1 Introduction

This chapter begins with the basic concepts and properties of stochastic dominance. It then gives examples of applications of stochastic dominance to various fields in economics: welfare analysis, finance, industrial organization, labor, international, health, and agricultural economics. The final subsection gives an overview of the subsequent chapters.

1.1 Concepts of Stochastic Dominance

1.1.1 Definitions

First-Order Stochastic Dominance (FSD) Let X_1 and X_2 be two (continuous) random variables with the cumulative distribution functions (CDFs) given by F_1 and F_2, respectively.[1] In economic applications, they typically correspond to incomes or financial returns of two different populations, which may vary regarding time, geographical regions or countries, or treatments. For $k = 1, 2$, let $Q_k(\tau) = \inf\{x : F_k(x) \geq \tau\}$ denote the quantile function of X_k, respectively, and let $\mathcal{U}_1$ denote the class of all monotone increasing (utility or social welfare) functions. If the functions are assumed to be differentiable, then we may write

$$\mathcal{U}_1 = \{u(\cdot) : u' \geq 0\}.$$

Definition 1.1.1 *The random variable X_1 is said to first-order stochastically dominate the random variable X_2, denoted by $F_1 \succeq_1 F_2$ (or X_1 FSD X_2),[2] if*

[1] Stochastic dominance can be defined for discrete or mixed continuous–discrete distributions. However, for the purpose of explanation, we shall mainly focus on continuous random variables, unless it is stated otherwise.

[2] To denote stochastic dominance relations, it is a convention to freely exchange the random variables with their respective distribution functions. For example, for first-order stochastic dominance, we may write $X_1 \succeq_1 X_2$ or F_1 FSD F_2. The same rule applies to the other concepts of stochastic dominance defined later.

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

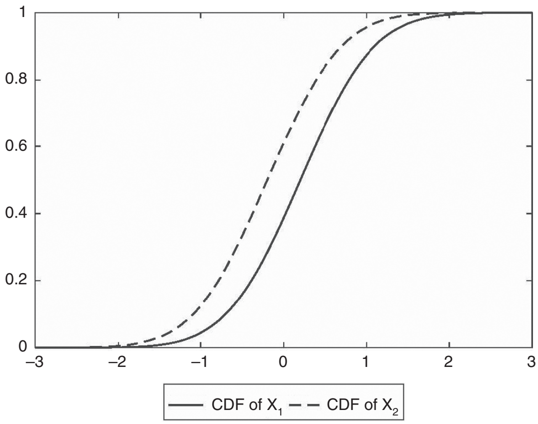

Figure 1.1 X_1 first-order stochastically dominates X_2

any of the following equivalent conditions holds: (1) $F_1(x) \leq F_2(x)$ *for all* $x \in \mathbb{R}$*; (2)* $E[u(X_1)] \geq E[u(X_2)]$ *for all* $u \in \mathcal{U}_1$*; and (3)* $Q_1(\tau) \geq Q_2(\tau)$ *for all* $\tau \in [0, 1]$.

This is the definition of *weak* stochastic dominance. If the inequalities hold with strict inequality for some $x \in \mathbb{R}$, some $u \in \mathcal{U}_1$, and some $\tau \in [0, 1]$, then the above serves as the definition of *strong* stochastic dominance, while one has *strict* stochastic dominance if the inequalities hold with strict inequality for all $x \in \mathbb{R}$, all $u \in \mathcal{U}_1$, and all $\tau \in [0, 1]$.[3] The equivalence of the three definitions will be discussed below.

Figure 1.1 illustrates two distributions with a first-order stochastic dominance relation. It shows that, when X_1 FSD X_2, the CDF of X_1 lies below that of X_2. To interpret the FSD relation, suppose that the random variables correspond to incomes of two different populations. Then, the inequality $F_1(x) \leq F_2(x)$ implies that the proportion of individuals in population 1 with incomes less than or equal to an income level x is not larger than the proportion of such individuals in population 2. If we measure poverty by the proportion of individuals earning less than a predetermined level of income (poverty line) x, then this implies that, whatever poverty line we choose, we have less poverty in F_1 than in F_2.[4] Therefore, the distribution F_1 would be preferred by any social planner having a welfare function that respects monotonicity ($u \in \mathcal{U}_1$),

[3] This classification is adopted from McFadden (1989, p. 115). The distinction among weak, strong, and strict dominance could be important in theoretical arguments. However, from a statistical point of view, the theoretically distinct hypotheses often induce the same test statistic and critical region, and hence the distinction is not very important; see McFadden (1989) and Davidson and Duclos (2000) for this point.

[4] See Section 5.2 for a general discussion about the relationship between poverty and SD concepts.

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

explaining the fact that we say that F_1 first-order stochastically dominates F_2 when the dominance of the CDFs as functions is the other way around.

To explain the FSD relation in an alternative perspective, write the (weak) first-order stochastic dominance relation $F_1 \succeq_1 F_2$ as

$$P(X_1 > x) \geq P(X_2 > x) \text{ for all } x \in \mathbb{R}. \tag{1.1.1}$$

Consider a portfolio choice problem of an investor and suppose that the random variables denote returns of some financial assets. Then, (1.1.1) implies that, for all values of x, the probability of obtaining returns not less than x is larger under F_1 than under F_2. Such a probability would be desired by every investor who prefers higher returns, explaining again the first-order stochastic dominance of F_1 over F_2. Conversely, if the two CDFs intersect, then (1.1.1) does not hold. In this case, one could find an investor with utility function $u \in \mathcal{U}_1$ such that $E[u(X_1)] > E[u(X_2)]$, and another investor with utility function $v \in \mathcal{U}_1$ such that $E[v(X_1)] < E[v(X_2)]$, violating the FSD of F_1 over F_2.

Second-Order Stochastic Dominance (SSD) To define the second-order stochastic dominance, let $\mathcal{U}_2$ denote the class of all monotone increasing and concave (utility or social welfare) functions. If the functions are assumed to be twice differentiable, then we may write

$$\mathcal{U}_2 = \{u(\cdot) : u' \geq 0, \ u'' \leq 0\}.$$

Definition 1.1.2 *The random variable X_1 is said to second-order stochastically dominate the random variable X_2, denoted by $F_1 \succeq_2 F_2$ (or X_1 SSD X_2), if any of the following equivalent conditions holds: (1) $\int_{-\infty}^{x} F_1(t)dt \leq \int_{-\infty}^{x} F_2(t)dt$ for all $x \in \mathbb{R}$; (2) $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2$; and (3) $\int_0^{\tau} Q_1(p)dp \geq \int_0^{\tau} Q_2(p)dp$ for all $\tau \in [0, 1]$.*

For SSD, the accumulated area under F_1 must be smaller than the counterpart under F_2 below any value of x. If X_1 first-order dominates X_2, or equivalently, if $F_1(x)$ is smaller than $F_2(x)$ for all x, then it is easy to see that X_1 second-order dominates X_2, but the converse is not true.

