, effort will increase expected utility if and only if effort increases the mean income (condition (ii)(a)) and reduces the "risk function" (condition (ii)(b)). This result is often used in the finance literature; it has been called Second Order Monotonic Stochastic Dominance (SOMSD).

In this paper, we will work with a restatement of Proposition 2.1, which emphasizes that conditions (i) and (ii) are actually symmetric conditions. The following proposition is equivalent to Proposition 2.1 (where Δ^1 be the space of probability distributions defined on $\Re$, and $\overline{\Re}$ indicates the extended real line, that is, $\Re \cup \{-\infty, \infty\}$):

Proposition 2.1′ *The following two conditions are equivalent for all* $F(\cdot;e) \in \Delta^1$:

(i) For all π *in the set* $\Pi^{SOM} \equiv \{\pi | \pi : \Re \to \Re, \text{ nondecreasing, concave}\}$, $\int_s \pi(s) \cdot dF(s;e)$ *is nondecreasing in e.*

(ii) For all γ *in the set* $\Gamma^{SOM} \equiv \{\gamma | \gamma(s) = \min(a,s),\ a \in \overline{\Re}\}$, $\int_s \gamma(s) \cdot dF(s;e)$ *is nondecreasing in e.*

Conditions (i) and (ii) of this Proposition simply restate conditions (i) and (ii) of Proposition 2.1, except that we have replaced $-\int_{-\infty}^{a} F(s;e)$ with $\int_{-\infty}^{\infty} \min(a,s) \cdot dF(s;e)$ in condition (ii). It is straightforward to verify (using integration by parts) that the former term is nondecreasing if and only if the latter term is. This condition is not usually associated with SOMSD (see Brumelle and Vickson (1975) for an exception). This way of writing the SOMSD theorem illustrates the mathematical structure underlying the stochastic dominance theorem. As written in Proposition 2.1′, the result says that instead of checking that $\int_s \pi(s) \cdot dF(s;e)$ is nondecreasing in θ_1 for all payoff functions in the relatively large set, Π^{SOM}, it is equivalent to check that $\int_s \gamma(s) \cdot dF(s;e)$ is nondecreasing in e for all payoff functions in the smaller set, Γ^{SOM}.

The two sets of payoff functions are pictured in Figure 2. The relationship between the two sets of payoff functions can be described as follows: Π^{SOM} is equal to the closed convex cone of the set which is the union of Γ^{SOM} and the functions $\gamma(s) = 1$ and $\gamma(s) = -1$. That is, by taking positive combinations and limits of sequences (or nets) of elements of the latter set, we can generate any function in Π^{SOM}, which is itself a closed convex cone. In Section 3, we will show that this "closed convex cone" relationship between Π^{SOM} and Γ^{SOM} holds for all stochastic dominance theorems.

Now, let us ask a different, but related, question: When is $\int_s \pi(s) \cdot dF(s;e)$ concave in effort? This problem was addressed by Jewitt (1988), who analyzes conditions under which the First Order Approach (FOA) to analyzing principal-agent problems is valid. The FOA requires that if the agent's FOCs are satisfied, then the agent's choice of effort must be optimal. Extending Jewitt's analysis, we can show that the sufficient conditions he derives are in fact necessary.[5] The following result is analogous to Proposition 2.1:

Proposition 2.2 *The following two conditions are equivalent for all* $F(\cdot;e) \in \Delta^1$*:*

(i) For all π *in the set* Π^{SOM}, $\int_s \pi(s) \cdot dF(s;e)$ *is concave in e.*

(ii) For all γ *in the set* Γ^{SOM}, $\int_s \gamma(s) \cdot dF(s;e)$ *is concave in e.*

Proof: This will be established as a simple corollary of Theorem 4.1 below, together with Proposition 2.1. *Q.E.D.*

---

[5]Jewitt first derives conditions under which the agent's utility function is increasing and concave in output, given the optimal contract; he then addresses the question of concavity of the expected utility function in effort. It is only the latter question which we study here.

Proposition 2.2 says that expected payoffs are concave in e for all payoff functions in Π^{SOM} if and only if expected payoffs are concave in e for all payoff functions in Γ^{SOM}. Again, this theorem is useful because condition (ii) is easier to check that condition (i): the set Γ^{SOM} is much smaller. Condition (ii) can be interpreted as requiring that there are decreasing returns to e in terms of increasing the mean and decreasing the "risk function." Note that the pair of sets of payoff functions, $(\Pi^{SOM}, \Gamma^{SOM})$, is the same in both propositions.

Now we turn to ask a final question: When is the optimal choice of effort monotone nondecreasing in t, which parameterizes the stochastic production technology? More precisely, how can t affect the probability distribution over output in such a way that monotonicity of the optimal effort in t is ensured, without any additional information about the cost of effort function or the agent's preferences? This question has not been answered in the existing literature; thus, the following proposition provides a new insight into the comparative statics problem. Note that this proposition is true irrespective of whether expected utility is monotone in effort.

Proposition 2.3 *The following three conditions are equivalent for all $F(\cdot;e,t) \in \Delta^1$:*

(MCS) For all cost functions c and all π in Π^{SOM}, $e^*(t) \equiv \arg\max_e \int_s \pi(s) \cdot dF(s;e,t) - c(t)$ is monotone nondecreasing in t.[6]

(i) For all $\pi \in \Pi^{SOM}$, $\int_s \pi(s) \cdot dF(s;e,t)$ is supermodular in (e,t).

(ii) For all $\gamma \in \Gamma^{SOM}$, $\int_s \gamma(s) \cdot dF(s;e,t)$ is supermodular in (e,t).

Proof: The equivalence of parts (MCS) and (i) follows directly from Milgrom and Shannon (1994), as stated in Theorem A.1 in the Appendix. The equivalence of (i) and (ii) will be established as a corollary of Theorem 4.1 below together with Proposition 2.1.

Q.E.D.

The formal definition of supermodularity (and the comparative statics theorem which relies upon it) can be found in the Appendix. Intuitively, $V(e,t)$ is supermodular if t increases the returns to effort. Proposition 2.3 provides necessary and sufficient conditions for monotone comparative statics in this problem. If (ii) is violated, then we can construct payoff functions and cost functions such that the choice of effort is not monotonic in t. Thus, we have identified the exact conditions which ensure monotone comparative statics. Condition (ii) requires that e and t are complementary in terms of increasing the mean of the probability distribution and in terms of reducing the risk. The intuition is straightforward: since a risk-averse, income-loving agent likes high expected returns and low risk (as

[6]In general, the optimal e may be a set. Then, this theorem requires that the set be nondecreasing in the Strong Set Order, as defined in the Appendix.

shown in SOMSD), variables which are complementary in increasing the mean and decreasing the risk are complementary in increasing expected utility of such an agent. Note that if either one of e and t does not affect the mean or the riskiness of the agent's income, the corresponding complementarity conditions are satisfied trivially.

Notice that Propositions 2.2 and 2.3 have a structure which is very similar to the existing stochastic dominance result, as stated in Proposition 2.1. In Section 3, we will build a framework for analyzing Proposition 2.1 and other stochastic dominance theorems. In Section 4, we will formalize the relationship between Propositions 2.1 through 2.3.

3 MONOTONICITY OF STOCHASTIC OBJECTIVE FUNCTIONS

The goal of this section is to provide a unified framework for analyzing stochastic dominance theorems. In Section 3.1, we introduce a framework which incorporates the existing stochastic dominance literature, arguing that each stochastic dominance theorem describes a relationship between two sets of payoff functions. In Section 3.2, we prove a result which characterizes a mathematical relationship between two sets of payoff functions which is equivalent to the relationship determined by the stochastic dominance theorem. Section 3.3 extends the result to the case of conditional stochastic dominance.

3.1 A Unified Framework for Stochastic Dominance

In this section, we introduce the framework which we will use to discuss stochastic dominance theorems as an abstract class of theorems, and to draw precise parallels between stochastic dominance theorems and other types of theorems.

Let us first consider another well-known example of a stochastic dominance theorem, First Order Stochastic Dominance (FOSD). This theorem can be stated as follows (where $I_A(s)$ is the indicator function for the set A):

Proposition 3.1 *The following two conditions are equivalent for all* $F(\cdot;\theta)\in\Delta^1$*:*

(i) For all $\pi\in\Pi^{FO}\equiv\{\pi|\pi:\Re\to\Re,\ \text{nondecreasing}\}$, $\int_s \pi(s)dF(s;\theta)$ *is nondecreasing in* θ.

(ii) For all $\gamma\in\Gamma^{FO}\equiv\{\gamma|\gamma(s)=I_{[a,\infty)}(s),\ a\in\Re\}$, $\int_s \gamma(s)dF(s;\theta)$ *is nondecreasing in* θ.

This theorem has the same structure as the SOMSD theorem, Proposition 2.1′. The sets of payoff functions are illustrated in Figure 1. This theorem says that instead of checking that expected payoffs are nondecreasing in θ for all nondecreasing payoff functions, Π^{FO}, it is equivalent to check that expected payoffs are nondecreasing in θ for all payoff functions in the set Γ^{FO}, the set of indicator functions of upper intervals. The latter set of payoff functions is much smaller, and so condition (ii) is easier to check that condition (i); condition (ii) can be reduced to a restriction which requires that $1-F(a;\theta)$, the complement of the cumulative distribution function, is nondecreasing in θ for $a \in \Re$. The latter requirement is the more standard way of stating FOSD.

There are many other examples of stochastic dominance theorems, some with multiple random variables; we will discuss other examples below in Section 4.5. In general, stochastic dominance theorems all have a parallel structure, as illustrated in Propositions 2.1′ and 3.1. However, different stochastic dominance theorems are characterized by different (Π,Γ) pairs. The SOMSD theorem states that the pair $(\Pi^{SOM}, \Gamma^{SOM})$ satisfies a particular relationship; the FOSD theorem states that the pair (Π^{FO}, Γ^{FO}) satisfies the same relationship. To make this relationship precise, we introduce a formal definition of the statement "A pair of sets of payoff functions, (Π,Γ), satisfies a stochastic dominance theorem." If that statement is true, then we will say that (Π,Γ) is a *stochastic dominance pair*. We will use the abstract definition to make statements about the class of stochastic dominance theorems, and how that class of theorems relates to other classes of theorems. We allow for multidimensional payoff functions and probability distributions, using the following notation: the set of probability distributions on $\Re^n$ is denoted Δ^n, with typical element $F:\Re^n \to [0,1]$. Further, for a given parameter space Θ, we will use the notation Δ^n_Θ to represent the set of parameterized probability distributions $F:\Re^n \times \Theta \to [0,1]$ such that such that $F(\cdot;\theta) \in \Delta^n$ for all $\theta \in \Theta$.

Definition 3.1 Consider a pair of sets of payoff functions (Π,Γ), *with typical elements* $\pi:\Re^n \to \Re$ *and* $\gamma:\Re^n \to \Re$. *The pair* (Π,Γ) *is a* ***stochastic dominance pair*** *if conditions (i) and (ii) are equivalent for all parameter spaces* Θ *with a partial order and all* $F \in \Delta^n_\Theta$*:*

(i) For all $\pi \in \Pi$, $\int_s \pi(s)dF(s;\theta)$ *is* ***nondecreasing*** *in* θ.

(ii) For all $\gamma \in \Gamma$, $\int_s \gamma(s)dF(s;\theta)$ *is* ***nondecreasing*** *in* θ.

Further, we define the set Σ_{SDT} to be the set of all (Π,Γ) pairs which are stochastic dominance pairs, as follows:

$$\Sigma_{SDT} = \{(\Pi,\Gamma) | (\Pi,\Gamma) \text{ is a stochastic dominance pair}\}$$

Thus, when a given (Π,Γ) pair is a stochastic dominance pair, we write $(\Pi,\Gamma) \in \Sigma_{SDT}$.

Definition 3.1 clarifies the structure of stochastic dominance theorems. Stochastic dominance theorems identify pairs of sets of payoff functions which have the following property: given a parameterized probability distribution F, checking that all of the functions in the set $\left\{\int_s \pi(s)dF(s;\theta) | \pi \in \Pi\right\}$ are nondecreasing is *equivalent* to checking that all of the functions in the set

$\left\{\int_s \gamma(s)dF(s;\theta) \middle| \gamma \in \Gamma\right\}$ are nondecreasing. Stochastic dominance theorems are useful because, in general, the set Γ is smaller than the set Π.

We can think of this definition as a statement about the equality of two sets. First, let us define the set of admissible parameter spaces for the property "nondecreasing" together with probability distributions parameterized on those spaces as follows:

$$\mathcal{D}^n_{ND} \equiv \left\{(F,\Theta) \middle| \Theta \text{ has a partial order and } F \in \Delta^n_\Theta\right\}$$

Now, we can rephrase the definition as follows: (Π,Γ) is a stochastic dominance pair if

$$\left\{(F,\Theta) \in \mathcal{D}^n_{ND} \middle| \int_s \pi(s)dF(s;\theta) \text{ is nondecreasing on } \Theta \ \forall \pi \in \Pi\right\} =$$

$$\left\{(F,\Theta) \in \mathcal{D}^n_{ND} \middle| \int_s \gamma(s)dF(s;\theta) \text{ is nondecreasing on } \Theta \ \forall \gamma \in \Gamma\right\}$$

Definition 3.1 differs from the existing literature (i.e., Brumelle and Vickson, 1975), in that the existing literature generally compares the expected value of two probability distributions, say F and G, viewing $\int_s \pi(s)dF(s)$ and $\int_s \pi(s)dG(s)$ as two different linear functionals mapping payoff functions to $\Re$. In contrast, by parameterizing the probability distribution and viewing $\int_s \pi(s)dF(s;\theta)$ as a bilinear functional mapping payoff functions and (parameterized) probability distributions to the reals, we are able to create an analogy between stochastic dominance theorems and stochastic supermodularity theorems, an analogy which would not be obvious using the standard constructions. The utility of this definition will become more clear when we formalize the relationship between stochastic dominance and stochastic P theorems, for other properties P (such as supermodularity).

To provide specific examples of pairs of sets of payoff functions which satisfy stochastic dominance theorems, we summarize three univariate stochastic dominance theorems in Table II. There are potentially many other univariate stochastic dominance theorems (for example, theorems where the set of payoff functions imposes restrictions on the third derivative of the payoff function); however, we will simply report the three most familiar univariate stochastic dominance theorems here.

TABLE II

COMPONENTS OF UNIVARIATE STOCHASTIC DOMINANCE THEOREMS*

	Sets of Payoff Functions, Π	Sets of Payoff Functions, Γ
(i)	$\Pi^{FO} \equiv \{\pi \mid \pi : \Re \to \Re,\ \text{nondecreasing}\}$	$\Gamma^{FO} \equiv \{\gamma \mid \gamma(s) = I_{[a,\infty)}(s),\ a \in \overline{\Re}\}$
(ii)	$\Pi^{SO} \equiv \{\pi \mid \pi : \Re \to \Re,\ \text{concave}\}$	$\Gamma^{SO} \equiv \{\gamma \mid \gamma(s) = -s\}$ $\cup \{\gamma \mid \gamma(s) = \min(a,s),\ a \in \overline{\Re}\}$
(iii)	$\Pi^{SOM} \equiv \left\{\pi \,\middle\vert\, \begin{array}{l}\pi : \Re \to \Re,\ \text{nondecreasing},\\ \text{concave}\end{array}\right\}$	$\Gamma^{SOM} \equiv \{\gamma \mid \gamma(s) = \min(a,s),\ a \in \overline{\Re}\}$

*Each (Π,Γ) pair in Table II is a (univariate) stochastic dominance pair.

Table II (iii) corresponds to a SOMSD theorem, as shown in Proposition 2.1′, while Table II (i) corresponds to a FOSD theorem, Proposition 3.1. Let us now illustrate the interpretation of Definition 3.1 and Table II with a third example: Second Order Stochastic Dominance (SOSD), shown in Table II (ii). It is known that $\int_s \pi(s)dF(s;\theta)$ is nondecreasing in θ for all univariate, concave payoff functions if and only if $\int_s \min(a,s)dF(s;\theta)$ is nondecreasing in θ for all $a \in \Re$, and $\int_s s \cdot dF(s;\theta)$ does not depend on θ. Observe that both $\gamma(s) = s$ and $\gamma(s) = -s$ are included in Γ^{SO}; this forces the mean of the distribution to be both nonincreasing and nondecreasing, and hence constant in θ.

There are many other stochastic dominance theorems in addition to the univariate examples given above. Levy and Paroush (1974) derive results for bivariate functions, while Meyer (1990) extends these results and examines some multivariate stochastic dominance theorems as well. We will report some of these results in Section 4.5, where the main objective is to apply these results to problems of stochastic supermodularity.

3.2 *Exact Conditions for a Stochastic Dominance Theorem*

In this section, we study necessary and sufficient conditions for the pair (Π,Γ) to be a stochastic dominance pair. We want to specify the exact mathematical relationship which is equivalent to the statement that $(\Pi,\Gamma) \in \Sigma_{SDT}$. We will first discuss our result and its implications; then, in Sections 3.2.1 and 3.2.2, we will provide the mathematical arguments underlying the result.

The main result of this section is that $(\Pi,\Gamma) \in \Sigma_{SDT}$ if and only if the following statement is true: the closure (under an appropriate topology) of the convex cone of $\Pi \cup \{1,-1\}$ is equal to the closure (under that topology) of the convex cone of $\Gamma \cup \{\mathbf{1},\mathbf{-1}\}$, where $\{\mathbf{1},\mathbf{-1}\}$ denotes the set containing the two constant functions, $\{\pi(s)=1\} \cup \{\pi(s)=-1\}$. We formalize this using the following notation:

$$\overline{cc(\Pi \cup \{\mathbf{1},\mathbf{-1}\})} = \overline{cc(\Gamma \cup \{\mathbf{1},\mathbf{-1}\})} \tag{3.1}$$

In the context of specific sets of payoff functions Π, the existing literature identifies similar, but stronger sufficient conditions for the corresponding stochastic dominance theorems, using a more restrictive notion of closure (i.e., a topology with more open sets) than the one which we will identify below. For example, Brumelle and Vickson (1975) argue that (3.1) is sufficient for the (Π,Γ) pairs shown in Table II to be stochastic dominance pairs under the topology of monotone convergence. In this paper, using the abstract definition we have developed for a "stochastic dominance pair," we are able to formally prove that the sufficient conditions hypothesized by Brumelle and Vickson (1975) are in fact sufficient for *any* stochastic dominance theorem, not just particular examples. Further, the result that (3.1) is also necessary for (Π,Γ) to be a stochastic dominance pair is a new contribution of this paper.