Figure 1.2 illustrates that, even when there is no first-order stochastic dominance between them (i.e., when the two CDFs intersect), X_1 may second-order stochastically dominate X_2.

To have second-order stochastic dominance $F_1 \succeq_2 F_2$, for any negative area ($F_2 < F_1$) there should be a positive area ($F_1 < F_2$) which is greater than or equal to the negative area and which is located before the negative area. To relate this to the second definition (2) of SSD, consider the expression

$$E[u(X_1)] - E[u(X_2)] = \int_{-\infty}^{\infty} [F_2(x) - F_1(x)]\, u'(x)dx,$$

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

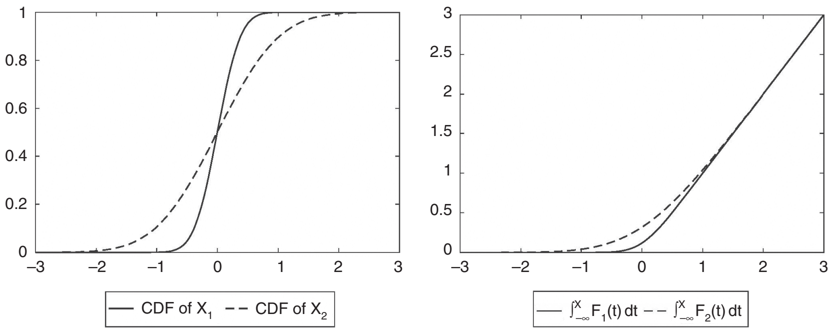

Figure 1.2 X_1 does not first-order stochastically dominate X_2, but X_1 second-order stochastically dominates X_2

which follows from integration by parts under regularity conditions (lemma 1 of Hanoch and Levy 1969; see also Equation 1.1.6). Whenever u' is a decreasing function (i.e., $u'' < 0$), the positive area is multiplied by a larger number $u'(x) > 0$ than the negative area which comes later on, so that the total integral becomes non-negative, establishing the second-order stochastic dominance of X_1 over X_2 under Definition 1.1.2 (2).

In the analysis of income distributions, the concavity assumption $u'' \leq 0$ implies that a transfer of income from a richer to a poorer individual always increases social welfare, which is a weaker form of the transfer principle (Dalton 1920). In the portfolio choice problem, on the other hand, the concavity assumption reflects risk aversion of an investor. That is, a risk-averse investor would prefer a portfolio with a guaranteed payoff to a portfolio without the guarantee, provided they have the same expected return. Therefore, the definition implies that any risk-averse investor would prefer a portfolio which dominates the other in the sense of SSD, because it yields a higher expected utility.

Higher-Order Stochastic Dominance The concept of stochastic dominance can be extended to higher orders. Higher-order SD relations correspond to increasingly smaller subsets of utility functions. Davidson and Duclos (2000) offer a very useful characterization of stochastic dominance of any order.

For $k = 1, 2$, define the *integrated CDF* and the *integrated quantile function* to be

$$F_k^{(s)}(x) = \begin{cases} F_k(x) & \text{for } s = 1 \\ \int_{-\infty}^{x} F_k^{(s-1)}(t)dt & \text{for } s \geq 2 \end{cases} \tag{1.1.2}$$

and

$$Q_k^{(s)}(x) = \begin{cases} Q_k(x) & \text{for } s = 1 \\ \int_{0}^{x} Q_k^{(s-1)}(t)dt & \text{for } s \geq 2. \end{cases}, \tag{1.1.3}$$

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

respectively. For $s \geq 1$, let

$$\mathcal{U}_s = \{u(\cdot) : u' \geq 0, u'' \leq 0, \ldots, (-1)^{s+1}u^{(s)} \geq 0\}$$

denote a class of (utility or social welfare) functions, where $u^{(s)}$ denotes the sth-order derivative of u (assuming that it exists).

Definition 1.1.3 *The random variable X_1 is said to stochastically dominate the random variable X_2 at order s, denoted by $F_1 \succeq_s F_2$, if any of the following equivalent conditions holds: (1) $F_1^{(s)}(x) \leq F_2^{(s)}(x)$ for all $x \in \mathbb{R}$ and $F_1^{(r)}(\infty) \leq F_2^{(r)}(\infty)$ for all $r = 1, ..., s-1$; (2) $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_s$; and (3) $Q_1^{(s)}(\tau) \geq Q_2^{(s)}(\tau)$ for all $\tau \in [0,1]$ and $Q_1^{(r)}(1) \geq Q_2^{(r)}(1)$ for all $r = 1, ..., s-1$.*

Whitmore (1970) introduces the concept of third-order stochastic dominance ($s = 3$, TSD) in finance; see also Whitmore and Findlay (1978). Levy (2016, section 3.8) relates the additional requirement $u''' \geq 0$ to the skewness of distributions and shows that TSD may reflect the preference for "positive skewness," i.e., investors dislike negative skewness but like positive skewness. Shorrocks and Foster (1987) show that the addition of a "transfer sensitivity" requirement leads to TSD ranking of income distributions. This requirement is stronger than the Pigou–Dalton principle of transfers since it makes regressive transfers less desirable at lower income levels.

If we let $s \to \infty$, then the class $\mathcal{U}_\infty$ of utility functions has marginal utilities that are completely monotone. This leads to the concept of infinite-order stochastic dominance, which is the weakest notion of stochastic dominance; see Section 5.4.3 for details.

Equivalence of the Definitions We now show the equivalence of the conditions that appear in the definitions of SD. For simplicity, we discuss the case of FSD and SSD, and assume that X_1 and X_2 have a common compact support, say $\mathcal{X} = [0, 1]$.[5]

We first establish the following lemma:

Lemma 1.1.1 *If $F_1(x) \leq F_2(x)$ for all $x \in \mathbb{R}$, then $EX_1 \geq EX_2$.*

Proof: Recall that, for any nonnegative random variable X with CDF F,

$$EX = \int_0^\infty P(X > t)\,dt = \int_0^\infty [1 - F(t)]\,dt; \tag{1.1.4}$$

[5] The equivalence results can be extended to general random variables, possibly with unbounded supports; see Hanoch and Levy (1969) and Tesfatsion (1976). The proofs in this subsection are based on Wolfstetter (1999, chapter 4) and Ross (1996, chapter 9). For a proof of strong stochastic dominance, see Levy (2016, section 3).

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

see, e.g., Billingsley (1995, equation 21.9). Therefore,

$$EX_1 - EX_2 = \int_0^{\infty} [P(X_1 > t) - P(X_2 > t)]\, dt$$
$$= \int_0^{\infty} [F_2(t) - F_1(t)]\, dt \geq 0. \quad (1.1.5)$$

□

The following theorem establishes the equivalence of (1) and (2) in Definition 1.1.1:[6]

Theorem 1.1.2 *$F_1(x) \leq F_2(x)$ for all $x \in \mathbb{R}$ if and only if $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_1$.*

Proof: Suppose that $F_1(x) \leq F_2(x)$ for all $x \in \mathbb{R}$ and let $u \in \mathcal{U}_1$ be an increasing function. Let $u^{-1}(z) = \inf\{x : u(x) > z\}$. For any $z \in \mathbb{R}$, we have

$$\begin{aligned} P(u(X_1) > z) &= P\left(X_1 > u^{-1}(z)\right) \\ &= 1 - F_1\left(u^{-1}(z)\right) \\ &\geq 1 - F_2\left(u^{-1}(z)\right) \\ &= P\left(X_2 > u^{-1}(z)\right) \\ &= P(u(X_2) > z). \end{aligned}$$

Therefore, by Lemma 1.1.1, we have $E[u(X_1)] \geq E[u(X_2)]$ for any $u \in \mathcal{U}_1$.

Conversely, suppose that $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_1$. Let

$$u_x(z) = \begin{cases} 1 & \text{if } z > x \\ 0 & \text{if } z \leq x \end{cases}.$$

Clearly, $u_x(\cdot) \in \mathcal{U}_1$ for each x. Therefore, for each $x \in \mathbb{R}$,

$$\begin{aligned} P(X_1 > x) &= E[u_x(X_1)] \\ &\geq E[u_x(X_2)] \\ &= P(X_2 > x). \end{aligned}$$

□

For SSD, the following theorem establishes the equivalence of (1) and (2) in Definition 1.1.2:[7]

Theorem 1.1.3 *$\int_0^x F_1(t)dt \leq \int_0^x F_2(t)dt$ for all $x \in \mathcal{X}$ if and only if $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2$.*

[6] The equivalence of the conditions (1) and (3) easily follows from monotonicity of the CDFs.
[7] For a proof of the equivalence of the conditions (1) and (3), see Thistle (1989, proposition 4).

Proof: Suppose that $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2$. Consider the following function:

$$u_x(z) = \begin{cases} z & \text{if } z \leq x \\ x & \text{if } z > x \end{cases} .$$

Obviously, for each $x \in \mathcal{X}$, $u_x(\cdot) \in \mathcal{U}_2$ so that

$$\begin{aligned} 0 &\leq E\left[u_x(X_1)\right] - E\left[u_x(X_2)\right] \\ &= \int_0^x \left[1 - F_1(t)\right] dt - \int_0^x \left[1 - F_2(t)\right] dt \\ &= \int_0^x \left[F_2(t) - F_1(t)\right] dt. \end{aligned}$$

Conversely, suppose that $\int_0^x F_1(t)dt \leq \int_0^x F_2(t)dt$ for all $x \in \mathcal{X}$. Since monotonicity implies differentiability almost everywhere (a.e.), we have $u' > 0$ and $u'' \leq 0$ a.e. for each $u \in \mathcal{U}_2$. Therefore, by integration by parts, we have

$$\begin{aligned} \Delta u &:= E\left[u(X_1)\right] - E\left[u(X_2)\right] \\ &= -\int_0^1 u(x)d\left[F_2(x) - F_1(x)\right] \\ &= \int_0^1 u'(x)\left[F_2(x) - F_1(x)\right] dx \qquad (1.1.6) \\ &= u'(1)\int_0^1 \left[F_2(t) - F_1(t)\right] dt \qquad (1.1.7) \\ &\quad - \int_0^1 u''(t) \int_0^t \left[F_2(s) - F_1(s)\right] ds dt. \end{aligned}$$

Since $u' > 0$ and $u'' \leq 0$, the assumed condition $\int_0^x \left[F_2(t) - F_1(t)\right] dt \geq 0$ for all $x \in \mathcal{X}$ implies immediately $\Delta u \geq 0$. This establishes Theorem 1.1.3. □

1.1.2 Basic Properties of Stochastic Dominance

While stochastic dominance relations compare *whole* distribution functions, they are also related to the moments and other aspects of distributions.

Let $supp(F)$ denote the support of distribution F. The following theorem gives sufficient and necessary conditions for the first-order stochastic dominance.

Theorem 1.1.4 *Let X_1 and X_2 be random variables with distribution functions F_1 and F_2, respectively. (1) If $P(X_2 \leq X_1) = 1$, then X_1 FSD X_2; (2) If* $\min\{supp(F_1)\} \geq \max\{supp(F_2)\}$, *then X_1 FSD X_2; (3) If X_1 FSD X_2, then $EX_1 \geq EX_2$ and* $\min\{supp(F_1)\} \geq \min\{supp(F_2)\}$.

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

(1) and (2) in the above theorem give sufficient conditions for the FSD. (1) holds because, if X_1 is not smaller than X_2 (with probability 1), then

$$\begin{aligned} &X_1 \leq x \text{ implies } X_2 \leq x \text{ for all } x \\ &\quad \Longrightarrow \{X_1 \leq x\} \subseteq \{X_2 \leq x\} \text{ for all } x \\ &\quad \Longrightarrow P(X_1 \leq x) \leq P(X_2 \leq x) \text{ for all } x \\ &\quad \Longrightarrow F_1(x) \leq F_2(x) \text{ for all } x.^8 \end{aligned}$$

For example, if $X_1 = X_2 + a$ for a constant $a > 0$, then (1) implies that X_1 FSD X_2. (2) says that if the minimum of the support of F_1 is not less than the maximum of the support of F_2, then we have first-order stochastic dominance of X_1 over X_2. This follows directly from (1).