We now argue that this result tells us when we cannot do better than the closed convex cone method. First, observe that unless Γ is a subset of Π, there is no guarantee that stochastic dominance theorems provide conditions which are easier to check than $\int_s \pi(s)dF(s;\theta)$ nondecreasing in θ for all $\pi \in \Pi$. For example, Π might be a set which is not a closed convex cone. Its closed convex cone might be much larger, and it might not be possible to find a subset, Γ, of that closed convex cone for which is easier to check that expected payoffs are nondecreasing in θ. Thus, (3.1) indicates that stochastic dominance theorems, as defined in Definition 3.1, are most likely to be useful when Π is a closed convex cone. For example, in the case of FOSD, we consider the set of payoff functions Π^{FO}. It is easy to verify that positive scalar multiples and convex combinations of nondecreasing functions are nondecreasing functions, as are limits of sequences or nets of nondecreasing functions. Finally, constant functions are also in Π^{FO}.

When Π contains the constant functions and is a closed convex cone, (3.1) becomes:

$$\Pi = \overline{cc(\Gamma \cup \{\mathbf{1},\mathbf{-1}\})} \tag{3.2}$$

In principle, the most useful Γ is the smallest set whose closed convex cone is Π. However, in general, there will not be a unique smallest set. To see this, consider the case of FOSD, where the set Γ^{FO} contains indicator functions of upper intervals. By taking limits of sequences of convex combinations of elements of $\Gamma^{FO} \cup \{\mathbf{1},\mathbf{-1}\}$, and appropriately scaling these functions, we can generate any nondecreasing function. However, we can also define the set $\tilde{\Gamma}^{FO} \equiv \left\{\gamma \middle| \gamma(s) = I_{[a,\infty)}(s),\ a \in Q\right\}$, where Q represents the rationals, and note that $\Pi^{FO} = \overline{cc(\tilde{\Gamma}^{FO} \cup \{\mathbf{1},\mathbf{-1}\})}$. While $\tilde{\Gamma}^{FO} \subset \Gamma^{FO}$, in practice the smaller set $\tilde{\Gamma}^{FO}$ is not any easier to check. Thus, stochastic dominance theorems are generally stated so that Γ is the smallest *closed* set whose closed convex cone is Π; we will call such a set the "extreme points" of Π.

Finally, because (3.2) is necessary and sufficient for (Π,Γ) to be a stochastic dominance pair when $\Pi=\overline{cc(\Pi \cup \{\mathbf{1},-\mathbf{1}\})}$, we know that we cannot do any better than letting Γ be the set of extreme points of Π: there is no smaller or easier-to-check closed set of payoff functions, $\hat{\Gamma}$, such that $(\Pi,\hat{\Gamma})$ is a stochastic dominance pair. This is what we mean when we say that we have proved that the closed convex cone method is the "right" approach.

In the next two subsections, we prove that (3.1) characterizes stochastic dominance pairs. Consider the problem of ordering two probability distributions, F^1 and F^2, where we say that a set of payoff functions Π orders F^1 higher (lower) than F^2 if $\int_s \pi(s)dF^1(s) \geq (\leq) \int_s \pi(s)dF^2(s)$ for all π in Π; if neither inequality holds for all π in Π, then we say that the distributions are not ordered by Π. In Section 3.1.1, we first show that Π orders two distributions exactly the same as the set $\overline{cc(\Pi \cup \{\mathbf{1},-\mathbf{1}\})}$. We then show that two sets, Π and Γ, order arbitrary pairs of probability distributions in the same way *if and only if* (3.1) holds, that is, their closed convex cones are the same. In Section 3.2.2, we use these results to prove the characterization of stochastic dominance pairs according to (3.1).

3.2.1 The Main Mathematical Results

We begin by proving some general mathematical results about linear functionals of the form $\int \pi \, d\mu$. Our results are variations on standard theorems from the theories of linear functional analysis, topological vector spaces, and linear algebra; the main contribution of this section is to define the appropriate function spaces and topology and restate the problem in such a way that we can adapt these theorems to solve the stochastic dominance problem.

We will work with a class of objects known as *finite signed measures* on $\Re^n$, denoted $\mathcal{M}^*$. Any finite signed measure μ has a "Jordan decomposition" (see Royden (1968), pp. 235-236), so that $\mu = \mu^+ - \mu^-$, where each component is a positive, finite measure. We will be especially interested in finite signed measures which have the property that $\int d\mu = 0$, so that $\int d\mu^+ = \int d\mu^-$. Denote the set of all non-zero finite signed measures which have this property by $\mathcal{Z}^*$; we are interested in this set because elements of this set can always be written as $\mu = \frac{1}{k}[F^1 - F^2]$, where k is a positive scalar, and F^1 and F^2 are probability distributions; likewise, for any two probability distributions F^1 and F^2, the measure $F^1 - F^2 \in \mathcal{Z}^*$.

In this section, we will prove results which involve inequalities of the form $\int \pi \, d\mu \geq 0$, where $\mu \in \mathcal{Z}^*$. Since positive scalar multiples will not affect this inequality, and because the integral operator is linear, we can without loss of generality interpret this inequality as $\int_s \pi(s)dF^1(s) \geq \int_s \pi(s)dF^2(s)$ for the appropriate pair of probability distributions. The former notation will be easier to work with in terms of proving our main results, and further it will be useful in proving results in Section 4 about stochastic P theorems for properties P other than "nondecreasing."

Now let us begin our formal analysis. Let $\mathcal{P}^*$ be the set of bounded, measurable payoff functions on $\Re^n$. Define the bilinear functional $\beta : \mathcal{P}^* \times \mathcal{M}^* \to \Re$ by $\beta(\pi,\mu) = \int \pi d\mu$. Then $\beta(\mathcal{P}^*,\mathcal{M}^*)$ is a *separated duality*: that is, for any $\mu_1 \neq \mu_2$, there is a $\pi \in \mathcal{P}^*$ such that $\beta(\pi,\mu_1) \neq \beta(\pi,\mu_2)$, and for any $\pi_1 \neq \pi_2$, there is a $\mu \in \mathcal{M}^*$ such that $\beta(\pi_1,\mu) \neq \beta(\pi_2,\mu)$.[7] Our choice of $(\mathcal{P}^*,\mathcal{M}^*)$ is somewhat arbitrary: all of our results hold if we let A^n be a subset of measurable payoff functions on $\Re^n$ and we let B^n be any subset of $\mathcal{M}^*$, so long as $\beta(A^n,B^n)$ is a separated duality. For consistency we will use the pair $(\mathcal{P}^*,\mathcal{M}^*)$ in our formal analysis.

We now construct our topology, where would like to find the coarsest topology (that is, the topology with the fewest open sets) such that the set of all continuous linear functionals on $\mathcal{P}^*$ (the *dual* of $\mathcal{P}^*$) is exactly the set $\{\beta(\cdot,\mu) | \mu \in \mathcal{M}^*\}$; this is the weak topology $\sigma(\mathcal{P}^*,\mathcal{M}^*)$ on $\mathcal{P}^*$. By Bourbaki (1987, p. II.43), this topology uses as a basis neighborhoods of the form $N(\pi;\varepsilon,(\mu_1,\ldots,\mu_k)) = \left\{\hat{\pi} \,\middle|\, \max_{i=1,\ldots,k} |\beta(\pi - \hat{\pi},\mu_i)| < \varepsilon \right\}$, where there is a neighborhood corresponding to each finite set $(\mu_1,\ldots,\mu_k) \subset \mathcal{M}^*$ and each $\varepsilon > 0$. We will return to clarify the relationship between this topology and other topologies in the discussion following Theorem 3.4, below.

Let $cc(A)$ denote the convex cone of a set A, and let $\overline{A}$ denote the closure of A (where the topology is understood to be $\sigma(\mathcal{P}^*,\mathcal{M}^*)$ in the discussion below, unless noted). We now use the fact that the functional $\beta(\pi,\mu) = \int \pi d\mu$ is linear and continuous in its first argument to prove the following simple lemma. The proof of this lemma is elementary, but we state it here because all of the mathematical results in this paper build upon it.

Lemma 3.2 *Consider a set of payoff functions $\Pi \subseteq \mathcal{P}^*$. Then the following two conditions are equivalent for all $\mu \in \mathcal{Z}^*$:*

(i) For all $\pi \in \Pi$, $\int \pi \, d\mu \geq 0$.

---

[7]The boundedness assumption guarantees that the integral of the payoff function exists. It is possible to place other restrictions on the payoff functions and the space of finite signed measures so that the pair is a separated duality, in which case the arguments below would be unchanged; for example, it is possible to restrict the payoff functions and the signed measures using a "bounding function." For more discussions of separated dualities, see Bourbaki (1987, p. II.41).

(ii) For all $\pi \in \overline{cc(\Pi \cup \{1,-1\})}$, $\int \pi \, d\mu \geq 0$.

Proof: First, we show (i) implies (ii). Fix a measure $\mu \in \mathfrak{F}^*$. Then the following implications hold:

$$\textit{For all } \pi \in \Pi, \ \int \pi \, d\mu \geq 0$$

$$\Rightarrow \qquad \textit{For all } \pi \in \Pi \cup \{1,-1\}, \ \int \pi \, d\mu \geq 0$$

(Since $\int d\mu = -\int d\mu = 0$).

$$\Rightarrow \qquad \textit{For all } \pi \in cc(\Pi \cup \{1,-1\}), \ \int \pi \, d\mu \geq 0$$

(Since $\int \pi_1 \, d\mu \geq 0$ and $\int \pi_2 \, d\mu \geq 0$ implies $\int [\alpha_1 \pi_1 + \alpha_2 \pi_2] \, d\mu = \alpha_1 \int \pi_1 \, d\mu + \alpha_2 \int \pi_2 \, d\mu \geq 0$ when $\alpha_1, \alpha_2 \geq 0$).

$$\Rightarrow \qquad \textit{For all } \pi \in \overline{cc(\Pi \cup \{1,-1\})}, \ \int \pi \, d\mu \geq 0$$

(Recall that for a continuous function *f*, $f(\overline{A}) \subseteq \overline{f(A)}$. The implication then follows because the half-space $\{x \in \Re | x \geq 0\}$ is closed, and the linear functional $\beta(\cdot;\mu)$ is continuous for all μ.)

That (ii) implies (i) follows because $\Pi \subseteq \overline{cc(\Pi \cup \{1,-1\})}$. *Q.E.D.*

Lemma 3.2 can be restated another way: the set of measures $\mu \in \mathfrak{F}^*$ for which $\int \pi \, d\mu \geq 0$ for all $\pi \in \Pi$ is exactly the same as the set of measures $\mu \in \mathfrak{F}^*$ for which $\int \pi \, d\mu \geq 0$ for all $\pi \in \overline{cc(\Pi \cup \{1,-1\})}$. Formally,

$$\left\{ \mu \in \mathfrak{F}^* \middle| \int \pi \, d\mu \geq 0 \ \ \forall \pi \in \Pi \right\} = \left\{ \mu \in \mathfrak{F}^* \middle| \int \pi \, d\mu \geq 0 \ \ \forall \pi \in \overline{cc(\Pi \cup \{1,-1\})} \right\}$$

As discussed above, because positive scalar multiples do not reverse the inequalities and because the functional β is linear in its second argument, the latter equality is equivalent to the following:

$$\left\{ F^1, F^2 \in \Delta^n \middle| \int \pi \, dF^1 \geq \int \pi \, dF^2 \ \ \forall \pi \in \Pi \right\} = \left\{ F^1, F^2 \in \Delta^n \middle| \int \pi \, dF^1 \geq \int \pi \, dF^2 \ \ \forall \pi \in \overline{cc(\Pi \cup \{1,-1\})} \right\}$$

That is, for any two probability distributions, Π orders the two distributions exactly the same as $\overline{cc(\Pi \cup \{1,-1\})}$.

Building from this lemma, we turn to prove the main mathematical theorem underlying the characterization of stochastic dominance theorems, as well as the characterizations of other properties in Section 4. This theorem makes use of the linearity of the functional $\beta(\pi,\mu)$ in π. The proof that (ii) implies (i) relies on Lemma 3.2, while the proof that (i) implies (ii) makes use of a standard separating hyperplane argument. Note that the choice of topology, which determines the meaning of closure, is critical for the application of the separating hyperplane theorem.

Theorem 3.3 Consider a pair of sets of payoff functions (Π,Γ), *where* Π *and* Γ *are subsets of* $\mathcal{P}^*$. *Then the following two conditions are equivalent:*

(i) $\left\{\mu \in \mathcal{F}^* \middle| \int \pi \, d\mu \geq 0 \;\; \forall \pi \in \Pi\right\} = \left\{\mu \in \mathcal{F}^* \middle| \int \gamma \, d\mu \geq 0 \;\; \forall \gamma \in \Gamma\right\}$.

(ii) $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$.

Proof: First consider (ii) implies (i). Suppose that $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$. Then we have:

$$\left\{\mu \in \mathcal{F}^* \middle| \int \pi \, d\mu \geq 0 \;\; \forall \pi \in \Pi\right\}$$

$$= \left\{\mu \in \mathcal{F}^* \middle| \int \pi \, d\mu \geq 0 \;\; \forall \pi \in \overline{cc(\Pi \cup \{1,-1\})}\right\}$$

(By Lemma 3.2.)

$$= \left\{\mu \in \mathcal{F}^* \middle| \int \gamma \, d\mu \geq 0 \;\; \forall \gamma \in \overline{cc(\Gamma \cup \{1,-1\})}\right\}$$

(By assumption.)

$$= \left\{\mu \in \mathcal{F}^* \middle| \int \gamma \, d\mu \geq 0 \;\; \forall \gamma \in \Gamma\right\}$$

(By Lemma 3.2.)

Now we prove that (i) implies (ii). Define $\tilde{\Pi} \equiv \overline{cc(\Pi \cup \{1,-1\})}$ and $\tilde{\Gamma} \equiv \overline{cc(\Gamma \cup \{1,-1\})}$. Suppose (without loss of generality) that there exists a $\hat{\gamma} \in \tilde{\Gamma}$ such that $\hat{\gamma} \notin \tilde{\Pi}$. We know that the $\sigma(\mathcal{P}^*,\mathcal{M}^*)$ topology is generated from a family of open, convex neighborhoods. Recall from above that the set of continuous linear functionals on $\mathcal{P}^*$ is exactly the set $\left\{\beta(\cdot,\mu) \middle| \mu \in \mathcal{M}^*\right\}$. Using these facts, a corollary to the Hahn-Banach theorem[8] implies that

[8]See Dunford and Schwartz (1957, p. 421), Kothe (1969, p. 244) for discussions of the relevant theorems about the separation of convex sets. See also McAfee and Reny (1992) for a related application of the Hahn-Banach theorem, where the separating hyperplane also takes the form of an element of $\mathcal{F}^*$.

since $\tilde{\Pi}$ is closed and convex, there exists a constant c and a $\mu_* \in \mathcal{M}^*$ (a separating hyperplane) so that $\beta(\pi,\mu_*) \geq c$ for all $\pi \in \tilde{\Pi}$, and $\beta(\hat{\gamma},\mu_*) < c$.

Since $\{1,-1\} \in \tilde{\Pi}$ and is $\tilde{\Pi}$ convex, $0 \in \tilde{\Pi}$ as well. Thus, $\beta(0,\mu_*) = 0 \geq c$. Now we will argue we can take $c = 0$ without loss of generality. Suppose not. Then there exists a $\hat{\pi} \in \tilde{\Pi}$ such that $c \leq \beta(\hat{\pi},\mu_*) = \hat{c} < 0$. Choose any positive scalar ρ such that $\rho > \frac{c}{\hat{c}} \geq 1$ (which implies that $\rho\hat{c} < c$). Since $\tilde{\Pi}$ is a cone, $\rho\hat{\pi} \in \tilde{\Pi}$. But, $\beta(\rho\hat{\pi},\mu_*) = \rho\hat{c} < c$, contradicting the hypothesis that $\beta(\pi,\mu_*) \geq c$ for all $\pi \in \tilde{\Pi}$. So, we let $c = 0$.

Because $\{1,-1\} \in \tilde{\Pi}$, and $\beta(\pi,\mu_*) \geq 0$ for all $\pi \in \tilde{\Pi}$, we conclude that $\beta(1,\mu_*) = -\beta(1,\mu_*) = 0$, and thus $\int d\mu_* = 0$. So, $\mu_* \in \mathcal{Z}^*$, and we have shown that $\int \pi \, d\mu_* \geq 0$ for all $\pi \in \tilde{\Pi}$, but $\int \hat{\gamma} d\mu_* < 0$, which violates condition (i). *Q.E.D.*

The proof of Theorem 3.3 is analogous to the proof of the "bipolar theorem" from the theory of topological vector spaces (see Schaefer, 1980, p. 126). This result is different because $\mathcal{M}^*$, not $\mathcal{Z}^*$, is the dual of the space $\mathcal{P}^*$; this accounts for the inclusion of the constant functions in condition (ii) of Theorem 3.3.

As above, we can restate condition (i) of Theorem 3.3 as follows, without loss of generality:

$$\left\{F^1,F^2 \in \Delta^n \middle| \int_S \pi \, dF^1 \geq \int_S \pi \, dF^2 \quad \forall \pi \in \Pi\right\} = \left\{F^1,F^2 \in \Delta^n \middle| \int_S \gamma dF^1 \geq \int_S \gamma dF^2 \quad \forall \gamma \in \Gamma\right\}$$

The theorem first states that if two sets of payoff functions have the same closed convex cone, then they will order any pair of probability distributions the same way. Second, the theorem states that if two sets of payoff functions order all pairs of probability distributions the same way, those two sets of payoff functions *must* have the same closed convex cone.

3.1.2 Exact Characterization of Stochastic Dominance Theorems

Building from the mathematical results of the last subsection, we now state the main result of Section 3. This theorem applies Theorem 3.3 to give necessary and sufficient conditions for a pair (Π,Γ) to satisfy a stochastic dominance theorem.

Theorem 3.4 *Consider a pair of sets of payoff functions* (Π,Γ), *where* Π *and* Γ *are subsets of* $\mathcal{P}^*$. *Then the following two conditions are equivalent:*

(i) The pair (Π,Γ) *is a stochastic dominance pair.*

(ii) $\overline{cc(\Pi \cup \{\mathbf{1},-\mathbf{1}\})} = \overline{cc(\Gamma \cup \{\mathbf{1},-\mathbf{1}\})}$.

Proof: To see that (ii) implies (i), fix a pair (Π,Γ) and assume that (ii) holds. Consider an arbitrary probability distribution $F(\cdot;\theta) \in \Delta^n$, and choose $\theta_H > \theta_L$. Define $F^1 \equiv F(\cdot;\theta_H)$ and $F^2 \equiv F(\cdot;\theta_L)$. Let $\mu = F^1 - F^2$, and note that $\mu \in \mathfrak{Z}^*$.

Now, note that Theorem 3.3 implies that

$$\textit{For all } \pi \in \Pi, \int \pi \, d(F^1 - F^2) \geq 0$$

$$\Leftrightarrow \quad \textit{For all } \gamma \in \Gamma, \int \gamma \, d(F^1 - F^2) \geq 0$$

Since this must be true for all $\theta_H > \theta_L$, then (Π,Γ) is a stochastic dominance pair.

To see that (i) implies (ii), note that if (ii) fails, then (without loss of generality) there exists some $\hat{\gamma} \in \tilde{\Gamma}$ such that $\hat{\gamma} \notin \tilde{\Pi}$. But then, by Theorem 3.3 there exists a $\mu_* \in \mathfrak{Z}^*$ so that $\int \pi \, d\mu_* \geq 0$ for all $\pi \in \tilde{\Pi}$, but $\int \hat{\gamma} d\mu_* < 0$. Let $k = \int d\mu_*^+ = \int d\mu_*^-$. Let $\Theta = \{\theta_L, \theta_H\}$, where $\theta_H > \theta_L$, and define a parameterized probability distribution as follows: $F(\cdot;\theta_H) \equiv \frac{1}{k}\mu_*^+$ and $F(\cdot;\theta_L) \equiv \frac{1}{k}\mu_*^-$. But then, $\int_s \pi(s) dF(s;\theta)$ is nondecreasing in θ for all $\pi \in \Pi$, while $\int_s \hat{\gamma}(s)\, dF(s;\theta)$ is strictly decreasing in θ. *Q.E.D.*

Theorem 3.4 provides an algorithm for generating stochastic dominance theorems, and for checking whether (Π,Γ) pairs satisfy stochastic dominance theorems. It gives the weakest possible sufficient conditions on a particular (Π,Γ) pair to guarantee that it is a stochastic dominance pair: if $\overline{cc(\Pi \cup \{1,-1\})} \neq \overline{cc(\Gamma \cup \{1,-1\})}$, then there will always exist a parameterized probability distribution such that $\int_s \gamma(s) dF(s;\theta)$ is nondecreasing in θ for all $\gamma \in \Gamma$, but $\int_s \pi(s) dF(s;\theta)$ is strictly decreasing in θ for some $\pi \in \Pi$.

However, Theorem 3.4 makes use of the weak topology, which might be less familiar than some others; for this reason, we will now discuss the relationship between closure in the weak topology and closure in other topologies. In particular, some existing stochastic dominance theorems have been proved by showing that the closure under monotone convergence of the convex cones of the two sets are equal (Brumelle and Vickson, 1975; Topkis, 1968). This raises the question, what is the relationship between those results, and results using the notion of closure under the weak topology? The following corollary answers that question.

Corollary 3.4.1 *Let τ be any topology on $\mathcal{P}^*$ such that the functionals $\{\beta(\cdot,\mu) | \mu \in \mathcal{M}^*\}$, are continuous linear functionals. Suppose that (Π,Γ) is a pair of sets of payoff functions, each in $\mathcal{P}^*$.*

If $\overline{cc(\Pi \cup \{1,-1\})^{\tau}} = \overline{cc(\Gamma \cup \{1,-1\})^{\tau}}$ *(closure taken with respect to* τ*), then* (Π,Γ) *is a stochastic dominance pair.*

Proof: τ is finer than $\sigma(\mathcal{P}^*,\mathcal{M}^*)$ (by Bourbaki, TVS II.43). For any set A in $\mathcal{P}^*$, we know $A \subseteq \overline{A^{\tau}} \subseteq \overline{A^{\sigma}}$. Taking closures in $\sigma(\mathcal{P}^*,\mathcal{M}^*)$ of each set, we conclude that $\overline{A^{\sigma}} = \overline{\overline{A^{\tau}}^{\sigma}}$. Thus, if $\overline{cc(\Pi \cup \{1,-1\})^{\tau}} = \overline{cc(\Gamma \cup \{1,-1\})^{\tau}}$, then $\overline{cc(\Pi \cup \{1,-1\})^{\sigma}} = \overline{cc(\Gamma \cup \{1,-1\})^{\sigma}}$. This implies that (Π,Γ) is a stochastic dominance pair by Theorem 3.4. *Q.E.D.*

To see the intuition behind this result, first note that, given two topologies τ_1 and τ_2, where τ_1 is coarser than τ_2 (written $\tau_1 \subseteq \tau_2$), the closure of a set A under τ_2 ($\overline{A^{\tau_2}}$) is contained in the closure of A under τ_1 ($\overline{A^{\tau_1}}$). This is true because the finer topology has more closed sets; $\overline{A^{\tau_1}}$ is a closed set under τ_2, but since τ_2 is finer, there might be a closed set which is a strict subset of $\overline{A^{\tau_1}}$ but that still contains A.

Since the weak topology is coarser than any topology which makes the linear functionals $\beta(\pi,\mu)$ continuous in π, the closure of a set of payoff functions under monotone convergence is contained in the closure of the set under the weak topology. Thus, if the closed (under monotone convergence) convex cones of two sets of payoff functions are the same, then the closed (under the weak topology) convex cones of those two sets of payoff functions will be the same. This tells us that checking that the closed (under monotone convergence) convex cones of two sets of payoff functions are the same is sufficient to establish that the two sets of payoff functions satisfy a stochastic dominance theorem.

Thus, in practice, when checking whether (Π,Γ) satisfy sufficient conditions for a stochastic dominance theorem, it is possible to check whether the closure of the convex cones of the two sets are equivalent, using any topology which is convenient (provided of course that the topology guarantees continuity of the functional $\beta(\cdot,\mu)$). The topologies of monotone convergence, dominated convergence, and uniform convergence are examples of topologies which might be useful in applications.

Remark 1

We have characterized the relationship between Π^{SOM} and Γ^{SOM} as follows:

$$\Pi^{SOM} = \overline{cc(\Pi^{SOM} \cup \{1,-1\})} = \overline{cc(\Gamma^{SOM} \cup \{1,-1\})}$$

Since much of the existing literature involves integration by parts, it is useful to illustrate the relationship between integration by parts and our characterization. We can view the integration by parts as a means of discovering the extreme points of a set of payoff functions which forms a closed convex cone.

In the case where the distribution function is continuous and differentiable with respect to the parameters, the payoff function is differentiable, and the random variable has compact support $[0,\bar{s}]$, both Propositions 2.1′ and 3.1 can be proved directly by first rewriting expected profits using integration by parts, and then taking the partial derivative of this expression with respect to the parameter(s). It is the integration by parts which determines the relationship between Π^{SOM} and Γ^{SOM}, and establishing this relationship plays a direct role in the proofs of both Propositions 2.1′ and 3.1. We can rewrite the functional $\beta(\pi,F)=\int_{s=0}^{\bar{s}}\pi(s)dF(s)$ as follows, using integration by parts:

$$\int_{s=0}^{\bar{s}}\pi(s)dF(s)$$
$$=\left[\pi(\bar{s})-\pi'(\bar{s})\cdot\bar{s}+\bar{s}\cdot\int_{s=0}^{\bar{s}}\pi''(s)\cdot ds\right]\cdot\int_{s=0}^{\bar{s}}dF(s) \qquad (3.3)$$
$$+\pi'(\bar{s})\cdot\int_{s=0}^{\bar{s}}s\,dF(s)-\int_{s=0}^{\bar{s}}\pi''(s)\cdot\int_{v=0}^{\bar{s}}\min(s,v)dF(v)ds$$

Note that the functions $\gamma(s)=s$ and $\gamma(s)=\min(a,s)$, which determine the set of functionals Γ^{SOM}, appear in this expression, as do the functions π' and π'', which characterize the set of payoff functions, Π^{SOM}. In fact, we can see that the functional $\int_{s=0}^{\bar{s}}\pi(s)dF(s)$ is the limit of a sequence of positive combinations of the expected value of the functions in Γ^{SOM}, plus linear combinations of $\int_{s=0}^{\bar{s}}dF(s)$. The set Π^{SOM} is defined so that π' and $-\pi''$ are positive; thus, the functions $\gamma(s)=s$ and $\gamma(s)=\min(a,s)$ are positive in Γ^{SOM}. The constant functions $\{\mathbf{1},\mathbf{-1}\}$ are included because $\int_{s=0}^{\bar{s}}dF(s)$ is constant at 1, but the coefficient

$$\left[\pi(\bar{s})-\pi'(\bar{s})\cdot\bar{s}+\bar{s}\cdot\int_{s=0}^{\bar{s}}\pi''(s)\cdot ds\right]$$

can be positive or negative.

To see how this is used in the study of monotonicity, note that we can evaluate the functional given in (3.1) at a parameterized probability distribution $F(\cdot;\theta)$ (i.e., take $\beta(\pi,F(\cdot;\theta))$) and then differentiate, as follows:

$$\frac{\partial}{\partial\theta}\int_{s=0}^{\bar{s}}\pi(s)dF(s;\theta)$$
$$=\pi'(\bar{s})\cdot\frac{\partial}{\partial\theta}\int_{s=0}^{\bar{s}}s\,dF(s;\theta)-\int_{s=0}^{\bar{s}}\pi''(s)\cdot\frac{\partial}{\partial\theta}\int_{v=0}^{\bar{s}}\min(s,v)dF(v;\theta)ds \qquad (3.4)$$

This expression can be used to prove the SOMSD theorem directly. If the mean and $\int_{v=0}^{\bar{s}}\min(v,s)dF(v;\theta)$ are nondecreasing in θ, then for any nondecreasing, concave payoff function, this expression is clearly positive. On the other hand, if the mean or $\int_{v=0}^{\bar{s}}\min(v,s)dF(v;\theta)$ were decreasing in θ somewhere, we could construct a nondecreasing, concave payoff function which put all of the weight on the failures. What the integration by parts shows us is that instead of checking that $\beta(\pi,F(\cdot;\theta))$ is nondecreasing in θ for all $\pi\in\Pi^{SOM}$ (the left-hand side of (3.4)), we can check that $\beta(\gamma,F(\cdot;\theta))$ is nondecreasing in θ for all $\gamma\in\Gamma^{SOM}$ (as motivated by the right-hand side of (3.4)). This is useful because the second set of conditions is easier to check. The constant functions $\{1,-1\}$ from above do not appear in Γ^{SOM} because $\int_{s=0}^{\bar{s}} 1 \cdot dF(s;\theta) = 1$ is always constant in θ. We emphasize the fact that the sets Π^{SOM} and Γ^{SOM} are determined by the integration by parts (equation 3.3), *not* by the subsequent differentiation with respect to the parameters of the distribution.

In an analogous way, we can evaluate the functional $\int_{s=0}^{\bar{s}} \pi(s)dF(s)$ at the parameterized probability distribution $F(\cdot;\theta_1,\theta_2)$, using equation (3.3), and take the mixed partial derivative:

$$\begin{aligned} &\frac{\partial^2}{\partial\theta_1\partial\theta_2}\int_{s=0}^{\bar{s}} \pi(s)dF(s;\theta_1,\theta_2) \\ &= \pi'(\bar{s})\frac{\partial^2}{\partial\theta_1\partial\theta_2}\int_{s=0}^{\bar{s}} s\, dF(s;\theta_1,\theta_2) + \int_{s=0}^{\bar{s}} \pi''(s)\frac{\partial^2}{\partial\theta_1\partial\theta_2}\int_{v=0}^{\bar{s}} \min(s,v)dF(v;\theta_1,\theta_2)ds \end{aligned} \tag{3.5}$$

The proof of the stochastic supermodularity theorem follows similar arguments to that of SOMSD. This example clarifies the relationship between the new stochastic supermodularity theorem and the existing stochastic dominance result.

3.3 *Conditional Stochastic Dominance*

In this section, we show that Theorem 3.4 can be extended to the case of "conditional stochastic dominance theorems." We will use the following notation:

$$G(s;\theta,K) \equiv \frac{\int_{K\cap\{t|t<s\}} dF(t;\theta)}{\int_K dF(t;\theta)}$$

$$\bar{\pi}(\theta|K) \equiv \frac{\int_{s\in K} \pi(s)dF(s,\theta)}{\int_{s\in K} dF(s,\theta)} = \beta(\pi \cdot I_K, G(s;\theta,K))$$

We now introduce an object which is analogous to a stochastic dominance theorem:

Definition 3.2 *Consider a pair of sets of payoff functions* (Π,Γ), *with typical elements* $\pi:\Re^n \to \Re$ *and* $\gamma:\Re^n \to \Re$. *Let* K *be a collection of subsets of* $\Re^n$. *Then the pair* (Π,Γ) *is a* K*–* ***conditional stochastic dominance pair*** *if conditions (i) and (ii) are equivalent for all* Θ *with a partial order and all* $F:\Re^n \times \Theta \to [0,1]$ *such that such that* $F(\cdot;\theta) \in \Delta^n$*:*

(i) For all $\pi \in \Pi$ *and all* $K \in$ K, $\bar{\pi}(\theta|K)$ *is* ***nondecreasing in*** θ.

(ii) For all $\gamma \in \Gamma$ *and all* $K \in$ K, $\bar{\gamma}(\theta|K)$ *is* ***nondecreasing in*** θ.

We now extend Theorem 3.4.

Theorem 3.5 Consider a pair of sets of payoff functions (Π,Γ), *where* Π *and* Γ *are subsets of* $\mathcal{P}$. *Let* K *be a collection of subsets of* $\Re^n$, *where* $\Re^n \in$ K. *Then the following two conditions are equivalent:*

(i) *The pair* (Π,Γ) *is a* K*–conditional stochastic dominance pair.*

(ii) $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$.

Proof: We can apply the proof of Theorem 3.4 almost exactly. Let $\Pi_K = \{\pi \cdot I_K | \pi \in \Pi\}$, and likewise for Γ_K. Then note that $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$ implies that $\overline{cc(\Pi_K \cup \{I_K,-I_K\})} = \overline{cc(\Gamma_K \cup \{I_K,-I_K\})}$ for all K. Then, for every K we apply Theorem 3.4, (ii) implies (i). To show that (i) implies (ii), the arguments of Theorem 3.4 can be used to show that if (ii) fails, then (i) must fail for the case where $K = \Re^n$. *Q.E.D.*

One example of a conditional stochastic dominance theorem which has appeared in various forms in the literature (see Whitt (1982)) involves the set of nondecreasing payoff functions, as follows:

Theorem 3.6 The following conditions are equivalent:

(i) *For all* $\pi : \Re^n \rightarrow \Re$ *nondecreasing and all sublattices* $K \subseteq \Re^n$, $\bar{\pi}(\theta|K)$ *is nondecreasing in* θ.

(ii) *For all increasing sets* $A \subseteq \Re^n$ *and all sublattices* $K \subseteq \Re^n$, $\int_{A \cap K} dG(s;\theta,K)$ *is nondecreasing in* θ.

Proof: Apply Theorem 3.5 together with the fact that $\overline{cc(\{\pi|\pi \text{ nondecreasing}\} \cup \{1,-1\})} = \overline{cc(\{I_A|I_A(s) \text{ nondecreasing}\} \cup \{1,-1\})}$. *Q.E.D.*

This theorem is usually proved using algebraic arguments; but using the results of this paper, we see that it follows as an immediate corollary of Theorem 3.5. First, take the case where $n = 1$. Then condition (ii) is equivalent to requiring that $F(s;\theta)$ satisfies the Monotone Likelihood Ratio (MLR) Order (for a proof, see Whitt (1980)), defined as follows:

Definition 3.2 The parameter θ *indexes the probability distribution* $F(\cdot;\theta) \in \Delta^1$ *according to the* ***Monotone Likelihood Ratio Order*** *(MLR) if, for all* $\theta_H > \theta_L$, *there exist numbers* $-\infty \le a \le b \le \infty$ *and a nondecreasing function* $h : [a,b] \rightarrow \Re$ *such that* $F(a;\theta_H) = 0$, $F(b;\theta_L) = 1$, *and* $\int_K dF(s;\theta_H) = \int_K h(s) \cdot dF(s;\theta_L)$ *for all* $K \subseteq [a,b]$.

When the support of F is constant in θ, and F has a density f, then an equivalent requirement is that f satisfies the Monotone Likelihood Ratio Property (MLRP), as follows:

$$\frac{f(s^H;\theta)}{f(s^L;\theta)} \text{ is nondecreasing in } \theta \text{ for all } s^H \ge s^L$$

Milgrom (1981) introduced the MLRP to the economics literature in the context of signaling problems. The density f satisfies the MLRP if and only if the log of the density is supermodular. The MLR implies FOSD (but not the reverse); that is, if $F(s;\theta)$ satisfies MLR, then the distribution $F(s;\theta)$ is decreasing in θ pointwise. FOSD does not place any restrictions on the movements of $F(s;\theta)$ apart from the restriction that changes in θ do not cause the distribution to cross. The MLR, on the other hand, requires that for any sublattice K, the distribution conditional on K satisfies FOSD.

Now, suppose that $n > 1$. If we assume that the vector s is a vector of affiliated random variables (see Milgrom and Weber (1982), who define affiliation and show that if the distribution has a density, the log of that density must be supermodular[9]), then condition (ii) requires that each marginal, $F_i(s_i;\theta)$, satisfies MLR. The most general case, where the vector s is an arbitrary vector of random variables, has not to our knowledge been analyzed in the literature.

We have presented Theorem 3.6 to illustrate the relationship between several existing results. The framework developed in this paper shows that the proof of Theorem 3.6 is an immediate consequence; previous analyses have relied on more complicated arguments. But, Theorem 3.6 is just one example; in his analysis of firm entry and exit in a dynamic environment, Hopenhayn (1992) derives another conditional stochastic dominance result, which he terms monotone conditional dominance. This result considers the case of $\Pi = \Pi^{FO}$ and K is the set of indicator functions of upper intervals. Our theorem shows how to generate other conditional stochastic dominance theorems as applications arise.