On the other hand, (3) gives necessary conditions for FSD. That is, if X_1 FSD X_2, then the mean of X_1 is not smaller than the mean of X_2. This follows from the expression[9]

$$EX_1 - EX_2 = \int_{-\infty}^{\infty} [F_2(x) - F_1(x)]\, dx, \tag{1.1.8}$$

which is nonnegative, provided the integral exists; see also Lemma 1.1.1. Also, if X_1 FSD X_2, then the minimum of the support of F_1 is not smaller than that of F_2. This is called the "left tail" condition because it implies that F_2 has a thicker left tail than F_1. This result holds because, otherwise, there would exist a value x_0 such that $F_1(x_0) > F_2(x_0)$, and hence X_1 could not first-order stochastically dominate X_2.

For second-order stochastic dominance, analogous conditions can be established (the proofs are also similar):

Theorem 1.1.5 *Let X_1 and X_2 be random variables with distribution functions F_1 and F_2, respectively. (1) If X_1 FSD X_2, then X_1 SSD X_2; (2) If* $\min\{supp(F_1)\} \geq \max\{supp(F_2)\}$, *then X_1 SSD X_2; (3) If X_1 SSD X_2, then* $EX_1 \geq EX_2$ *and* $\min\{supp(F_1)\} \geq \min\{supp(F_2)\}$.

In the above theorem, (3) shows that $EX_1 \geq EX_2$ is a necessary condition for the SSD. Is there any general condition on variances which is also a necessary condition for the SSD? In general, the answer is no. However, for distributions with an equal mean, we can state a necessary condition for the SSD using their variances.

[8] Here, the notation '$A \Longrightarrow B$' means 'A implies B'.
[9] This holds because, for any random variable X with CDF F,

$$EX = \int_0^{\infty} [1 - F(x) - F(-x)]\, dx,$$

provided the integral exists.

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

Theorem 1.1.6 *Let X_1 and X_2 be random variables with $EX_1 = EX_2$. If X_1 SSD X_2, then $Var(X_1) \leq Var(X_2)$.*

To see this, take, for example, a quadratic utility function $u(x) = x + \beta x^2$ for $\beta < 0$, which certainly lies in $\mathcal{U}_2$. Then, $Eu(X_1) \geq Eu(X_2)$ and $EX_1 = EX_2$ together immediately imply that $Var(X_1) \leq Var(X_2)$. The mean-variance approach in the portfolio choice problem compares only the first two moments of distributions. A natural question would be whether F_1 second-order stochastically dominates F_2, if F_1 has larger mean and smaller variance than F_2. The answer is no, in general. This can be illustrated using the following counterexample (Levy, 1992, p. 567):

x	$P(X_1 = x)$	x	$P(X_2 = x)$
1	0.80	10	0.99
100	0.20	1000	0.01

Note that $EX_1 = 20.8 > EX_2 = 19.9$ and $Var(X_1) = 1468 < Var(X_2) = 9703$. Hence, X_1 dominates X_2 by the mean-variance criterion. However, X_1 does not second-order stochastically dominate X_2 because a risk-averse investor with utility function $u(x) = \log(x)$ would prefer X_2 over X_1, since $Eu(X_1) = 0.4 < Eu(X_2) = 1.02$; see the next subsection for another example with continuously distributed random variables.

The foregoing discussion implies that there is no direct relationship between the mean-variance approach and the stochastic dominance approach in general. However, in the special case of normal distributions, stochastic dominance can be related to mean-variance in the following sense:

Theorem 1.1.7 *Let X_1 and X_2 be random variables with normal distributions. Then, (1) $EX_1 > EX_2$ and $Var(X_1) = Var(X_2)$ if and only if X_1 FSD X_2; (2) if $EX_1 > EX_2$ or $Var(X_1) < Var(X_2)$, then X_1 SSD X_2.*

For more complete discussions on the properties of stochastic dominance, the reader may refer to Levy (2016) and Wolfstetter (1999, chapters 4–5).

1.1.3 A Numerical Example

The mean-variance criterion has been widely adopted in portfolio choice problems. It is a simple performance indicator comparing only the first two moments of distributions; whenever the mean is higher and the variance is lower for one distribution than for the other, the former distribution is preferred. However, it is well known that the criterion is valid only in certain

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

cases: (1) when the utility function is quadratic, and (2) when the distributions of the portfolios are all members of a two-parameter family; see Hanoch and Levy (1969). In reality, however, the assumptions are restrictive and the stochastic dominance approach provides an ordering of prospects under much less restrictive conditions.

To illustrate how the two approaches yield different results, we present a simple numerical example using two prospects, X and Y, with probability density functions (PDFs) and cumulative distribution functions (CDFs) given by

$$f_X(x) = 0.1 \cdot 1(0 \leq x < 1 \text{ or } 2 \leq x \leq 3) + 0.8 \cdot 1(1 \leq x < 2),$$
$$f_Y(x) = 0.5 \cdot 1(0.5 \leq x \leq 2.5)$$

and

$$\begin{aligned} F_X(x) &= 0.1x \cdot 1(0 \leq x < 1) + (0.8x - 0.7) \cdot 1(1 \leq x < 2) \\ &\quad + (0.1x + 0.7) \cdot 1(2 \leq x \leq 3) + 1(x > 3), \\ F_Y(x) &= 0.5(x - 0.5) \cdot 1(0.5 \leq x \leq 2.5) + 1(x > 2.5), \end{aligned}$$

respectively, where $1(\cdot)$ denotes the indicator function. Figure 1.3 depicts the PDFs and the CDFs of the prospects. Their expected values and variances are given by $E(X) = 3/2$, $Var(X) = 17/60$, $E(Y) = 3/2$, and $Var(Y) = 1/3$.

In terms of the mean-variance criterion, the prospect X is more efficient than the prospect Y. However, X does not second-order stochastically dominate Y, which can easily be observed from Figure 1.3. Since the value of the CDF of X is greater than that of Y over the region [0, 0.5], the integrated area of the distribution of X is greater than that of the distribution of Y. This violates the second-order stochastic dominance of X over Y.

In reality, we do not observe the population distributions F_X and F_Y, but rather a sample randomly drawn from the distributions. This motivates us

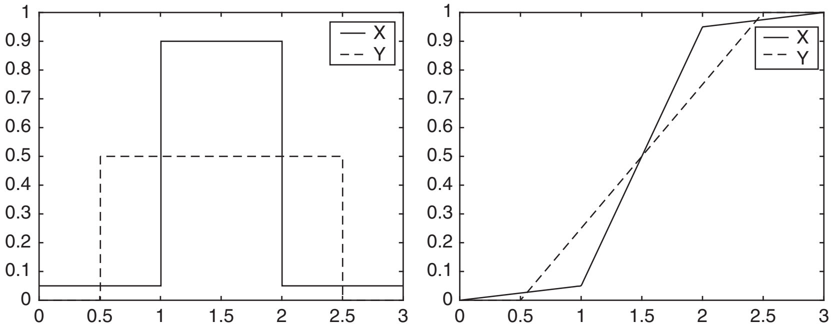

Figure 1.3 PDFs (left) and CDFs (right) for the simulation design

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

Table 1.1 *Mean-variance and SSD criteria*

Mean-Variance Criterion		
	Sample Mean	Sample Variance
X	1.5002	0.2831
Y	1.4997	0.3329
SSD Test		
Rejection rate	0.186	

to consider a statistical inference method: a scientific procedure to make a conjecture, based on samples, on the population distribution.