4 OTHER PROPERTIES OF STOCHASTIC OBJECTIVE FUNCTIONS

This section derives necessary and sufficient conditions for the objective function, $\int_s \pi(s)dF(s;\theta)$, to satisfy properties P other than "nondecreasing in θ," for example, the properties supermodular or concave in θ. We ask two questions: (i) For what properties P is the closed convex cone method of proving stochastic P theorems valid? (ii) For what properties P can we establish that the closed convex cone approach is exactly the right one, as in the case of stochastic dominance?

We proceed as follows. In Section 4.1, we introduce notation and definitions which extend our stochastic dominance framework to "stochastic supermodularity theorems," and for arbitrary properties P, "stochastic P theorems." Section 4.2 answers question (i), showing that the closed convex cone approach to proving stochastic dominance theorems is valid for all "stochastic P theorems," where P is a CCC property. In Section 4.3 we introduce a new class of properties, Linear Difference Properties (LDPs), which is a subset of CCC properties. We show that examples of LDPs include convexity,

[9]In fact, Milgrom and Weber (1982) also show that, for general distributions, affiliation is equivalent to log-supermodularity of the $\Re^n$ derivative on a product measure.

supermodularity, and all properties which place a sign restriction on a derivative (refer to Table I for a summary of CCC properties and LDPs). Section 4.4 addresses question (ii), proving that the closed convex cone approach is exactly the right one for all stochastic *P* theorems when *P* is an LDP. Section 4.5 applies the results of this section to the case of supermodularity, providing examples of (Π,Γ) pairs from the stochastic dominancc literature (which are also stochastic supermodularity pairs, by our main result) and discussing applications of stochastic supermodularity theorems.

4.1 Stochastic Supermodularity Theorems and Stochastic P Theorems

We now introduce a construct which we will call *stochastic P theorems*, which are precisely analogous to stochastic dominance theorems. Let us begin with a specific example, stochastic supermodularity theorems. Using two parallel definitions, we will be able to identify the close relationship between the stochastic dominance theorems and stochastic supermodularity theorems. Stochastic *dominance* theorems focus on monotonicity of expected profits with respect to a single parameter, while stochastic *supermodularity* theorems examine whether increasing one parameter increases the returns (in expected profits) to increasing the other parameter (recall that supermodularity can be checked pairwise). Proposition 2.3 is an example of a stochastic supermodularity theorem. The following definition is precisely analogous to Definition 3.1:

Definition 4.1 *Consider a pair of sets of payoff functions* (Π,Γ), *with typical elements* $\pi:\Re^n\rightarrow\Re$ *and* $\gamma:\Re^n\rightarrow\Re$. *The pair* (Π,Γ) *is a* ***stochastic supermodularity pair*** *if conditions (i) and (ii) are equivalent for all lattices* Θ *and all* $F\in\Delta^n_\Theta$:

(i) For all $\pi\in\Pi$, $\int_s \pi(s)dF(s;\theta)$ *is* ***supermodular*** *in* θ.

(ii) For all $\gamma\in\Gamma$, $\int_s \gamma(s)dF(s;\theta)$ *is* ***supermodular*** *in* θ.

Further, we define the set Σ_{SST} to be the set of all (Π,Γ) pairs which are stochastic supermodularity pairs, analogous to Σ_{SDT}. To place our example from Proposition 3.3 in this framework, the pair (Π^{SOM},Γ^{SOM}) is in Σ_{SST}.

When comparing Definition 3.1 (stochastic dominance theorem) and Definition 4.1 (stochastic supermodularity theorem), it is helpful to recall that stochastic dominance theorems and stochastic supermodularity theorems are both characterized by (Π,Γ) pairs corresponding to sets of linear functionals of the form $\beta(\pi,\cdot)$, which are then composed with parameterized probability distributions. Thus, it is possible to compare the two types of theorems directly, even though in a stochastic dominance theorem, the parameter space is any set with a partial order, while in a stochastic dominance theorem, the parameter space is any lattice.

More generally, we can define a class of theorems called *Stochastic P Theorems,* which are defined for an arbitrary property P. We are interested in properties P together with parameter spaces Θ_P which are defined so that, given a function $h:\Theta_P \rightarrow \Re$, the statement "$h(\theta)$ satisfies property P on Θ_P" is well-defined and takes on the following values: "true" or "false." Let $\overline{\Theta}_P$ denote the set of all such parameter spaces Θ_P. Further, for a given property P, we will define the set of admissible parameter spaces together with probability distributions parameterized on those spaces:

$$\mathcal{D}_P^n \equiv \left\{(F,\Theta_P) \middle| \Theta_P \in \overline{\Theta}_P \text{ and } F \in \Delta_{\Theta_P}^n\right\}.$$

Definition 4.2 Consider a pair of sets of payoff functions (Π,Γ), *with typical elements* $\pi:\Re^n \rightarrow \Re$ *and* $\gamma:\Re^n \rightarrow \Re$. *The pair* (Π,Γ) *is a* ***stochastic P pair*** *if conditions (i) and (ii) are equivalent for all* $(F,\Theta_P) \in \mathcal{D}_P^n$:

(i) For all $\pi \in \Pi$, $\int_s \pi(s)dF(s;\theta)$ ***satisfies property P on*** Θ_P.

(ii) For all $\gamma \in \Gamma$, $\int_s \gamma(s)dF(s;\theta)$ ***satisfies property P on*** Θ_P.

For example, for a stochastic concavity theorem, Θ_P can be any convex set, and condition (i) of Definition 4.2 is interpreted, "For all $\pi \in \Pi$, $\int_s \pi(s)dF(s;\theta)$ is concave in θ." As in the case of stochastic dominance, we let Σ_{SPT} be the set of all stochastic P pairs for a given property P. Also analogous to stochastic dominance, we can rewrite the requirement of Defintion 4.2 as follows:

$$\left\{(F,\Theta) \in \mathcal{D}_P^n \middle| \int \pi(s)dF(s;\theta) \text{ satisfies } P \text{ on } \Theta \ \forall \pi \in \Pi\right\}$$

$$= \left\{(F,\Theta) \in \mathcal{D}_P^n \middle| \int \gamma(s)dF(s;\theta) \text{ satisfies } P \text{ on } \Theta \ \forall \gamma \in \Gamma\right\}$$

In the next section, we show that the closed convex cone method of proving stochastic dominance theorems can be extended to all stochastic P theorems, if P is a closed convex cone property.

4.2 *The Closed Convex Cone Approach is Valid for all CCC Properties*

In this section, we identify a class of properties, which we call "Closed Convex Cone" (CCC) properties, such that the following statement is true: for all properties P which are CCC, if a pair (Π,Γ) is a stochastic dominance pair, then (Π,Γ) is a stochastic P pair; thus, for all CCC properties, the closed convex cone method of proving stochastic P theorems is valid. This result is useful because it allows us to use the existing theorems from the stochastic dominance literature to generate many new classes of theorems. CCC properties are properties are defined as follows:

Definition 4.3 A property P is a CCC property (written $P \in CCC$) if the set of functions $g:\Theta_P \to \Re$ which satisfy P forms a closed convex cone, where closure is taken with respect to the topology of pointwise convergence, and if constant functions satisfy P.

Note that we are using closure under the topology of pointwise convergence for properties P, while we are using closure under the weak topology (as defined in Section 3.1.1) for sets of payoff functions, Π and Γ.

The properties nondecreasing, concave, and supermodular are all CCC properties, as is the property "constant." Further, any property which places a sign restriction on a mixed partial derivative is CCC. Finally, since the intersection of two closed convex cones is itself a closed convex cone, any of these properties can be combined to yield another CCC property. For example, the property "nondecreasing and convex" is a CCC property. This property is useful in the context of comparative statics for the following reason. Consider two functions, $h:\Re \to \Re$ and $g:\Re^n \to \Re$. If $h(x)$ is convex and nondecreasing, and $g(y)$ is supermodular and monotone (nondecreasing *or* nonincreasing), then $h(g(y))$ is supermodular. Thus, if $\int_s \pi(s)dF(s;\theta)$ is nondecreasing and convex in θ for all π in Π, and θ is in fact determined by $\theta = g(y)$ (where g is supermodular and monotone), then $\int_s \pi(s)dF(s;g(y))$ is supermodular in y.

Now we prove a result which builds from Theorem 3.4:

Theorem 4.1 *Suppose property $P \in CCC$. If (Π,Γ) is a stochastic dominance pair, then (Π,Γ) is a stochastic P pair. Equivalently, if $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$, then (Π,Γ) is a stochastic P pair.*

Proof:

$$\textit{For all } \pi \in \Pi, \ \int_s \pi(s)dF(s;\theta) \ \textit{satisfies } P.$$

$$\Rightarrow \qquad \textit{For all } \pi \in \Pi \cup \{1,-1\}, \ \int_s \pi(s)dF(s;\theta) \ \textit{satisfies } P.$$

(Since $\int_s dF(s;\theta) = 1$, and constant functions satisfy P, for all $P \in$CCC).

$$\Rightarrow \qquad \textit{For all } \pi \in cc(\Pi \cup \{1,-1\}), \ \int_s \pi(s)dF(s;\theta) \ \textit{satisfies } P.$$

(Since $\int_s \pi_1(s)\, dF(s;\theta)$ satisfies P and $\int_s \pi_2(s)\, dF(s;\theta)$ satisfies P implies $\int_s [\alpha_1\pi_1(s) + \alpha_2\pi_2(s)]\, dF(s;\theta) = \alpha_1 \int_s \pi_1(s)\, dF(s;\theta) + \alpha_2 \int_s \pi_2(s)\, dF(s;\theta)$ satisfies P for $\alpha_1, \alpha_2 \geq 0$, since P is a CCC property).

$$\Rightarrow \qquad \textit{For all } \pi \in \overline{cc(\Pi \cup \{1,-1\})}, \ \int_s \pi(s)dF(s;\theta) \ \textit{satisfies } P.$$

(Recall that a subset of a topological space is closed if and only if it contains the limits of all the convergent nets of elements in that set.[10] Since the linear functional $\beta(\cdot;\mu)$ is continuous for all μ, for any net π_α in $cc(\Pi \cup \{1,-1\})$ such that $\pi_\alpha \to \pi$, then given θ_0, $\int_S \pi_\alpha(s)\, dF(s;\theta_0) \to \int_S \pi(s)\, dF(s;\theta_0)$ (Kothe, 1969, p.11). Since P is a CCC property, and $\int_S \pi(s) dF(s;\theta)$ is the pointwise limit of a net of functions which satisfy P, then $\int_S \pi(s) dF(s;\theta)$ satisfies P as well.)

$$\Rightarrow \qquad \textit{For all } \gamma \in \overline{cc(\Gamma \cup \{1,-1\})},\ \int_S \gamma(s) dF(s;\theta) \textit{ satisfies } P.$$

(By Theorem 3.6, since (Π,Γ) is a stochastic dominance pair.)

$$\Rightarrow \qquad \textit{For all } \gamma \in \Gamma,\ \int_S \gamma(s) dF(s;\theta) \textit{ satisfies } P.$$

(Since $\Gamma \subseteq \overline{cc(\Gamma \cup \{1,-1\})}$).

Precisely analogous arguments establish the symmetric implication. *Q.E.D.*

This theorem generates many new classes of stochastic P theorems, where the (Π,Γ) pairs which have been identified in the large literature on stochastic dominance are also stochastic P pairs. As a special case, this theorem generalizes a result by Topkis (1968), who proves that the (Π,Γ) pair corresponding to nondecreasing functions and indicator functions of nondecreasing sets, respectively, is a stochastic P pair for P in CCC.

What we have shown in this section is that if P is a CCC property, and Σ_{SPT} denotes the set of all stochastic P pairs, then $\Sigma_{SDT} \subseteq \Sigma_{SPT}$. However, note that *not* all properties P in CCC are such that $\Sigma_{SDT} = \Sigma_{SPT}$. In Remark 2 at then end of Section 4.4, we argue that when the property P is "nondecreasing and convex," then $\Sigma_{SDT} \subset \Sigma_{SPT}$. For now, let us consider a simpler example of a CCC property, the property "constant in θ." Take the case of Π^{FO}, the set of all univariate, nondecreasing payoff functions, and the set $\hat{\Gamma} = -\Pi^{FO}$. The pair $(\Pi^{FO}, \hat{\Gamma})$ satisfies a "stochastic constant theorem," since $\int \pi(s) dF(s;\theta)$ is constant in θ if and only if $-\int \pi(s) dF(s;\theta)$ is constant in θ. However, $(\Pi^{FO}, \hat{\Gamma})$ clearly is not a stochastic dominance pair. The next section identifies a class of properties which *will* satisfy equivalence relationships.

[10]See Kothe (1969, pp. 10-11) for a proof of this statement.

4.3 *Linear Difference Properties*

This section characterizes a subset of CCC properties, which we call Linear Difference Properties (LDPs). These properties are interesting because, as we will show in Section 4.4, for all P in LDP,

$$\Sigma_{SDT} = \Sigma_{SPT}. \tag{4.1}$$

Combining (4.1) with Theorem 3.4, we can conclude that, just as in the case of stochastic dominance, the "closed convex cone" method is exactly the right one for the study of stochastic P theorems when P is an LDP. That is, if $P \in$ LDP, then $(\Pi,\Gamma) \in \Sigma_{SPT}$ if and only if $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$.

Important examples of LDPs are supermodular and concave; others are summarized in Table I. Supermodularity is important because of its role in the analysis of monotone comparative statics predictions. We also emphasize the result about concavity, since concavity and convexity are also frequently encountered in economic contexts; for example, if an objective function is concave, then the First Order Conditions characterize the optimum. Further, concavity can be used to establish the existence of supporting prices in a resource allocation problem.

The result described in equation (4.1) is useful because it may be easier to verify whether or not (Π,Γ) is a stochastic supermodularity pair by drawing from the existing literature on stochastic dominance; by (4.1), checking to see if (Π,Γ) is a stochastic dominance pair answers the same question. For example, if the characteristics of a set Π are determined by an economic problem, and this set Π has been analyzed in the stochastic dominance literature, then the corresponding stochastic supermodularity theorem is immediate. This allows us to bypass the step of checking whether Π and Γ have the same closed convex cone directly (further, most of the stochastic dominance literature does not make such a statement explicitly).

We now begin to build our formal definition of Linear Difference Properties (LDPs). The first important feature of LDPs is that they can be represented in terms of sign restrictions on inequalities involving linear combinations of the function evaluated at different parameter values. For example, $g(\theta)$ is nondecreasing on Θ if and only if $g(\theta^H) - g(\theta^L) \geq 0$ for all $\theta^H \geq \theta^L$ in Θ. This statement specifies a set of inequalities, where each inequality is a difference between the function evaluated at a high parameter value and a low parameter value. To take another example, $g(\theta)$ is supermodular on Θ if and only if $g(\theta^1 \vee \theta^2) - g(\theta^1) + g(\theta^1 \wedge \theta^2) - g(\theta^2) \geq 0$ for all θ^1, θ^2 in Θ. Again, this statement specifies a set of inequalities, this time involving the sum of two differences.

In both cases, we can represent every inequality by the parameter values and the coefficients which are placed on the corresponding function values. For example, for the property nondecreasing, each inequality involves the coefficient vector (1,-1) and a parameter vector of the form (θ^H, θ^L), where 1 is the coefficient on $g(\theta^H)$ and -1 is the coefficient on $g(\theta^L)$. Thus, $g(\theta)$ is nondecreasing on Θ if and only if for all vector pairs $(\alpha,\phi) = ((1,-1),(\theta^H,\theta^L))$ such that $\theta^H \geq \theta^L$, $\sum_{i=1}^{2}\alpha_i \cdot g(\phi_i) = g(\theta^H) - g(\theta^L) \geq 0$.

Likewise, in the case of supermodularity, we are interested in vector pairs of the form $(\alpha,\phi) = ((1,-1,1,-1),(\theta^1 \vee \theta^2,\theta^1,\theta^1 \wedge \theta^2,\theta^2))$. In this case, $g(\theta)$ is supermodular on Θ if and only if, for all such vector pairs, $\sum_{i=1}^{4}\alpha_i \cdot g(\phi_i) = g(\theta^1 \vee \theta^2) - g(\theta^1) + g(\theta^1 \wedge \theta^2) - g(\theta^2) \geq 0$.

Motivated by this discussion, we present the following definition.

Definition 4.4: *A property P has a* ***linear inequality representation*** *if for any parameter space $\Theta_P \in \overline{\Theta}_P$, there exists a positive integer m and a collection of pairs of vectors, $\mathcal{C}_{\Theta_P}$, where $\mathcal{C}_{\Theta_P} \subset \left\{(\alpha,\phi) \middle| \alpha \in \Re^m,\ \phi \in \Theta_P^m\right\}$, so that conditions (i) and (ii) are equivalent for any $g:\Theta_P \to \Re$:*

(i) $g(\theta)$ satisfies P on Θ_P.

(ii) $\sum_{i=1}^{m}\alpha_i \cdot g(\phi_i) \geq 0$ for all $(\alpha,\phi) \in \mathcal{C}_{\Theta_P}$.

The definition of $\mathcal{C}_{\Theta_P}$ builds directly on the above discussion; we will refer to $\mathcal{C}_{\Theta_P}$ as the "linear inequality representation of P on Θ_P." Note first that in the construction of $\mathcal{C}_{\Theta_P}$, we have allowed the coefficients α to vary with the vectors of parameters; this will be useful when we show that multivariate concavity has a linear inequality representation. In the above examples, nondecreasing and supermodular both have linear inequality representations. In the case of nondecreasing, given a parameter space Θ_{ND},

$$\mathcal{C}_{\Theta_{ND}} = \left\{(\alpha,\phi) \middle| \alpha = (1,-1);\ \phi = (\theta^H,\theta^L);\ \theta^H,\theta^L \in \Theta_{ND};\ \theta^H \geq \theta^L\right\}.$$

For the property supermodular, given a parameter space Θ_{SPM}, the appropriate set is

$$\mathcal{C}_{\Theta_{SPM}} = \left\{(\alpha,\phi) \middle| \alpha = (1,-1,1,-1);\ \phi = (\theta^1 \vee \theta^2,\theta^1,\theta^1 \wedge \theta^2,\theta^2);\ \theta^1,\theta^2 \in \Theta_{SPM}\right\}.$$

Note that all properties which have a linear inequality representation are CCC properties, so long as $\sum_{i=1}^{m}\alpha_i = 0$. Observe that in the examples of nondecreasing and supermodular, the components of α sum to zero. It is easy to show that if $\sum_{i=1}^{m}\alpha_i = 0$, then $\sum_{i=1}^{m}\alpha_i \cdot x_i$ can be represented as a linear combination of differences between x_i's. We will require $\sum_{i=1}^{m}\alpha_i = 0$ in our formal definition of an LDP, motivating the word "Difference" in the name.

Proposition 4.2 *If property P has a linear inequality representation and $\sum_{i=1}^{m}\alpha_i = 0$, then $P \in CCC$.*

Proof: Consider $\Theta_P \in \overline{\Theta}_P$. If $g(\theta)$ is constant in θ on Θ_P, then $\sum_{i=1}^{m} \alpha_i \cdot g(\phi_i) = 0$ for all $(\alpha,\phi) \in \mathcal{C}_{\Theta_P}$, and thus $g(\theta)$ satisfies P on Θ_P by Definition 4.4. Now suppose that $g_1(\theta)$ and $g_2(\theta)$ satisfy P on Θ_P. Then for any $(\alpha,\phi) \in \mathcal{C}_{\Theta_P}$ and any $a_1, a_1 > 0$, $\sum_{i=1}^{m} \alpha_i \cdot [a_1 g_1(\phi_i) + a_2 g_2(\phi_i)] = a_1 \sum_{i=1}^{m} \alpha_i \cdot g_1(\phi_i) + a_2 \sum_{i=1}^{m} \alpha_i \cdot g_2(\phi_i)$, which is nonnegative by assumption and by Definition 4.4. This argument can be extended to sequences or nets of functions which satisfy P. *Q.E.D.*

Our properties of interest, LDPs, are a subset of those properties which have linear inequality representations and satisfy $\sum_{i=1}^{m} \alpha_i = 0$ (thus, all LDPs are also CCC properties):

Definition 4.5: *P is an LDP (written $P \in$LDP) if (1) P has a linear inequality representation, and (2) for any parameter space $\Theta_P \in \overline{\Theta}_P$, if we take the linear inequality representation $\mathcal{C}_{\Theta_P}$ of P on Θ_P, then each pair $(\alpha,\phi) \in \mathcal{C}_{\Theta_P}$ satisfies the following two conditions:*

(A) $\sum_{i=1}^{m} \alpha_i = 0$.

(B) If we let $\hat{\hat{\Theta}}_P = \{\phi_1, \phi_2, \ldots, \phi_m\}$, *then given any* $F \in \Delta^n_{\hat{\hat{\Theta}}_P}$, *there exists a* $(G, \hat{\Theta}_P) \in \mathcal{D}^n_P$ *such that* $\hat{\hat{\Theta}}_P \subseteq \hat{\Theta}_P$, *G agrees with F on* $\hat{\hat{\Theta}}_P$, *and (i) and (ii) are equivalent for any* $\pi \in \mathcal{P}^*$:

(i) $\int_s \pi(s) dG(s;\theta)$ *satisfies P on* $\hat{\Theta}_P$.

(ii) $\sum_{i=1}^{m} \alpha_i \int \pi(s) dG(s;\phi_i) \geq 0$.

Now let us interpret condition (B) of Definition 4.4 (we will explain why this requirement is necessary in the following section). Part (B) requires that for each vector pair (α,ϕ), there exists an appropriate parameter space and probability distribution so that the inequality corresponding to (α,ϕ) is critical in determining whether $\int_s \pi(s) dG(s;\theta)$ satisfies P for arbitrary payoff functions π. Part (B) is not trivial because, for some properties P, $\hat{\hat{\Theta}}_P \notin \overline{\Theta}_P$. That is, the property P is not always well-defined on the restricted parameter space which is defined by the components of ϕ.

We illustrate Definition 4.5 with several examples.

Example 1 *Nondecreasing is an LDP.*

Proof: We argued above that nondecreasing has a linear inequality representation and that Part (A) is satisfied. To see (B), pick a $\Theta_{ND} \in \overline{\Theta}_{ND}$ and define $\mathcal{C}_{\Theta_{ND}}$ as above. Then take a vector pair $(\alpha,\phi) = ((1,-1), (\theta^H, \theta^L))$ and define $\hat{\Theta}_{ND} = \{\theta^H, \theta^L\}$. Then, given any parameterized probability distribution $F \in \Delta^n_{\hat{\Theta}_{ND}}$, we can let G=F. Then $\int_s \pi(s) dG(s;\theta)$ is

nondecreasing on $\hat{\Theta}_{ND}$ if and only if $\sum_{i=1}^{m} \alpha_i \int \pi(s) dG(s;\phi_i) =$
$\int \pi(s) dG(s;\theta^H) - \int \pi(s) dG(s;\theta^L) \geq 0$, as required. *Q.E.D.*

Example 2 Supermodular is an LDP.

Proof: We argued above that supermodular has a linear inequality representation and that Part (A) is satisfied. To see (B), pick a $\Theta_{SPM} \in \overline{\Theta}_{SPM}$ and define $\mathcal{C}_{\Theta_{SPM}}$ as above. If we take a vector pair $(\alpha,\phi) = ((1,-1,1,-1),(\theta^1 \vee \theta^2, \theta^1, \theta^1 \wedge \theta^2, \theta^2))$, we can define $\hat{\Theta}_{SPM} =$ $\{\theta^1 \vee \theta^2, \theta^1, \theta^1 \wedge \theta^2, \theta^2\}$. Then, if we take any $F \in \Delta^n_{\hat{\Theta}_{SPM}}$, we can let $G=F$. Then $\int \pi(s) dG(s;\theta)$ is nondecreasing on $\hat{\Theta}_{SPM}$ if and only if

$$\sum_{i=1}^{m} \alpha_i \int \pi(s) dG(s;\phi_i) =$$

$$\int \pi(s) dG(s;\theta^1 \vee \theta^2) - \int \pi(s) dG(s;\theta^1) + \int \pi(s) dG(s;\theta^1 \wedge \theta^2) - \int \pi(s) dG(s;\theta^2) \geq 0. \quad \textit{Q.E.D.}$$

Requirement (B) does rule out certain properties, however. For example, nothing in the definition of a linear inequality representation rules out the following scenario: the same parameter vector ϕ appears twice in $\mathcal{C}_{\Theta_P}$, but with two different parameter vectors, α^1 and α^2. In that scenario, condition (B) (ii), which might only involve α^1 but not α^2, would not be sufficient to check the property P on the parameter space defined by the componenets of ϕ. Take the example "constant." The linear inequality representation of this property is as follows:

$$\mathcal{C}_{\Theta_{CNS}} = \left\{(\alpha,\phi) \middle| \alpha = (1,-1);\ \phi = (\theta^1,\theta^2);\ \theta^1,\theta^2 \in \Theta_{CNS}\right\}.$$

However, there is no way to verify that an arbitrary function is constant using just one inequality. If we let $\hat{\Theta}_{CNS} = \{\theta^1,\theta^2\}$ and take any probability distribution with that parameter space, we cannot verify whether $\int_s \pi(s) dG(s;\theta)$ is constant in θ by checking that $\sum_{i=1}^{m} \alpha_i \int \pi(s) dG(s;\phi_i)$ $=$ $\int_s \pi(s) dG(s;\theta^1) - \int_s \pi(s) dG(s;\theta^2) \geq 0$. The reverse inequality must be checked as well.

We now give several other important examples of Linear Difference Properties.

Example 3: "Concave" is an LDP.

Proof: To see part (A) of the definition of LDP, given any convex set Θ_{CV}, let

$$\mathcal{C}_{\Theta_{CV}} = \left\{(\alpha,\phi) \middle| \alpha = (1,-\lambda,\lambda-1);\ \phi = (\lambda\theta^1 + (1-\lambda)\theta^2, \theta^1, \theta^2);\ \lambda \in (0,1);\ \theta^1,\theta^2 \in \Theta_{CV}\right\}.$$

Then $g(\theta)$ is concave on Θ_{CV} if and only if $g\left(\lambda\theta^1 + (1-\lambda)\theta^2\right) - \lambda g(\theta^1) - (1-\lambda) g(\theta^2) \geq 0$ for all $\lambda \in (0,1)$ and all $\theta^1,\theta^2 \in \Theta_{CV}$, which in turn is true if and only if $\sum_{i=1}^{m} \alpha_i \cdot g(\phi_i) = 0$ for all (α,ϕ) in $\mathcal{C}_{\Theta_{CV}}$.

Now we examine part (B). For any (α,ϕ) in $\mathcal{C}_{\Theta_{CV}}$, let $\hat{\hat{\Theta}}_{CV} = \{\phi_1,\phi_2,\phi_3\}$ and let $\hat{\Theta}_{CV}$ =convex hull($\hat{\hat{\Theta}}_{CV}$). Let $\lambda = -\alpha_2$. Then take any $F \in \Delta^n_{\hat{\hat{\Theta}}_{CV}}$, and define a parameterized probability distribution $G \in \Delta^n_{\hat{\Theta}}$ so that G agrees with F on $\hat{\hat{\Theta}}_{CV}$. Then, for any $\hat{\lambda} \in (\lambda,1)$, let

$G\left(\cdot;\hat{\lambda}\phi_2 + (1-\hat{\lambda})\phi_3\right) \equiv \dfrac{\hat{\lambda}-\lambda}{1-\lambda}F^2 + \dfrac{1-\hat{\lambda}}{1-\lambda}F^1$, and for any $\hat{\lambda} \in (0,\lambda)$, let

$G\left(\cdot;\hat{\lambda}\phi_2 + (1-\hat{\lambda})\phi_3\right) \equiv \dfrac{\lambda-\hat{\lambda}}{\lambda}F^3 + \dfrac{\hat{\lambda}}{\lambda}F^1$. This is a piecewise linear mapping through $G(\cdot;\phi_1)$.

Then, $\int_s \pi(s)dG(s;\theta)$ is concave on $\hat{\Theta}_{CV}$ if and only if

$$\int \pi(s)dG(s;\phi_1) - \lambda\int \pi(s)dG(s;\phi_2) - (1-\lambda)\int \pi(s)dG(s;\phi_3) \geq 0.$$

as required. *Q.E.D.*

We now consider properties which place a sign restriction on an arbitrary partial derivative. First, we treat the discrete generalizations of such properties. To do so, we introduce the following notation: $u^i = (0,..,1,0..,0)$, where the 1 is in the ith component of the vector, and $\Delta^i_\varepsilon f(x) = f(x) - f(x - \varepsilon \cdot u^i)$. Note that, for a sufficiently differentiable f on $\Re^n$, $\frac{\partial^k}{\partial x_1 \ldots \partial x_k} f(x) \geq 0$ everywhere if and only if, for all $\varepsilon_1..,\varepsilon_k > 0$ and all x, $\Delta^1_{\varepsilon_1} \cdot\cdot \Delta^k_{\varepsilon_k} f(x) \geq 0$. Thus, for arbitrary functions f, we will refer to the latter condition as DPk, the discrete generalization of the property $\frac{\partial^k}{\partial x_1 \ldots \partial x_k} f(x) \geq 0$.

Example 4: Any discrete generalization of a sign restriction on one mixed partial derivative is an LDP.

Proof: We have already proved that supermodularity is an LDP, which implies that DP2, the discrete generalization of $\frac{\partial^2}{\partial x_1 \partial x_2} f(x) \geq 0$, is an LDP. We will consider DP3, the discrete generalization of $\frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} f(x) \geq 0$; other mixed partials are analogous. To see part (A) of the definition of LDP, given any $\Theta_{DP3} \in \overline{\Theta}_{DP3}$, let

$$\mathcal{C}_{\Theta_{DP3}} = \left\{ (\alpha,\phi) \left| \begin{array}{l} \alpha = (1,-1,-1,-1,1,1,1,-1); \\ \phi = \begin{pmatrix} \theta,\ \theta-\varepsilon_1 u^1,\ \theta-\varepsilon_2 u^2,\ \theta-\varepsilon_3 u^3, \\ \theta-\varepsilon_1 u^1-\varepsilon_2 u^2,\ \theta-\varepsilon_1 u^1-\varepsilon_3 u^3,\ \theta-\varepsilon_2 u^2-\varepsilon_3 u^3,\ \theta-\sum_{i=1}^3 \varepsilon_i u^i \end{pmatrix}; \\ \varepsilon_i \in \Re_{++};\ \theta \in \Theta_{DP3};\ \phi_i \in \Theta_{DP3},\ i=1,..,8 \end{array} \right. \right\}$$

Then $g(\theta)$ is satisfies DP3 on Θ_{DP3} if and only if $\Delta^1_{\varepsilon_1}\Delta^2_{\varepsilon_2}\Delta^3_{\varepsilon_3} g(\theta) \geq 0$ for all $\varepsilon_i \in \Re_{++}$ and all $\theta \in \Theta_{DP3}$, which in turn is true if and only if $\sum_{i=1}^m \alpha_i \int_s \pi(s)dG(s;\phi_i) \geq 0$ for all (α,ϕ) in $\mathcal{C}_{\Theta_{DP3}}$.

Now we examine part (B). For any (α,ϕ) in $\mathcal{C}_{\Theta_{DP3}}$, let $\hat{\Theta}_{DP3}=\{\phi_1,..,\phi_8\}$. DP3 is well-defined on this set. Then take any $F\in\Delta^n_{\hat{\Theta}_{DP3}}$, and let $G=F$. Then, $\int_s \pi(s)dG(s;\theta)$ satisfies DP3 on $\hat{\Theta}_{DP3}$ if and only if $\Delta^1_{\varepsilon_1}\Delta^2_{\varepsilon_2}\Delta^3_{\varepsilon_3}g(\theta)\geq 0$ for all $\varepsilon_i\in\Re_{++}$ and all $\theta\in\hat{\Theta}_{DP3}$ such that $\phi_i\in\hat{\Theta}_{DP3}$, which reduces to $\sum_{i=1}^{m}\alpha_i\int_s \pi(s)dG(s;\phi_i)\geq 0$, as required. *Q.E.D.*

Example 5: Any sign restriction on one mixed partial derivative is an LDP.

Proof: Let us consider the property $\frac{\partial^3}{\partial x_1\partial x_2\partial x_3}f(x)\geq 0$. Part (A) follows immediately from Example 3, by the definition of DP3. Now, fix (α,ϕ) in $\mathcal{C}_{\Theta_{DP3}}$, and let $\hat{\hat{\Theta}}_{DP3}=\{\phi_1,..,\phi_8\}$. Part (B) requires the construction of a new parameterized probability distribution on the space $\hat{\Theta}_{DP3}$, which we now define to be the simplex generated by $\hat{\hat{\Theta}}_{DP3}$. In other words, $\hat{\Theta}_{DP3}=\left\{\theta\middle|\exists\lambda_1,..,\lambda_8\in[0,1]\ s.t.\ \sum_{i=1}^{8}\lambda_i\ =1 \text{ and } \sum_{i=1}^{8}\lambda_i\phi_i\ =\theta\right\}$. Then, we can construct a continuous vector-valued function $\chi:\hat{\Theta}_{DP3}\rightarrow[0,1]^8$ so that for any $\theta\in\hat{\Theta}_{DP3}$, $\sum_{i=1}^{8}\chi_i(\theta)\cdot\phi_i\ =\theta$. Then take any take any $F\in\Delta^n_{\hat{\hat{\Theta}}_{DP3}}$ and define a parameterized probability distribution $G\in\Delta^n_{\hat{\Theta}_{DP3}}$ so that G agrees with F on $\hat{\hat{\Theta}}_{DP3}$, and further, $G(\cdot;\theta)=\sum_{i=1}^{8}\chi_i(\theta)\cdot G(\cdot;\phi_i)$. Note that this function is indeed a probability distribution. It is then straightforward (albeit tedious) to verify that, for any π, $\int_s \pi(s)dG(s;\theta)$ will then satisfy DP3 on $\hat{\Theta}_{DP3}$ (and thus $\frac{\partial^3}{\partial\theta_1\partial\theta_2\partial\theta_3}\int\pi(s)dG(s;\theta)\geq 0$) if and only if $\sum_{i=1}^{m}\alpha_i\int\pi(s)dG(s;\phi_i)\geq 0$, as required. *Q.E.D.*

It is interesting to note that in general, the intersection of two LDPs is not itself an LDP. For example, the property "nondecreasing *and* concave" is not an LDP. This is because part (B) of the definition of LDP requires that any single vector pair (α,ϕ) must be critical for determining whether the property holds for some parameterized function. However, it is not possible to check "nondecreasing and concave" with a single inequality. This differentiates LDPs from CCC properties, where the combination of any two CCC properties is CCC (refer to Table I for a comparison between the two classes of properties). On the other hand, since LDPs are a subset of CCCs, the intersection of two LDPs is a CCC property; thus, the closed convex cone approach will always be valid for finite combinations of LDPs. Given this discussion, it is interesting to note that supermodularity, which for a suitably differentiable function $f(x)$ can be defined as requiring that $\frac{\partial^2}{\partial x_i\partial x_j}f(x)\geq 0$ for all $i\neq j$, is in fact an LDP. This is true because supermodularity is a property which (i) can be defined on an arbitrary lattice and (ii) when the lattice has four or fewer points, supermodularity of a function on that lattice can be expressed in terms of a single inequality, so that part (B) of the definition of LDP can be satisfied.

The next section proves that the closed convex cone method of proving stochastic *P* theorems is exactly right for LDPs.

4.4 LDPs Satisfy the Necessary and Sufficient Conditions for the Closed Convex Cone Approach

To begin, we recall from Section 3.1.1 that (Π,Γ) is a stochastic dominance pair if and only if the following equality holds (recalling that $\mathcal{Z}^*$ is the set of finite signed measures μ such that $\int d\mu = 0$):

$$\left\{\mu \in \mathcal{Z}^* \middle| \int \pi\, d\mu \geq 0 \quad \forall \pi \in \Pi\right\} = \left\{\mu \in \mathcal{Z}^* \middle| \int \gamma\, d\mu \geq 0 \quad \forall \gamma \in \Gamma\right\} \tag{4.2}$$

This is true because every measure which is a difference between two probability distributions is in $\mathcal{Z}^*$, and further every element of $\mathcal{Z}^*$ has a representation as a (scaled) difference between two probability distributions; thus, $\mathcal{Z}^*$ is the "right" set of measures to check when we are interested in the property "nondecreasing," which involves pairwise comparisons between probability distributions.