To give an early idea of SD tests, we draw a sample of size 500 from each distribution, then test the null hypothesis that "X second-order stochastically dominates Y" using a standard SD test (Barrett and Donald [BD] Test; see Section 2.2.2) at the significance level 0.10. We repeat this procedure 1,000 times. Table 1.1 summarizes the simulation results.

Sample means and sample variances are very close to their population values, implying that prospect X dominates prospect Y in the sample in terms of the mean-variance efficiency criterion.[10] On the other hand, the estimated rejection rate of the BD test in the simulation experiment is about 0.186. This means that the null hypothesis is rejected 186 out of 1,000 times. We may interpret this result as statistical evidence against second-order stochastic dominance of X over Y, which is consistent with the population relationship. This implies that SD tests can yield rankings of prospects that are fundamentally different from those based on the traditional mean-variance criterion.

1.1.4 Extensions and Some Related Concepts

In this subsection, we present a brief overview of some of the extensions and related concepts of stochastic dominance, which will be discussed in detail in the subsequent chapters.

Suppose that $F_1(y|x)$ and $F_2(y|x)$ denote the conditional CDFs of random variables Y_1 and Y_2 given $X = x$, respectively. We say that *the distribution of Y_1 (first-order) stochastically dominates the distribution of Y_2, conditional on X*, if

$$F_1(y|x) \leq F_2(y|x) \text{ for all } y, x.$$

[10] More formally, we cannot reject the null hypothesis of "$E(X) = E(Y)$ and $Var(X) \leq Var(Y)$."

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

This concept of *conditional stochastic dominance* is useful when one wishes to compare distributions of two population subgroups defined by some observed covariates X.

Stochastic monotonicity is a continuum version of the stochastic dominance hypothesis for conditional distributions. Let Y and X denote two random variables whose joint distribution is absolutely continuous with respect to a Lebesgue measure on $\mathbb{R}^2$. Let $F_{Y|X}(\cdot|x)$ denote the distribution of Y conditional on $X = x$. The hypothesis of stochastic monotonicity is defined to be

$$F_{Y|X}(y|x) \leq F_{Y|X}\left(y\middle|x'\right) \text{ for all } y \text{ and } x \geq x'.$$

For example, if X is some policy or other input variable, it amounts to the hypothesis that its effect on the distribution of Y is increasing.

The *Lorenz curve (LC)* plots the percentage of total income earned by various portions of the population when the population is ordered by their incomes, i.e., from the poorest to the richest. It is a fundamental tool for the analysis of economic inequality. The Lorenz curve is defined to be

$$L_k(p) = \frac{\int_0^p Q_k(t)dt}{\mu_k},$$

for $p \in [0, 1]$, where $Q_k(p)$ and μ_k denote the quantile function and the mean of F_k, respectively, for the population $k = 1, 2$. *Lorenz dominance of* L_1 *over* L_2 is defined by

$$L_2(p) \leq L_1(p) \text{ for all } p \in [0, 1].$$

Kahneman and Tversky (1979) criticize the expected utility theory and introduce an alternative theory, called prospect theory. They argue that individuals would rank prospects according to the expected value of S-shaped utility functions $u \in \mathcal{U}_P \subseteq \mathcal{U}_1$ for which $u''(x) \leq 0$ for all $x > 0$ but $u''(x) \geq 0$ for all $x < 0$. We say that X_1 *prospect stochastically dominates* X_2 if

$$\int_y^x F_1(t)dt \leq \int_y^x F_2(t)dt$$

for all pairs (x, y) with $x > 0$ and $y < 0$.

On the other hand, Levy and Levy (2002) discuss the concept of *Markowitz stochastic dominance*. In this case, individuals rank outcomes according to the expected value of reverse S-shaped utility functions $u \in \mathcal{U}_M \subseteq \mathcal{U}_1$ for which $u''(x) \geq 0$ for all $x > 0$ but $u''(x) \leq 0$ for all $x < 0$. We say that X_1 *Markowitz stochastically dominates* X_2 if

$$\left(\int_{-\infty}^y + \int_x^\infty\right) F_1(t)dt \leq \left(\int_{-\infty}^y + \int_x^\infty\right) F_2(t)dt$$

for all pairs (x, y) with $x > 0$ and $y < 0$.

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

Let F_1 and F_2 denote two CDFs, both continuous and with supports on $[0, \infty)$. We say that F_1 *initially dominates* F_2 up to an (unknown) point x_1, if

$$F_1(x) \leq F_2(x) \text{ for all } x \in [0, x_1)$$

with strict inequality for some $x \in (0, x_1)$. x_1 is called the "maximal point of initial dominance," when F_1 initially dominates F_2 up to x_1 and $F_1(x) > F_2(x)$ for all $x \in (x_1, x_1 + \varepsilon)$ for some sufficiently small $\varepsilon > 0$. This concept is useful to determine whether a stochastic dominance restriction holds over some unknown subset of the supports.

Let $X = (X_1, \ldots, X_K)^\intercal$ be a vector of K prospects (e.g., asset returns), and let Y be a benchmark portfolio. We say that Y is *stochastic dominance efficient* if there does not exist any portfolio $\{X^\intercal\lambda \ : \lambda \in \Lambda_0\}$ that stochastically dominates Y, where Λ_0 is a set of portfolio weights. This concept allows for full diversification of portfolios and provides a method of inference based on the first-order optimality condition of an investor's expected utility maximization problem.

Convex stochastic dominance, suggested by Fishburn (1974), is an extension of stochastic dominance to mixtures or convex combinations of distributions. If there are multiple choices available, a decision-maker can make pairwise comparisons with all prospects and establish whether a given choice is "SD admissible" in the sense that it is not dominated by any of the other alternatives. However, it is possible that some members of the SD admissible set may never be chosen by any individual with utility function in the hypothesized class. Instead, there may exist an "SD optimal" set, which is a subset of the SD admissible set that consists of the elements that *will* be chosen by some individuals with a utility function in the hypothesized class. The concept of convex stochastic dominance is useful to identify the SD optimal set.