Other properties, however, might correspond to sets of measures other than $\mathcal{Z}^*$. Let us now consider how the relationship between Π and Γ changes if we choose a different set of measures. Let $\mathcal{S}^*$ be an arbitrary subset of the set of finite signed measures, $\mathcal{M}^*$. The following equation then describes a relationship between a (Π,Γ) pair which is a variation on the relationship described in (4.2):

$$\left\{\mu \in \mathcal{S}^* \middle| \int \pi\, d\mu \geq 0 \quad \forall \pi \in \Pi\right\} = \left\{\mu \in \mathcal{S}^* \middle| \int \gamma\, d\mu \geq 0 \quad \forall \gamma \in \Gamma\right\} \tag{4.3}$$

If a property *P* has a corresponding set of measures $\mathcal{S}^*$ such that (Π,Γ) are a stochastic *P* pair if and only if (4.3) holds (analogous to the case of the property monotonicity and (4.2)), then $\Sigma_{SDT} = \Sigma_{SPT}$ if and only if (4.2) and (4.3) are equivalent. The following lemma tells us which sets $\mathcal{S}^*$ will be such that (4.2) and (4.3) are equivalent; this in turn can be used to determine which properties *P* are such that $\Sigma_{SDT} = \Sigma_{SPT}$.

Lemma 4.3 *Condition (ii) implies (i), and further, if* $\overline{cc(\mathcal{S}^*)} = -\overline{cc(\mathcal{S}^*)}$, *then (i) implies (ii):*

(i) (4.2) and (4.3) are equivalent for all (Π,Γ) *pairs.*

(ii) $\overline{cc(\mathcal{S}^*)} = \mathcal{Z}^*$.

Proof: Fix a (Π,Γ) pair. First, note that (4.3) is equivalent to the following:

$$\left\{\mu \in \overline{cc(\mathcal{S}^{\bullet})} \middle| \int \pi \, d\mu \geq 0 \quad \forall \pi \in \Pi\right\} = \left\{\mu \in \overline{cc(\mathcal{S}^{\bullet})} \middle| \int \gamma \, d\mu \geq 0 \quad \forall \gamma \in \Gamma\right\} \tag{4.4}$$

That (4.4) implies (4.3) is clear; that (4.3) implies (4.4) uses straightforward linear algebra arguments similar to those of Lemma 3.2.

Next, observe that (4.2) is equivalent to (4.4) under condition (ii); thus, condition (ii) of this lemma implies condition (i).

Now, suppose that (ii) fails and $\overline{cc(\mathcal{S}^{\bullet})} = -\overline{cc(\mathcal{S}^{\bullet})}$. We will show that (4.3) can be true even when (4.2) fails. Without loss of generality, consider a nonzero $v \in \mathfrak{Z}^{\bullet}$ such that $v \notin \overline{cc(\mathcal{S}^{\bullet})}$ (and thus $-v \notin \overline{cc(\mathcal{S}^{\bullet})}$). Note first that this implies (recalling that $c(A)$ denotes cone of A)

$$c(v) \cap \overline{cc(\mathcal{S}^{\bullet})} = \varnothing \text{ and } c(-v) \cap \overline{cc(\mathcal{S}^{\bullet})} = \varnothing. \tag{4.5}$$

Denote the dual cone of v as follows:

$$d(v) \equiv \left\{\pi \in \mathcal{P}^{\bullet} \middle| \int \pi \, dv \geq 0\right\}$$

And denote the dual cone of $d(v)$ as follows:

$$d(d(v)) \equiv \left\{\mu \in \mathcal{M}^{\bullet} \middle| \int \pi \, d\mu \geq 0 \quad \forall \pi \in d(v)\right\}$$

By a corollary to the Hahn-Banach theorem (see Schaefer, 1980, p. 126) $d(d(v)) = \overline{cc(v)} = c(v)$. Let $\Pi = d(v)$ and let $\Gamma = -\Pi$. Observe that $d(-\Pi) = c(-v)$. Then, by (4.5), (4.4) is satisfied for this (Π,Γ) pair because both sets are empty. However, it is clear that $\overline{cc(\Pi \cup \{1,-1\})} \neq \overline{cc(\Gamma \cup \{1,-1\})}$. But then, Theorem 3.3 implies that (4.2) must fail.

Q.E.D.

Thus, if $\overline{cc(\mathcal{S}^{\bullet})} = \mathfrak{Z}^{\bullet}$, then the closed convex cone method is exactly the right one for checking whether (4.3) holds; that is, (4.3) holds for a (Π, Γ) pair if and only if $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$. Further, Lemma 4.3 says that if our set of "test" measures, $\mathcal{S}^{\bullet}$, is such that $\overline{cc(\mathcal{S}^{\bullet})}$ leaves out a measure in $\mathfrak{Z}^{\bullet}$ *and* the negative of that measure, then there will exist a pair of sets of payoff functions so that (4.2) and (4.3) return different answers. Recall that, if a measure separates two sets, the negative of that measure will separate the sets as well.

Lemma 4.3 motivates our definition of what we will refer to as Σ-Properties, that is, the largest set of properties such that $\Sigma_{SDT} = \Sigma_{SPT}$. We are looking for properties P which correspond to sets of measures $\mathcal{S}^{\bullet}_P$ which satisfy Lemma 4.3 (ii), and further are such that (4.3) holds for $\mathcal{S}^{\bullet} = \mathcal{S}^{\bullet}_P$ *if and only if* (Π,Γ) is a stochastic P pair.

Definition 4.6 *A property P is a Σ-Property if for every positive integer n there exists a set of measures* $\mathcal{S}_P^*$ *such that* $\overline{cc(\mathcal{S}^*)} = \mathcal{Z}^*$*, and further statements (A) and (B) are true:*

(A) For any $(F,\Theta) \in \mathcal{D}_P^n$*, there exists a set* $S_P^n \subseteq \mathcal{S}_P^*$ *so that the following two conditions are equivalent for any* $\pi \in \mathcal{P}^*$*:*

(i) $\int \pi \, d\mu \geq 0 \quad \forall \mu \in S_P^n$.

(ii) $\int_s \pi(s) dF(s;\theta)$ *satisfies P on* Θ.

(B) For any $\mu \in \mathcal{S}_P^*$*, there exists a pair* $(G,\hat{\Theta}) \in \mathcal{D}_P^n$ *so that the following two conditions are equivalent for any* $\pi \in \mathcal{P}^*$*:*

(i) $\int \pi \, d\mu \geq 0$.

(ii) $\int_s \pi(s) dG(s;\theta)$ *satisfies P on* $\hat{\Theta}$.

This definition formalizes what it means for a property P to "correspond" to a set of measures $\mathcal{S}_P^*$. There are two parts to this definition. Part (A) guarantees that, given a (Π,Γ) pair, if (4.3) holds, then (Π,Γ) is a stochastic P pair. It says that any time we are given a parameterized distribution F and need to check whether $\int_s \pi(s) dF(s;\theta)$ satisfies P, we can reduce that problem to checking whether $\int \pi \, d\mu \geq 0$ for all μ in some subset of $\mathcal{S}_P^*$. Part (A) guarantees that $\mathcal{S}_P^*$ is "big enough," and we can think of (A) as defining a mapping from $\mathcal{D}_P^n$ to subsets of $\mathcal{S}_P^*$. Part (B) guarantees that, if (Π,Γ) is a stochastic P pair, then (4.3) holds. It requires that for each element of $\mathcal{S}_P^*$, we can find some parameterized probability distribution so that $\int_s \pi(s) dG(s;\theta)$ satisfies P if and only if $\int \pi \, d\mu \geq 0$. This part requires that $\mathcal{S}_P^*$ is not "too big," and we can think of (B) as describing a mapping from $\mathcal{S}_P^*$ to $\mathcal{D}_P^n$.

We now prove a lemma which is used to prove that all LDPs are Σ-properties. This lemma shows that a set of measures which can be written as a linear combination of differences between probability distributions has a cone (and thus closed convex cone) equal to $\mathcal{Z}^*$.

Lemma 4.4 *Given any positive integer m>1 and any nonzero vector of constants* $\alpha \in \Re^m$ *such that* $\sum_{i=1}^m \alpha_i = 0$*, define the following set:*

$$\mathcal{L}_\alpha \equiv \left\{ \mu \middle| \exists F^1, \ldots, F^m \in \Delta^n \text{ s.t. } \mu = \sum_{i=1}^m \alpha_i F^i \right\}$$

Then $\mathcal{Z}^* = c(\mathcal{L}_\alpha)$.

Proof: First, pick $\nu \in c(\mathcal{L}_\alpha)$, and note that probability distributions are finite signed measures. Then, for some $b>0$ and some $F^1,\ldots,F^m \in \Delta^n$,

$$\int d\nu = \int d\left[b \cdot \sum_{i=1}^{m} \alpha_i F^i\right] = b \cdot \sum_{i=1}^{m} \alpha_i \int dF^i = b \cdot \sum_{i=1}^{m} \alpha_i = 0,$$ and we conclude that $\nu \in \mathcal{F}^*$.

Then, pick $\mu \in \mathcal{F}^*$. Let $I^+ = \{i \in 1,..,m | \alpha_i > 0\}$ and let $I^- = \{i \in 1,..,m | \alpha_i < 0\}$. Then define $k = \sum_{i \in I^+} \alpha_i = -\sum_{i \in I^-} \alpha_i$. Now consider the Jordan decomposition of μ, $\mu = \mu^+ - \mu^-$, and define $a = \int d\mu^+ = \int d\mu^-$. Finally, define $F^{i1} = \frac{1}{a}\mu^+$ and $F^{i2} = \frac{1}{a}\mu^-$ for $i = 1,..,m$. Then, let $\nu \equiv \sum_{i=1}^{m} \alpha_i F^i$, and note that $\mu = \frac{a}{k} \cdot \nu \in \mathcal{L}_\alpha$, which in turn implies that $\mu \in c(\mathcal{L}_\alpha)$ since $\frac{a}{k}$ is positive. *Q.E.D.*

Theorem 4.5: If P is an LDP, then P is a Σ-property.

Proof: Let $A = \{\alpha | \exists \Theta_P \in \overline{\Theta}_P \text{ s.t. } (\alpha,\phi) \in \mathcal{C}_{\Theta_P}\}$, and define $\mathcal{S}_P^n = \bigcup_{\alpha \in A} \mathcal{L}_\alpha$ (where $\mathcal{L}_\alpha$ is defined in Lemma 4.4). Then, by Lemma 4.4 and since α is nonzero by the definition of $\mathcal{C}_{\Theta_P}$, $c(\mathcal{S}_P^n) = \bigcup_{\alpha \in A} c(\mathcal{L}_\alpha) = \mathcal{F}^*$.

Part (A): For any $(F, \Theta_P) \in \mathcal{D}_P^n$, let $S_P^n \equiv \{\mu | \mu = \sum_{i=1}^{m} \alpha_i F(\cdot;\phi^i) \text{ for some } (\alpha,\phi) \in \mathcal{C}_{\Theta_P}\}$. Then conditions (i) and (ii) of part (A) are equivalent by the definitions of S_P^n and $\mathcal{C}_{\Theta_P}$.

Part (B): Take a $\mu \in \mathcal{S}_P^*$. By definition, there exist $H^1,\ldots,H^m \in \Delta^n$ and $\Theta_P \in \overline{\Theta}_P$ such that $\mu = \sum_{i=1}^{m} \alpha_i H^i$ for some $(\alpha,\phi) \in \mathcal{C}_{\Theta_P}$. Take this (α,ϕ), and define $\hat{\hat{\Theta}}_P = \{\phi_1,\ldots,\phi_m\}$. Now define a parameterized probability distribution $F \in \Delta^n_{\hat{\hat{\Theta}}_P}$ such that $F(\cdot;\phi_i) = H^i$. Then by part (B) of Definition 4.5, there exists a $(G,\hat{\Theta}_P) \in \mathcal{D}_P^n$ such that $F = G$ on $\hat{\hat{\Theta}}_P$ and, for any $\pi \in \mathcal{P}^*$, $\int_S \pi(s)dG(s;\theta)$ satisfies P on $\hat{\Theta}_P$ if and only if $\sum_{i=1}^{m} \alpha_i \int_S \pi(s)dG(s;\phi_i) \geq 0$. But by the definition of G, and since $\mu = \sum_{i=1}^{m} \alpha_i F(\cdot;\phi_i)$, the latter condition is true if and only if $\int \pi \, d\mu \geq 0$, as required. *Q.E.D.*

We now state the final result of this section, which is that if P is a Σ-property, then the same mathematical structure underlies the stochastic P theorem as a stochastic dominance theorem. It is interesting to note that this theorem can be proved without reference to any of the results in Section 3; there is no topology specified in the result and no discussion of the closed convex cone relationship between Π and Γ.[11] The relationship between the different classes of theorems relies solely on the linear structure of the integral.

Theorem 4.6: *If P is a Σ-Property, then conditions (i) and (ii) are equivalent:*

(i) (Π,Γ) *is a stochastic P pair.*

(ii) (Π,Γ) *is a stochastic dominance pair.*

Proof: Fix (Π,Γ). Recall that we can rewrite condition (i) as follows:

$$\left\{(F,\Theta)\in\mathcal{D}_P^n \,\middle|\, \int_s \pi(s)dF(s;\theta) \text{ satisfies } P \text{ on } \Theta\ \forall\pi\in\Pi\right\}$$

$$= \quad \left\{(F,\Theta)\in\mathcal{D}_P^n \,\middle|\, \int_s \gamma(s)dF(s;\theta) \text{ satisfies } P \text{ on } \Theta\ \forall\gamma\in\Gamma\right\} \tag{4.6}$$

We now argue that (4.6) is equivalent to (4.7):

$$\left\{v\in\mathcal{S}_P^n \,\middle|\, \int\pi\, dv\geq 0\ \forall\pi\in\Pi\right\} = \left\{v\in\mathcal{S}_P^n \,\middle|\, \int\gamma dv\geq 0\ \forall\gamma\in\Gamma\right\} \tag{4.7}$$

First, let us prove that (4.7) implies (4.6). Pick any $(F,\Theta)\in\mathcal{D}_P^n$. Part (A) of the definition of a Σ-property implies that there exists a set $S_P^n\subseteq\mathcal{S}_P^*$ such that $\int_s\pi(s)dF(s;\theta)$ satisfies P on Θ $\forall\pi\in\Pi$ if and only if $\int\pi\, d\mu\geq 0\ \ \forall\mu\in S_P^n,\ \ \forall\pi\in\Pi$. But, by (4.7), the latter statement is true if and only if $\int\gamma d\mu\geq 0\ \ \forall\mu\in S_P^n,\ \ \forall\gamma\in\Gamma$. Finally, applying part (A) of the definition of a Σ-property again, this is equivalent to $\int_s\gamma(s)dF(s;\theta)$ satisfies P on Θ $\forall\gamma\in\Gamma$.

Now, let us argue that (4.6) implies (4.7). Pick any $\mu\in\mathcal{S}_P^*$. Then, by part (B) of the definition of a Σ-property, there exists a $(G,\hat{\Theta}_P)\in\mathcal{D}_P^n$ so that $\int\pi\, d\mu\geq 0\ \ \forall\pi\in\Pi$ if and only if $\int_s\pi(s)dG(s;\theta)$ satisfies P on $\hat{\Theta}$ $\ \forall\pi\in\Pi$. By (4.6), this is equivalent to $\int_s\gamma(s)dF(s;\theta)$ satisfies P on $\hat{\Theta}$ $\forall\gamma\in\Gamma$. But, applying part (B) of the definition of a Σ-property again, this is equivalent to $\int\gamma d\mu\geq 0\ \ \forall\gamma\in\Gamma$.

Finally, we can apply Lemma 4.3 to conclude that (4.7) is equivalent to:

$$\left\{\mu\in\mathfrak{Z}^* \,\middle|\, \int\pi\, d\mu\geq 0\ \ \forall\pi\in\Pi\right\} = \left\{\mu\in\mathfrak{Z}^* \,\middle|\, \int\gamma d\mu\geq 0\ \ \forall\gamma\in\Gamma\right\} \tag{4.8}$$

Recall from Section 3.1.1 that (4.8) is equivalent to the statement that (Π,Γ) is a stochastic dominance pair. *Q.E.D.*

---

[11] Although we refer to Lemma 4.3 in the proof of Theorem 4.6, we only use (ii) implies (i) of that Lemma, which does not rely on the choice of topology.

The consequence of Theorem 4.6 can be stated as follows: if P is a Σ-property (and thus if P is an LDP), then $\Sigma_{SPT} = \Sigma_{SDT}$. This result together with Theorem 3.4 establishes that the closed convex approach to stochastic P theorems exactly the right one for this class of properties: $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$, *if and only if* $\int_S \pi(s)dF(s;\theta)$ satisfies P $\forall \pi \in \Pi$ exactly when $\int_S \gamma(s)dF(s;\theta)$ satisfies P $\forall \gamma \in \Gamma$. This result then generates entirely new classes of theorems, only a few of which have appeared in the economics literature to date.

Note that Theorem 4.6 holds without any assumptions about continuity, differentiability, or other such properties. Even the restrictions on the boundedness of the payoff functions have been dropped.[12] Assumptions on sets of payoff functions might be relevant for a particular stochastic dominance pair (Π,Γ); these assumptions would then be inherited by the corresponding stochastic supermodularity theorem. However, the relationship between stochastic dominance and stochastic supermodularity is defined exactly by (4.1) without any restrictions beyond the structure imposed by Definitions 3.1 and 4.1. Linearity of the functional $\beta(\pi,\mu)$ in μ, however, is critical.

In the next section, we discuss applications of some of the new stochastic supermodularity theorems which can be derived using Theorem 4.6.

Remark 2

We argued in Section 4.3 that combinations of LDPs are not themselves LDPs (or Σ-Properties). In this remark we use an example which we call *NDC*, the univariate, discrete version of "nondecreasing and convex," to illustrate what goes wrong when properties are combined. The minimal set on which *NDC* will be considered well-defined is a 3-point subset of the real line (otherwise convexity is not well-defined).

So, what we want to show is that the fact that (Π,Γ) is a stochastic *NDC* pair does not necessarily imply that $\overline{cc(\Pi \cup \{1,-1\})} = \overline{cc(\Gamma \cup \{1,-1\})}$.