The SD rule provides rankings of distributions based on *all* utility functions in a certain class. However, this can be restrictive in practice, because a small violation of the SD rule can make the ranking invalid. *Almost stochastic dominance,* suggested by Leshno and Levy (2002), is a weaker concept than the standard stochastic dominance. It applies to *most* rather than *all* decision-makers by eliminating economically pathological preferences. For example, let $\mathcal{U}_2^*(\varepsilon)$ denote the class of all increasing and concave functions with second derivatives satisfying some restrictions for some $\varepsilon > 0$. We say that X_1 *ε-almost second-order stochastically dominates* X_2, if $E[u(X_1)] \geq E[u(X_2)]$ for all $u \in \mathcal{U}_2^*(\varepsilon)$, or, equivalently,

$$\int \left[F_1^{(2)}(x) - F_2^{(2)}(x)\right]_+ dx \leq \varepsilon \int \left|F_1^{(2)}(x) - F_2^{(2)}(x)\right| dx,$$

where $[x]_+ = \max\{x, 0\}$. This concept also allows us to construct a measure of disagreement with stochastic dominance.

Approximate stochastic dominance, introduced by Álvarez-Esteban, del Barrio, Cuesta-Albertos, and Matrán (2016), is another weaker concept of the stochastic dominance relationship based on contaminated models. This concept is useful for determining whether one can establish stochastic dominance after trimming away some fraction of the (possibly contaminated) observations at the tails. The minimum level of contamination needed to establish SD can also serve as an alternative measure of disagreement with stochastic dominance.

Infinite-order stochastic dominance, characterized by Thistle (1993), is the weakest notion of stochastic dominance. It is defined by letting $s \to \infty$ in the definition of the sth-order stochastic dominance. Since the SD efficient set or the SD optimal sets are monotonically decreasing in s, the choice based on infinite-order stochastic dominance yields the smallest SD efficient or optimal set, which is a useful property for the portfolio choice problem.

There are many related concepts of stochastic dominance, other than the aforementioned ones. Let X_1 and X_2 be random variables with the absolute continuous distributions F_1 and F_2 with the densities f_1 and f_2, respectively. Then, we say that X_1 *density ratio dominates* X_2 if

$$\frac{f_2(t)}{f_1(t)} \text{ is nonincreasing over } t.$$

It can be shown that this holds if and only if

$$X_1|\,(a \le X_2 \le b) \text{ first-order stochastically dominates } X_2|\,(a \le X_1 \le b)$$

whenever $a \le b$, where $X|A$ denotes the conditional distribution of X given A. This concept of density ratio (or likelihood ratio) ordering has been frequently applied in portfolio choice, insurance, mechanism design, and auction theory.

We say that X_1 *uniformly stochastically dominates* X_2, if

$$\frac{1 - F_2(t)}{1 - F_1(t)} \text{ is nonincreasing in } t \in [0, b_{F_2}),$$

where $1 - F_k$ denotes the survival function for $k = 1, 2$ and $b_{F_2} = \inf\{x : F_2(x) = 1\}$. Uniform stochastic ordering is useful in many applications in which risks change dynamically over time. For example, in choosing medical treatments, the survival time of treatment A might stochastically dominate that of treatment B at the initial stage, but it may not be true when patients are examined at later stages. Also, the concept is used to compare distributions in the upper tails of financial returns. It is a stronger concept than the FSD.

The concept of SD is also closely related to various concepts of dependence among random variables. Let X and Y be two random variables with the CDFs F and G, respectively. We say that Y *is positive quadrant dependent on* X, if

$$P\,(X \le x, Y \le y) \ge P\,(X \le x)\,P\,(Y \le y) \text{ for all } (x, y).$$

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

We also say that Y *is positive expectation dependent on* X, if

$$E(Y) - E(Y|X \leq x) \geq 0 \text{ for all } x.$$

These concepts have been applied extensively in the literature of finance, insurance, and risk management. Expectation dependence (Wright, 1987) is a weaker concept than quadrant dependence, but is a stronger concept than correlation. Expectation dependence also plays an important role in portfolio theory, because it is used to determine the necessary and sufficient conditions for portfolio diversification.

Central dominance (Gollier, 1995) provides the conditions under which a change in distribution increases the optimal value of a decision variable for all risk-averse agents. It implies a deterministic change in optimal decision variables, such as demand for risky assets or a social welfare policy, when the distribution changes. It is shown that stochastic dominance is neither sufficient nor necessary for central dominance.

Spatial dominance (Park, 2005) is a generalization of stochastic dominance from the frequency domain to the spatial domain. It allows us to compare, for example, performances of two assets over a given time interval. It has important implications for optimal investment strategies that may be horizon-dependent.

1.2 Applications of Stochastic Dominance

The concept of stochastic dominance and related concepts have been applied to real data in many areas of social and natural sciences. The literature has been expanding quite rapidly and it is beyond the scope of this book to cover all of the applications; see Mosler and Scarsini (1993) for an extensive bibliography on the literature up to the early 1990s. See also Wolfstetter (1999, chapter 5) for an overview of theoretical developments of SD in economics.

In this section, we briefly summarize some recent empirical applications of stochastic dominance, mainly to economics. They are chosen to illustrate the wide applicability of SD concepts and are not meant to be comprehensive. Some of them will be discussed in more detail in subsequent chapters.

1.2.1 Welfare Analysis

One of the main applications of stochastic dominance is welfare analysis. In particular, the issues of inequality, poverty, and polarization have frequently been analyzed using the concept of SD. For an excellent overview of the literature, see Cowell and Flachaire (2015).

Anderson (2003) uses SD criteria to examine improvements to poverty alleviation in the United States in the 1980s using the PSID data. Anderson (2004*a*), on the other hand, examines enhancements to polarization,

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

welfare, and poverty based on per-capita Gross National Products (GNP) of different countries. Anderson and Leo (2009) study changes in child poverty, investment in children, and generational mobility since the introduction of China's One Child Policy (OCP) in 1979. They compare child poverty in China (1987, 2001) with that in other countries such as Canada (1997, 2004), the United Kingdom (1996, 2002), and India (1994, 2004) by utilizing SD methods. Evolution of poverty is also studied by Contreras (2001) using Chilean data from 1990 to 1996 (a period of rapid growth in Chile).