The linear inequality representation of *NDC* is as follows:

$$\mathcal{C}_{\Theta_{NDC}} = \left\{(\alpha,\phi) \middle| \alpha = (\lambda, 1-\lambda, -1);\ \phi = (\theta^1, \theta^2, \lambda\theta^1 + (1-\lambda)\theta^2);\ \lambda \in (0,1);\ \theta^1, \theta^2 \in \Theta_{NDC}\right\}$$
$$\cup \left\{(\alpha,\phi) \middle| \alpha = (1,-1,0);\ \phi = (\theta^1, \theta^2, \theta^3);\ \lambda \in (0,1);\ \theta^1 \geq \theta^2;\ \theta^1, \theta^2 \in \Theta_{NDC}\right\}$$

Consider the case where the parameter space is a 3-point subset of the real line. The main problem is that checking that a function is nondecreasing and convex on that space requires three inequalities. Given that, the intuition about why *NDC* fails to be a Σ-Property as follows:

[12] We must, however, be careful so that the probability distributions and payoff functions are chosen to make the expected value of the payoff function well-defined.

suppose $\overline{cc(\Pi \cup \{1,-1\})} \neq \overline{cc(\Gamma \cup \{1,-1\})}$. This inequality can hold if there is only one point in $\overline{cc(\Pi \cup \{1,-1\})}$ but outside $\overline{cc(\Gamma \cup \{1,-1\})}$, and in that case the Hahn-Banach theorem guarantees the existence of only *one* measure which is a separating hyperplane. However, generating a parameterized distribution which contradicts the hypothesis that (Π,Γ) is a stochastic *NDC* pair requires three distributions corresponding to the three points in the parameter space; these three distributions will in turn induce *three* measures, as specified in the linear inequality representation of $\mathcal{C}_{\Theta_{NDC}}$.

4.5 Applications of Stochastic Supermodularity Theorems

In this section, we explore applications of the result that $\Sigma_{SDT} = \Sigma_{SST}$, that is, the result that the set of stochastic dominance pairs is the same as the set of stochastic supermodularity pairs. We look at stochastic dominance pairs from the existing literature and interpret the new stochastic supermodularity theorems which can be derived from these. We show how these new theorems can be used to derive sufficient, or necessary and sufficient, conditions for monotone comparative statics predictions in several economic examples.

4.5.1 Examples of Univariate Stochastic Supermodularity Theorems

We have already discussed univariate stochastic dominance theorems at some length. Using Theorem 4.6, we can apply the results of Table II to problems of stochastic supermodularity. Recall that the example of the agent's choice of effort, discussed in Section 2, uses the stochastic supermodularity theorem corresponding to Table II (iii); the equivalence of parts (i) and (ii) in Proposition 2.3 follows as a direct corollary of Theorem 4.6. We can analyze other properties P in an analogous way.

Let us now consider a second example. Suppose that a firm chooses to invest in research and development (denoted r) to improve its production process, and in particular it searches for ways to reduce its unit production costs (denoted $c \in \Re_+$). The returns to research and development are inherently uncertain, and the probability distribution over the firm's future production costs is parameterized by t. Suppose that the firm's payoffs are nonincreasing in its production costs, and that the firm's investment in r has a cost, $k(r)$. Then the firm's expected profits can be written as follows:

$$\int_c \pi(c)dF(c;r,t) - k(r)$$

Now we address the comparative statics question: What are necessary and sufficient conditions for the optimal investment in research to be monotone nondecreasing in the shift parameter t, for all investment cost functions k?

Proposition 4.7 *The following three conditions are equivalent:*

(MCS) $r^*(t) \equiv \arg\max_r \int_c \pi(c)dF(c;r,t) - k(r)$ *is monotone nondecreasing in t for all cost functions k and all* π *nonincreasing.*

(i) $\int_c \pi(c)dF(c;r,t)$ *is supermodular in (r,t) for all* π *nonincreasing.*

(ii) For all $c \in \Re_+$, $F(c;r,t)$ *is supermodular in (r,t).*

Proof: Theorem A.1 in the Appendix establishes that (MCS) is equivalent to (i). Theorem 4.6, together with Table II (i), establishes that (i) holds if and only if $-\int I_{[b,\infty)}(c)dF(c;r,t) = -1 + F(b;r,t)$ is supermodular in (r,t) for all $b \in \Re_+$. *Q.E.D.*

Thus, for any parameter t which is complementary with r in terms of increasing the probability distribution pointwise, increasing t will lead to an increase in the optimal choice of r. Further, *no other class of shifts in the probability distribution will always increase the firm's choice of r.* Intuitively, the goal of the firm's investment in research is to shift probability weight towards lower realizations of its unit cost; the parameter t measures the "sensitivity" of the probability distribution to investments in research. Higher values of t correspond to probability distributions where research is more effective at lowering unit costs.

4.5.2 Examples of Bivariate Stochastic Supermodularity Theorems

The stochastic dominance theorems for bivariate payoff functions are perhaps less familiar, but they also fit in the framework of Theorem 4.6; in the existing literature, the proofs of the continuous versions of these theorems use integration by parts. Levy and Paroush (1974) and Atkinson and Bourguignon (1982) first reported these results for continuous objective functions; Meyer (1990) extends some of their results to discrete problems. Table III summarizes the stochastic dominance results. By Theorem 4.6, each of the (existing) stochastic dominance results corresponds to a (new) stochastic supermodularity result, which may then be applied to solve comparative statics problems.

	Sets of payoff functions, Π	Sets of payoff functions, Γ
(i)	$\{\pi \mid \pi : \Re^2 \to \Re, \text{ nondecreasing}\}$	$\left\{\gamma \middle\vert \begin{array}{l} \gamma(s_1,s_2) = I_A(s_1,s_2), \text{ where } A \subseteq \Re^2 \\ \text{and } I_A(s_1,s_2) \text{ nondecreasing} \end{array}\right\}$
(ii)	$\left\{\pi \middle\vert \begin{array}{l} \pi : \Re^2 \to \Re, \text{ nondecreasing,} \\ \text{supermodular} \end{array}\right\}$	$\{\gamma \mid \gamma(s_1,s_2) = I_{[a_1,\infty)}(s_1) \cdot I_{[a_2,\infty)}(s_2),\ a_1,a_2 \in \overline{\Re}\}$
(iii)	$\{\pi \mid \pi : \Re^2 \to \Re, \text{ supermodular}\}$	$\{\gamma \mid \gamma(s_1,s_2) = I_{[a_1,\infty)}(s_1) \cdot I_{[a_2,\infty)}(s_2),\ a_1,a_2 \in \overline{\Re}\}$ $\cup \{\gamma \mid \gamma(s_1,s_2) = -I_{[a_1,\infty)}(s_1),\ a_1 \in \Re\}$ $\cup \{\gamma \mid \gamma(s_1,s_2) = -I_{[a_2,\infty)}(s_2),\ a_2 \in \Re\}$

Table III Bivariate Stochastic Dominance/ Supermodularity Theorems

Each (Π,Γ) pair in Table III satisfies a (bivariate) stochastic dominance theorem as well as a (bivariate) stochastic supermodularity theorem.

Table III (i) is simply a bivariate generalization of FOSD (this result also generalizes to the multivariate case). We will discuss the results associated with Table III (iii) carefully, and the intuition for Table III (ii) is similar. First, we will interpret the stochastic dominance theorem associated with Table III (iii); then, we will consider the stochastic supermodularity theorem.

The set Γ in Table III (iii) contains both the indicator functions for each upper interval of the random variable s_i, and the negative of the indicator functions for each upper interval. Note that $\iint_{s_1,s_2} I_{[a_1,\infty)}(s_1) dF(s_1,s_2) = F_1(a_1)$, where $F_1(s_1)$ is the marginal distribution of s_1. Since we have not specified whether the payoff function π is monotonic, it is not possible to verify whether changing the marginal distribution will raise or lower expected profits. Thus, if a shift in the probability distribution must raise expected payoffs *for all* supermodular payoff functions, that shift in the distribution must not affect the marginal distributions (i.e., $F_1(a_1;\theta)$ must be constant in θ for all a_1).

Now, consider a partition of the space $(s_1,s_2) \in \Re^2$ into four quadrants, delineated by the axes $s_1 = a_1$ and $s_2 = a_2$. Then the requirement that the function $\iint_{s_1,s_2} I_{[a_1,\infty)}(s_1) \cdot I_{[a_2,\infty)}(s_2) \cdot dF(s_1,s_2;\theta)$ must be nondecreasing in θ specifies that θ shifts probability mass into the northeast quadrant. We will refer to this type of shift as an increase in the "interdependence" of the random variables; such a shift is beneficial when the random variables are complementary in increasing the payoff function.

The intuition for the stochastic supermodularity theorem corresponding to Table III (iii) is similar: for two parameters to be complementary in increasing the expected value of a supermodular payoff function, they must not interact in the marginal distribution functions, and further they must be complementary in increasing the interdependence of the random variables.

To see a special case of how Table III (ii) might be used, consider the following example. Suppose that we are interested in the set of bivariate, supermodular payoff functions. Suppose further that the random variables have a bivariate normal distribution with a positive covariance ($(s_1,s_2) \sim BVN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\sigma_{12});\ \sigma_{12} \geq 0$). Then checking the conditions given in Table III (ii) establishes that the expected profits are supermodular in the means of each marginal distribution. That is, $\int_{s_1,s_2} \pi(s_1,s_2)dF(s_1,s_2;\mu_1,\mu_2)$ is supermodular in (μ_1,μ_2), for all π supermodular, if $\sigma_{12} \geq 0$. Intuitively, when the payoff function is such that under certainty, one random variable increases the returns to the other, then in a stochastic environment, raising the mean of one random variable increases the returns to raising the mean of the other. Further, since the random variables have a positive covariance, increasing the mean of one does not decrease the effectiveness of the other in terms of shifting probability weight into regions where the random variables realize high or low values together. The requirement that the parameters do not interact in the marginal distribution functions is satisfied trivially.

Let us apply this to an economic problem, in particular a coordination problem within a firm. Suppose that the normally distributed random variables, s_1 and s_2, represent the qualities of two different components of a final product. The components fit together in such a way that increasing the quality of one component increases the returns to quality in the other component. (For example, the fidelity of the amplifier in a stereo system is complementary with the fidelity of the tuner). Suppose that two product design teams work to develop the two different components. The qualities which each team will be able to achieve is stochastic, due to the inherent uncertainty in innovative activity, but the outputs of the two teams are correlated (perhaps due to realizations of random events which affect the whole firm, or due to communication between the two teams).

Now consider the firm's problem of setting target qualities (or incentive contracts) for each group, where an increase in the target quality increases the expected quality the group will produce. The above result states that increasing the target quality for one group increases the returns to increasing the target quality of the second group. Thus, if there is an exogenous decrease in the cost of producing quality for team one, then the firm will find it optimal to raise the targets for *both* teams.

4.5.3 Examples of Multivariate Stochastic Supermodularity Theorems

Multivariate payoff functions pose a particularly complicated problem because changes in the joint probability distribution can potentially affect the co-movements of many random variables simultaneously; thus, the high-order mixed partial derivatives between all of the arguments of the payoff function are relevant for stochastic dominance results. However, it is often difficult to place economic interpretations on such derivatives; thus, when analyzing multivariate problems, it is useful to impose additional structure. First, we discuss the case with multivariate statistical dependencies, but we restrict the interactions between variables in the payoff function.[13] Then, we discuss the case where multivariate interactions are permitted in the payoff function, but the random variables are constrained to be statistically independent.

To begin, following Meyer's (1990) work on stochastic dominance, we examine *pairwise separable* payoff functions. We define pairwise separable functions as functions which can be written $\pi(s_1, s_2, ..., s_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \pi^{ij}(s_i, s_j)$. It turns out that the results reported in the bivariate case generalize to this set of functions, since (using the linearity of the integral operator) we can write the expected value of a pairwise separable payoff function as the sum of the expected values of bivariate payoff functions. If each bivariate problem is supermodular (monotonic), then the sum of them will be supermodular (or monotonic) as well.

There are potential applications of these theorems in welfare economics; for example, Meyer and Mookherjee (1987) and Meyer (1990) discuss applications of bivariate and pairwise separable stochastic dominance results to welfare economics. Consider a social planner who values equity between agents in the economy. Meyer and Mookherjee (1987) and Meyer (1990) argue that this can be represented as a social welfare function, $W(s_1, ..., s_n)$, which is supermodular in $(s_1, ..., s_n)$, the vector of individual agents' incomes; Meyer (1990) proposes the simplifying assumption that $W(s_1, ..., s_n)$ be pairwise separable. These authors then ask the question, What types of shifts in the probability distribution over consumer incomes will improve expected welfare of the social planner? In this context, Table III (ii) shows that a shift in the distribution of income which holds the marginal distributions fixed, but increases "interdependence" between the random variables, increases expected welfare.

By Theorem 4.6, we can build from these results to perform comparative statics analysis, addressing the following questions: When will two social policies which affect the joint distribution of agents' income be complementary? For which exogenous shifts in the distribution of income will the social planner tend to increase her use of a particular policy? If the social welfare function is also monotone in each agent's income, the answer to these questions will be determined by checking supermodularity conditions on the joint distribution of each pair of random variables, as shown in Table III (ii). Thus, two policies which are complementary in (1) increasing the interdependence of

[13] In the next subsection, we will take the opposite approach, eliminating the multivariate statistical dependencies while allowing for multivariate payoff functions.

utilities in the economy and (2) improving each agent's marginal distribution of income, will be complementary in increasing expected welfare. (The case where the two policies do not interact in determining the marginal distribution of each agent's income is clearly a special case of the second restriction).

In the previous example, we restricted the multivariate interactions in the payoff function in order to analyze problems where there are statistical dependencies between several random variables. Now, we turn to show that when a subset of the random variables are statistically independent from the rest, it often possible to allow for more complex interactions in the payoff function.

When one or two random variables are independent of the others, we can apply the univariate and bivariate stochastic dominance/supermodularity theorems described above. To see an example of this, consider rewriting a stochastic objective function as follows

$$\int_s \pi(s) dF_{ij}(s_i, s_j; \theta) dF_{n\setminus ij}(s_{n\setminus ij})$$
$$= \int_{s_{n\setminus ij}} \left[\int_{s_i, s_j} \pi(s_i, s_j; s_{n\setminus ij}) dF_{ij}(s_i, s_j; \theta) \right] dF_{n\setminus ij}(s_{n\setminus ij})$$

Note that monotonicity as well as supermodularity are preserved by integration, so that supermodularity in θ of the inner integral in the above expression will guarantee that the objective function is supermodular in θ.

In the following example, we use this technique to derive a useful result for multivariate, supermodular payoff functions and independent random variables.

Theorem 4.8 *Let s be a vector of independent random variables. Further, suppose that θ is a vector of parameters such that for i=1,...,n, the marginal distribution of s_i is given by $F_i(s_i;\theta_i)$. Then the following two conditions are equivalent:*

(i) For all supermodular payoff functions $\pi : \Re^n \to \Re$, $\int_s \pi(s) dF(s;\theta)$, is supermodular in θ.

(ii) Either (a) or (b) is true:

(a) For $i = 1,..,n$, $F_i(a_i;\theta_i^H) \geq F_i(a_i;\theta_i^L) \;\; \forall \theta_i^H \geq \theta_i^L, \;\; \forall a_i \in \Re$.

(b) For $i = 1,..,n$, $F_i(a_i;\theta_i^H) \leq F_i(a_i;\theta_i^L) \;\; \forall \theta_i^H \geq \theta_i^L, \;\; \forall a_i \in \Re$.

Proof: For all (i,j) pairs, rewrite the expectation as

$$\int_{s_i, s_j} \left[\int_{s_{n\setminus ij}} \pi(s_{n\setminus ij}; s_i, s_j) dF_{n\setminus ij}(s_{n\setminus ij}; \theta_{n\setminus ij}) \right] \cdot dF(s_i;\theta_i) \cdot dF(s_j;\theta_j).$$

Note that the inner integral is supermodular in (s_i, s_j). Apply Table III (ii) and check that the supermodularity condition on the distributions reduces to (ii) (a) or (b).

This result says that parameters which induce FOSD shifts in independent random variables are complementary in increasing the expected value of supermodular payoff functions.[14] Thus, if a payoff function is supermodular in a group of random variables, increasing one variable in the stochastic sense (of FOSD) is complementary with increasing the others stochastically as well. This result has potential applications in the study of coordination problems in firms (recall the product design example from above) as well as in general investment problems. Athey and Schmutzler (1995) apply this result to analyze a firm's choices over investments in product design and process innovation.

5 SUPERMODULARITY OF $\int_s \pi(x,s) \cdot dF(s;\theta)$

This section studies the relationships between the parameters which enter the payoff function (x) and the parameters which shift the probability distribution (θ). To begin, in Section 5.1 we will consider interactions between the components of the vector x; in Section 5.2, we will study supermodularity of the objective function in (x,θ).

5.1 When is $\int_s \pi(x,s) \cdot dF(s)$ supermodular in x?

As noted in Section 2, arbitrary sums of supermodular functions are supermodular. Thus, if $h: Y \times S \to \Re$ is supermodular in x, then $h(y;s^1) + h(y;s^2)$ is supermodular in y. Using this fact, we can show that the integral (over a subset of the arguments) of any supermodular function is supermodular in the remaining arguments. So, if $\pi(x,s)$ is supermodular in x for all s, then $\int_s \pi(x,s) \cdot dF(s)$ is supermodular in x as well.[15] Further, there is no weaker property of $\pi(x,s)$ which guarantees supermodularity of the objective function for all distribution functions. To see this, suppose that $\pi(x,s)$ is not supermodular in x for a particular vector s^0. Then, consider the distribution function which places all of the probability weight on s^0: clearly the expected value of $\pi(x,s)$ will not be supermodular in x in that case.

[14]Since a supermodular function is also supermodular in the negative of all of its arguments, we allow for either case (ii)(a) or (ii)(b) in Theorem 4.6.

[15]Note that weighting the payoff function by a nonnegative function which does not depend on x does not disturb the supermodularity in x.

5.2 When is $\int_s \pi(x,s) \cdot dF(s;\theta)$ supermodular in (x,θ)?

We have already analyzed interactions between the components of x and between the components of θ. Since supermodularity can be checked pairwise, we focus on an arbitrary pair (x,θ), suppressing all other parameters in our notation. In turns out that finding the relevant conditions for supermodularity in this case is especially straightforward. For reference, we state the following simple lemma.

Lemma 5.1 *Let* $\Delta\pi(s) = \pi(x^H,s) - \pi(x^L,s)$. *Then* $\int_s \pi(x,s) \cdot dF(s;\theta)$ *is supermodular in* (x,θ) *if and only if* $\int_s \Delta\pi(s) \cdot dF(s;\theta)$ *is monotone nondecreasing in* θ *for all* $x^H \geq x^L$.

In this way, we rephrase a problem of supermodularity as a problem of monotonicity, and we can exploit the existing body of stochastic dominance and other stochastic monotonicity results.

To show how this can be used, we present the following theorem, which is based on a stochastic monotonicity result.

Theorem 5.2 *Let s be a vector of independent random variables. Then the following two conditions are equivalent:*

(i) For all $\pi : \Re \times \Re^n \to \Re$ *such that* π *is supermodular in* (x,s_i) *for* $i = 1,\ldots,n$, $\int_s \pi(x,s) \cdot dF(s;\theta)$, *is supermodular in* (x,θ).

(ii) For $i = 1,\ldots,n$, $F_i(a_i;\theta^H) \leq F_i(a_i;\theta^L)$ $\forall \theta^H \geq \theta^L$, $a_i \in \Re$

Proof: Pick $x^H \geq x^L$. Using Lemma 5.1, if n=1, then this theorem is equivalent to a FOSD theorem. Now suppose that the theorem holds for $n = m$ and consider $n = m + 1$. In the continuous case, we can write

$$\frac{\partial}{\partial\theta}\int \Delta\pi(s) dF(s;\theta)$$

$$= \frac{\partial}{\partial\theta}\int_{s_{m+1}}\int_{s_{1\ldots m}} \Delta\pi(s) \cdot dF_{1,\ldots,m}(s_{1,\ldots,m};\theta) \cdot dF_{m+1}(s_{m+1};\theta)$$

$$= \int_{s_{m+1}} \frac{\partial}{\partial\theta}\left[\int_{s_{1\ldots m}} \Delta\pi(s) \cdot dF_{1,\ldots,m}(s_{1,\ldots,m};\theta)\right] \cdot dF_{m+1}(s_{m+1};\theta)$$

$$- \int_{s_{m+1}}\int_{s_{1\ldots m}} \frac{\partial}{\partial s_{m+1}}\Delta\pi(s) \cdot dF_{1,\ldots,m}(s_{1,\ldots,m};\theta) \cdot \frac{\partial}{\partial\theta}F_{m+1}(s_{m+1};\theta) \cdot ds_{m+1}$$

Using a similar approach to the proof of FOSD, we can establish that under the inductive hypothesis, this expression is positive for all nondecreasing $\Delta\pi(s)$ if and only if $\frac{\partial}{\partial\theta}F_{m+1}(s_{m+1};\theta) \leq 0$. (In particular, $\Delta\pi(s)$ could be constant in the first m arguments, or be non-zero only at a point where $\frac{\partial}{\partial\theta}F_{m+1}(s_{m+1};\theta) > 0$). Thus, by induction, the theorem holds.

This result shows that, when maximizing the expected value of a supermodular function, increasing a parameter of the payoff function is complementary with a FOSD increase in all of the random variables.[16]

6 CONCLUSIONS

In this paper, we have created a framework for analyzing stochastic dominance theorems. The existing literature proves stochastic dominance theorems in an ad hoc manner, either making use of assumptions about the differentiability of payoff functions or else resorting to discrete distributions. This paper formally proves that these assumptions are unrelated to stochastic dominance conclusions. Using linear functional analysis, we prove a result which underlies all stochastic dominance theorems, and further generalizes to conditional stochastic dominance theorems as well.

Once we have developed the unifying framework for stochastic dominance, we then draw analogies between theorems about stochastic dominance and what we call stochastic P theorems, which are theorems that characterize when a stochastic objective function satisfies a given property P. We show that when P is a "Linear Difference Property," the same mathematical structure underlies stochastic P theorems as stochastic dominance theorems. Examples of Linear Difference Properties include supermodularity, concavity, and properties which place a sign restriction on a mixed partial derivative (see Table I). Thus, this paper generates entire new classes of theorems, providing a complete characterization of a variety of properties in stochastic optimization problems.

Our result about Linear Difference Properties is important because it not only tells us where to look for results about stochastic supermodularity and concavity (that is, the stochastic dominance literature), but it also tells us that there is no other place to look. Thus, if we can show that there is no general stochastic dominance result, we will not be able to find a general stochastic supermodularity or concavity result. In a precise way, Theorem 4.6 says that the stochastic supermodularity and concavity problem has already been solved, to the extent that the stochastic dominance problem has been solved.

The theorems presented here have a structure which will be useful in evaluating robustness in applications. In particular, the theorems give necessary and sufficient conditions on probability distributions for a property to hold for all payoff functions in a given set; thus, we can verify robustness of comparative statics results across the relevant set of payoff functions. This contribution is especially useful in an area of inquiry where many economic theories currently rely on functional forms. For example, the methods of this paper allow us to systematically verify whether monotone

[16]Hadar and Russell (1978) and Ormiston and Schlee (1992) show that this is true for the case with only one random variable.

comparative statics derived in models with specific functional forms and probability distributions are robust. Further, applying these methods to particular economic models help us pinpoint the critical properties which must hold for a given monotone comparative statics hypothesis to be true.

Finally, we note a caveat to our analysis: because the techniques in this paper exploited the linearity of the integral operator, we examined only properties and classes of payoff functions which are closed convex cones. However, some properties we encounter frequently in economics are not preserved by positive combinations and sums. Athey (1995) characterizes one such property, the single crossing property, in stochastic optimization problems.

APPENDIX

To begin, we present some basic definitions from lattice theory. The operations "meet" (denoted $\wedge$) and "join" (denoted $\vee$) are defined as follows: for $x, x' \in X$,

$$x \vee x' \equiv \inf\{z | z \geq x \text{ and } z \geq x'\}$$
$$x \wedge x' \equiv \sup\{z | z \leq x \text{ and } z \leq x'\}$$

where the operations of supremum and infimum are defined using a given order. For the case of $\Re^n$ with the usual order, join is the component-wise maximum, and meet is the component-wise minimum. A *lattice* consists of space and a partial order where the meet and join always exist. We will often be interested in subsets of lattices which have a special structure:

Definition A.1 *A set $K \subseteq X$ is a **sublattice** of a lattice X if $x, x' \in K$ implies that $x \vee x' \in K$ and $x \wedge x' \in K$.*

For example, any set $[a_1, a_2] \times [b_1, b_2]$ is a sublattice of $\Re^2$.

In the study of monotone comparative statics, it is useful to be able to compare sets, such as constraint sets or sets of maximizers of a function. The following definition provides a partial order over sets.

Definition A.2 *A set $A \subseteq X$ is higher than a set $B \subseteq X$ in Veinott's **strong set order**, written $A \geq B$, if for all $x \in A$ and $y \in B$, $x \vee y \in A$ and $x \wedge y \in B$.*

If $A = [a_1, a_2]$ and $B = [b_1, b_2]$ are intervals of the real line, then $A \geq B$ implies that $a_1 \geq b_1$ and $a_2 \geq b_2$. In general, $A \geq B$ implies that the lowest element of *A* is higher than the lowest element of *B*, and the highest element of *A* is higher than the highest element of *B*. The order is not reflexive in general; $A \geq A$ if and only if *A* is a sublattice.

We now formally define supermodularity:

Definition A.3 *A function $h: X \to \Re$ is **supermodular** if for all $x, x' \in X$, $h(x) + h(x') \leq h(x \vee x') + h(x \wedge x')$.*

It turns out (Topkis, 1978) that a function is supermodular if and only if its arguments are pairwise *complementary* in the sense that increasing one increases the returns to increasing the other. For a twice differentiable function, this reduces to nonnegative cross-partial derivatives between each pair of variables. Supermodularity has many useful properties; for example, arbitrary sums of supermodular functions are supermodular.[17]

---

[17]Another useful property is that the maximized value (with respect to a subset of the variables) of a supermodular function is supermodular. This property makes supermodularity easy to work with two-stage optimization problems.

We now report a result which illustrate the usefulness of supermodularity in the study of monotone comparative statics.

Theorem A.1 *(Milgrom and Shannon, 1994) Conditions (i) and (ii) are equivalent.*

(i) The function $h(x,t)$ is supermodular.

(ii) For all functions $g^i : X^i \to \Re$, the set of x-maximizers $x^*(t,K) = \arg\max_{x \in K}\left[h(x,t) + g^1(x_1) + \cdots + g^n(x_n)\right]$ is monotone nondecreasing in (t,K).

Further, if $X \subseteq \Re$, then (i) is equivalent to (iii):

(iii) For all functions $g : X \to \Re$, the set of x-maximizers $x^*(t) = \arg\max_{x \in X}\left[h(x,t) + g(x)\right]$ is monotone nondecreasing in t.

Theorem A.1 states that if h is supermodular, then the comparative statics conclusion (ii) holds; further, if we would like to guarantee that the comparative statics prediction holds for all additively separable cost/benefit functions g^i, then h MUST be supermodular.

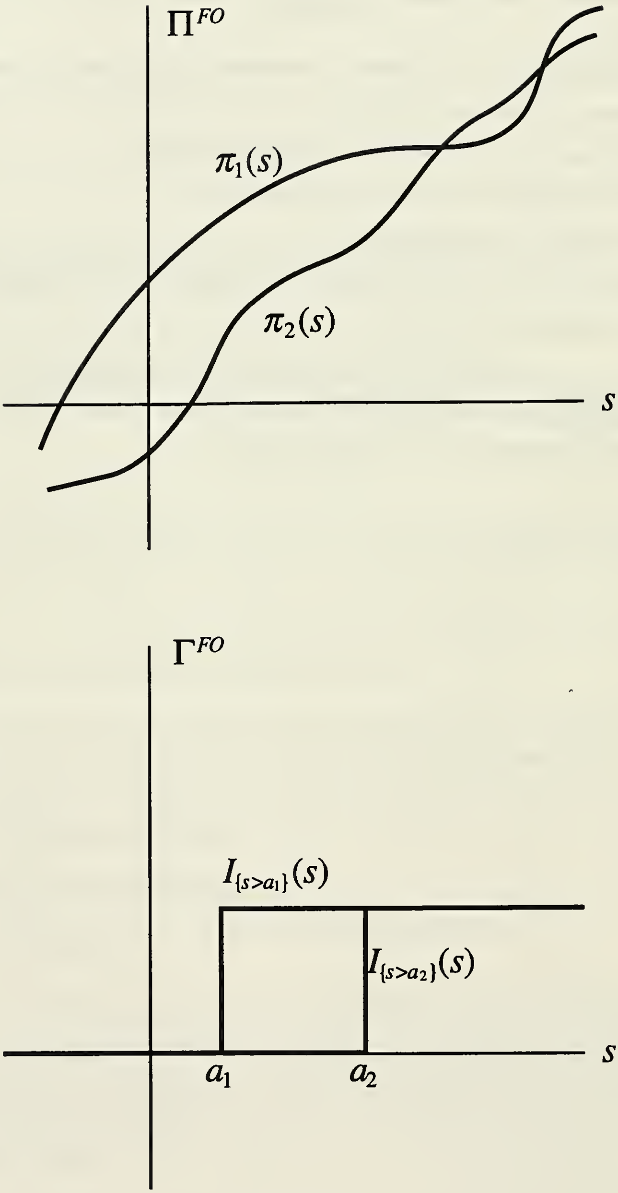

Figure 1

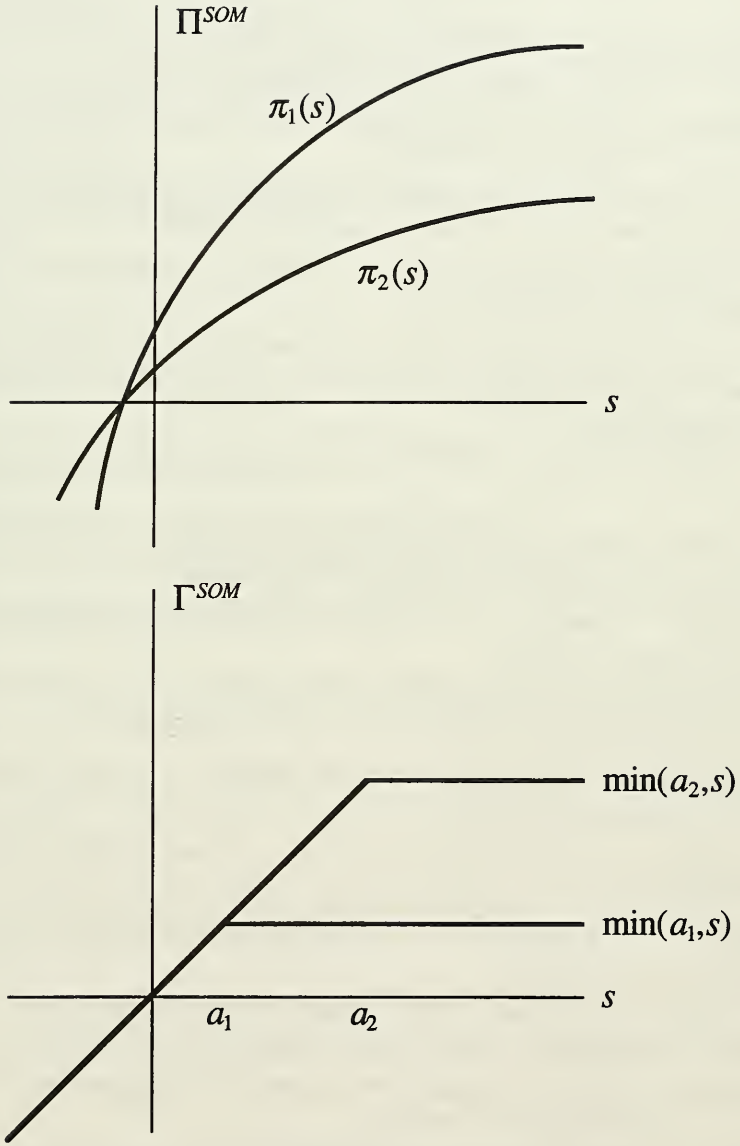

Figure 2

REFERENCES

Athey, Susan (1995), "Single Crossing and the Monotone Likelihood Ratio Order in Stochastic Optimization Problems." Mimeo, Massachusetts Institute of Technology.

Athey, Susan, and Armin Schmutzler (forthcoming), "Product and Process Flexibility in an Innovative Environment," *Rand Journal of Economics.*

Atkinson, A. B., and F. Bourguignon (1982), "The Comparison of Multi-Dimensioned Distributions of Economic Status." *Review of Economic Studies* (49): 183-201.

Bourbaki, N., (1987), *Topological Vector Spaces (Chapters 1-5).* Translated by H.G. Eggleston and S. Madan. Springer-Verlag: New York.

Brumelle, S. L., and R. G. Vickson (1975), "A Unified Approach to Stochastic Dominance," in *Stochastic Optimization Models in Finance*, W. T. Ziemba and R. G. Vickson, eds. New York: Academic Press.

Dunford, N. and J. T. Schwartz (1988), *Linear Operators*, Part I: *General Theory*, Wiley: New York.

Gollier, C. and M. Kimball (1995), "Toward a Systematic Approach to the Economic Effects of Uncertainty I: Comparing Risks," Mimeo, University of Toulouse.

Hadar, J. and W. R. Russel, (1971), "Stochastic Dominance and Diversification," *Journal of Economic Theory*, 3, 288-305.

Hopenhayn, Hugo, (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60 (5), 1127-1150.

Jewitt, Ian (1988), "Justifying the First Order Approach to Principal-Agent Problems," *Econometrica*, 56 (5), 1177-1190.

Kothe, Gottfried (1969), *Topological Vector Spaces I*, Translated by D.J.H. Garling, Springer-Verlag: New York.

Levy, Haim, and Jacob Paroush (1974), "Toward Multivariate Efficiency Criteria." *Journal of Economic Theory* 7, 129-142.

McAfee, Preston, and Phillip Reny (1992), "A Stone-Weierstrass Theorem Without Closure Under Suprema," *Proceedings of the American Mathematical Society* 114 (1): 61-67.

Meyer, Margaret, and Dilip Mookherjee, (1987), "Incentives, Compensation, and Social Welfare." *Review of Economic Studies,* 54, 209-226.

Meyer, Margaret (1990), "Interdependence in Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality Under Uncertainty." Nuffield College, Discussion Papers in Economics, No. 49, April.

Milgrom, Paul (1981), "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12 (2) Autumn. pp. 380-91.

Milgrom, Paul, and John Roberts (1990), "The Economics of Modern Manufacturing: Technology, Strategy, and Organization," *American Economic Review,* 80 (3): 511-528.

Milgrom, Paul, and John Roberts (1994), "Comparing Equilibria," *American Economic Review,* (June), 84 (3): 441-459.

Milgrom, Paul, and Chris Shannon (1994), "Monotone Comparative Statics," *Econometrica,* 62 (1), pp. 157-180.

Milgrom, Paul, and Robert Weber (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica* 50 (5): 1089-1122.

Rothschild, Michael and Joseph Stiglitz, (1970), "Increasing risk: I. A definition," *Journal of Economic Theory*, 2, 225-243.

Rothschild, Michael and Joseph Stiglitz, (1971), "Increasing risk: II. Its economic consequences," *Journal of Economic Theory*, 3, 66-84.

Royden, H.L., (1968), *Real Analysis*, 2nd ed., Macmillan: New York.

Schaefer, H.H., (1980), *Topological Vector Spaces*, Springer-Verlag: New York.

Shaked, Moshe and J. George Shanthikumar, (1994), *Stochastic Orders and Their Applications*. Academic Press: New York.

Topkis, Donald, (1968), "Ordered Optimal Solutions," Doctoral Dissertation, Stanford University.

Topkis, Donald, (1978), "Minimizing a Submodular Function on a Lattice," *Operations Research* 26: 305-321.

Whitt, Ward, (1980), "Uniform Conditional Stochastic Order." *Journal of Applied Probability*, 17, 112-123.

Whitt, Ward, (1982), "Multivariate Monotone Likelihood Ratio Order and Uniform Conditional Stochastic Order." *Journal of Applied Probability*, 19, 695-701.