Amin, Rai, and Topa (2003) use SD methods to evaluate whether microcredit programs in Bangladesh reach relatively poorer and more vulnerable people. They test the hypothesis that distributions of consumption and vulnerability of program participants stochastically dominate those of nonparticipants. They find that the microcredit program effectively reaches the poor, but it is not very successful at reaching the vulnerable. On the other hand, Skoufias and Di Maro (2008) investigate the effect of Progressa, which is a conditional cash transfer program implemented in Mexico, on poverty. They find that the income distribution of treated households first-order stochastically dominates that of controlled households, strengthening robustness of the other results obtained using different methods.

There are also several empirical studies that try to compare distributions of income and/or other socioeconomic variables using various SD tests. Heshmati and Rudolf (2014) examine inequality and poverty in Korea using distributions of income and consumption. Valenzuela, Lean, and Athanasopoulos (2014) study inequality in income and expenditure distributions in Australia from 1983 to 2010. Maasoumi and Heshmati (2000) consider multivariate generalizations of a univariate SD test to examine Sweden's income distributions for the whole population and its subgroups.

Maasoumi, Su, and Heshmati (2013) examine inequality and relative welfare levels over time and among population subgroups using the Chinese Household Nutrition Survey (CHNS) data. Chen (2008) employs SD methods to study regional income disparities in Canada. Zarco, Pérez, and Alaiz (2007) also use an SD method to investigate the effect of the European Union (EU)'s structural funds on convergence of the income distributions in Spanish regions for the time period from 1990 to 2003; see also Carrington (2006) and Liontakis, Papadas, and Tzouramani (2010) for related results about the regional income convergence.

Pinar, Stengos, and Topaloglou (2013) emphasize joint dependence among various attributes of welfare such as income, health, and education (see, e.g., Maasoumi 1999 or Fleurbaey 2009 for an overview). Using the concept of stochastic dominance efficiency (see Section 5.3.2), they propose an optimal weighting scheme for measuring human development. Compared to the traditional United Nation's Development Program's Human Development Index (HDI) which puts equal weights to three basic components (life expectancy, education, and GNI), the optimal index leads to a marked improvement of measured development. Based on the panel data of different countries for the period 1975 to 2000, they present new country rankings that are quite different from those based on the HDI.

1.2.2 Finance

Another major area of applications of stochastic dominance is financial economics. There are numerous applications of SD in the finance literature; see Levy (2016) and Sriboonchita, Wong, Dhompongsa, and Nguyen (2010) for extensive surveys. Below we mention just a few recent applications.

SD tests have been implemented to track down evidence of financial market inefficiency. One of the representative phenomena demonstrating market inefficiency is the calendar effect, i.e., investment strategies linked to particular times may earn more profits. Seyhun (1993) examines the *January Effect*, which refers to the unusually large, positive average stock returns during the last few days of December and the first week of January. Cho, Linton, and Whang (2007) find empirical evidence for the *Monday Effect*, the phenomenon that Monday stock returns are systematically smaller than returns on any other day of the week (Section 2.5.2).

SD methods have also been used to evaluate the profitability of investment strategies. Bali, Demirtas, Levy, and Wolf (2009), using the concept of almost stochastic dominance (see Section 5.4.1), find empirical evidence in favor of the popular practice of primarily allocating a greater proportion to stocks and then gradually relocating funds to bonds as the investment horizon shortens. Ibarra (2013) also finds evidence that bonds dominate stocks at short horizons, while stocks dominate bonds at long horizons, based on spatial dominance (see Section 5.5.6). Meanwhile, Abhyankar, Ho, and Zhao (2009) focus on investors' preference of value stocks to growth stocks, and demonstrate that the value premium is country- and sample-specific. Fong, Wong, and Lean (2005) test the hypothesis that there exist general asset pricing models that explain the *Momentum Effect*, which is a tendency for portfolios of stocks that have performed well in recent months to continue to earn positive returns over the next year. Fong (2010) examines profitability of the investment strategy of yen carry trade over the period 2001–09. Chan, de Peretti, Qiao, and Wong (2012) show the efficiency of the UK covered warrants market by comparing the returns of covered warrants and their underlying shares.

Post (2003) develops a statistical test of stochastic dominance efficiency (see Section 5.3.2) to test superior profitability of a given portfolio over other possible combinations of assets. He shows the inefficiency of the Fama and French market portfolio relative to the benchmark portfolios formed on market capitalization and book-to-market equity ratio; see also Post and van Vliet (2006) for related results. Li and Linton (2010) propose a method to construct a portfolio based on SD and demonstrate that the hedge fund portfolios constructed by stochastic dominance criteria outperform other randomly selected hedge fund portfolios and a mean-variance efficient hedge fund portfolio. Using the concept of stochastic dominance efficiency, Agliardi et al. (2012) construct a sovereign risk index in emerging markets by aggregating various risk factors such as economic, political, and financial risks; see also Agliardi, Pinar, and Stengos (2014) and Pinar, Stengos, and Yazgan (2012) for related applications.

1.2.3 Industrial Organization

Guerre, Perrigne, and Vuong (2009, GPV) investigate nonparametric identification of the first-price auction model with risk-averse bidders. They show that quantiles of the observed equilibrium bid distributions with different numbers of bidders should satisfy a set of inequality restrictions, which in turn implies a stochastic dominance relationship between the distributions.

Specifically, let $I_2 > I_1 \geq 2$ denote the two different numbers of bidders. For each τ such that $0 < \tau < 1$, let $Q_k(\tau)$ denote the τ quantile of the observed equilibrium bid distribution G_k when the number of bidders is I_k for $k \in \{1, 2\}$. If the auctions are homogeneous and the private values are independent of the number of bidders, then GPV (equation 5, p. 1201) show that, under some additional assumptions, the quantiles should satisfy:

$$\frac{I_1 - 1}{I_2 - 1} Q_2(\tau) + \frac{I_2 - I_1}{I_2 - 1} \underline{b} < Q_1(\tau) < Q_2(\tau) < \frac{I_2 - 1}{I_1 - 1} Q_1(\tau) + \frac{I_1 - I_2}{I_1 - 1} \underline{b} \tag{1.2.1}$$

for any $\tau \in (0, 1]$, where $\underline{b}$ is the left endpoint of the support of the observed bid distributions. The inequality (1.2.1) offers a testable implication: the observed bid distribution with I_2 bidders first-order stochastically dominates the distribution with I_1 bidders, i.e., $G_2 \succeq_1 G_1$. On the other hand, if the auctions are heterogeneous so that the private values are affected by (observed) characteristics, then one may consider conditionally exogenous participation with the conditional version of the restrictions:

$$\frac{I_1 - 1}{I_2 - 1} Q_2(\tau|x) + \frac{I_2 - I_1}{I_2 - 1} \underline{b} < Q_1(\tau|x) < Q_2(\tau|x) < \frac{I_2 - 1}{I_1 - 1} Q_1(\tau|x) + \frac{I_1 - I_2}{I_1 - 1} \underline{b} \tag{1.2.2}$$

for any $\tau \in (0, 1]$ and $x \in \mathcal{X}$, where $\mathcal{X}$ is the support of X, $Q_k(\tau|x)$ is the τth conditional quantile (given $X = x$) of the observed equilibrium bid distribution when the number of bidders is I_k, and X denotes the observed auction characteristics such as appraisal values (see section 3.2 of GPV). In this

https://doi.org/10.1017/9781108602204.002 Published online by Cambridge University Press

case, the testable implication is stochastic dominance between the conditional distributions, i.e., $G_2(\cdot|\cdot) \succeq_1 G_1(\cdot|\cdot)$. Using a timber auction data, Lee, Song, and Whang (2017) test the latter hypothesis and find no empirical evidence against (1.2.2); see Section 6.2 for details.

De Silva, Dunne, and Kosmopoulou (2003) compare the bidding patterns of entrants and incumbents in road construction auctions, using data from the Oklahoma Department of Transportation. They find that entrants bid more aggressively (in the sense of lower bids relative to engineering costs) and win auctions with much lower bids than incumbents. They also find that the difference is more prominent in the lower tail of the bid distribution. They justify this phenomenon theoretically by using an auction model with information asymmetries due to varying levels of experience and efficiency of the auction participants. In particular, they show that if the distribution of cost estimates of entrants first-order stochastically dominates that of incumbents, then the entrants will bid more aggressively relative to their cost estimates than the incumbents will. Based on quantile regressions controlling for auction heterogeneity, they find no evidence against the implication of their asymmetric auction model (see Section 2.2.4 for the relationship between SD tests and quantile regressions).

Pesendorfer (2000) studies collusive behaviors in first-price auctions, focusing on the Florida and Texas school milk cartels in the 1980s. He considers a theoretical model of cartel behavior and shows that if cartel firms and non-cartel firms are identical, and the cartel is efficient (in the sense that it selects the lowest-cost cartel firm), then there will be an induced asymmetry between selected cartel bidders and non-cartel firms. One of the testable implications of the bidding equilibrium is that the ex ante bid distribution of the cartel is first-order stochastically dominated by that of non-cartel bidders. Based on the data of 4,077 contracts, he tests the equality of the distributions of the residuals from the OLS regressions of cartel bids and non-cartel bids on covariates that represent auction heterogeneity. Figures 3 and 4 of the latter paper suggest that the distribution of cartel residuals is first-order stochastically dominated by that of non-cartel residuals. He does not perform a formal SD test, but employs a Chow test and a rank test and concludes that there is empirical evidence consistent with the implication of his theoretical model.

Aryal and Gabrielli (2013) provide a method to detect collusion in asymmetric first-price auctions. The basic idea is that, if the same bid data are used to recover the underlying latent cost, the cost under competition must stochastically dominate the cost under collusion, because collusion tends to increase the markup. This suggests that detecting collusion can boil down to testing for the first-order stochastic dominance. They implement the standard Kolmogorov–Smirnov and Mann–Whitney–Wilcoxon tests to the highway procurement data and find no evidence of collusion in the data.

1.2.4 Labor Economics

Maasoumi, Millimet, and Rangaprasad (2005) analyze the effect of class sizes on student achievements with a US data set using the SD approach. In their study, the actual class sizes and the distributions of test scores of the 8th and 10th grades are utilized as the indicators for school quality and measures of student achievements, respectively. The dominance relation of the test score distributions are tested by actual class sizes: small, medium, and large. Eren and Henderson (2008) and Eren and Millimet (2008) also employ SD methods to examine the factors that impact on student achievements (such as homework) and organizational structures of schools, respectively.

Millimet and Wang (2006) examine the gender earnings gap in China based on the generalized Kolmogorov–Smirnov test proposed by Linton, Maasoumi, and Whang (2005) (see Section 2.2.2) and Maasoumi and Heshmati (2000). The dominance relation for the distributions of annual earnings and hourly wages for male and female workers is investigated for the years 1988 and 1995. See also Maasoumi and Wang (2018) for a related study of the gender earnings gap for the US labor market over the last several decades.

Maasoumi, Millimet, and Sarkar (2009) investigate the marriage premium, which is a phenomenon that the average earnings of married men are higher than those of unmarried men, by employing the SD approach to examine whether the phenomenon appears for the whole wage distribution. They find that the marriage premium is confined primarily to the lower tails of the wage distribution and the majority of the premium can be explained by self-selection.

1.2.5 International Economics

Delgado, Farinas, and Ruano (2002) adopt the SD approach to examine the total factor productivity difference between exporting and non-exporting firms based on a sample of Spanish manufacturing firms from 1991 to 1996. The productivity of four groups (exporters, non-exporters, entering exporters, and exiting exporters) is compared in the paper. They find that the productivity distribution of small exporting firms stochastically dominates that of small non-exporting firms, but, in the case of large firms, no dominance relation seems to exist. A related research is conducted by Girma, Görg, and Strobl (2004) and Elliott and Zhou (2013).

Helpman, Melitz, and Yeaple (2004) provide a theoretical framework for the *Market Selection Hypothesis* that firms determine the types of markets in which they run their business depending on their own profitability (that is, high profitability = foreign direct investment, medium profitability = foreign markets through exporting, and low profitability = domestic markets). Girma, Kneller, and Pisu (2005) test the hypothesis using the SD approach. They find that the productivity distribution of multinational firms dominates that of export firms, which in turn dominates that of non-exporters. Wagner (2006) also compares the productivity distributions using a German data set and finds that the productivity distribution of foreign direct investors dominates that of exporters, which in turn dominates that of national market suppliers; see also Arnold and Hussinger (2010) for a related result.

1.2.6 Health Economics

Bishop, Formby, and Zeager (2000) examine the impact of food stamp cashout on undernutrition using a household data from cashout experiments in Alabama and San Diego. They compare the CDFs of nutrients for two population subgroups (cash recipients and food coupon recipients), truncated in the neighborhood of the RDA (Recommended Dietary Allowances). They apply the FSD test of Bishop, Formby, and Thistle (1989) to the data and find that a substantial proportion of household participants falls short of the recommended levels of food energy and a variety of nutrients. Furthermore, in Alabama, the cash recipients show higher deficiency of Vitamin E and B than the coupon recipients.

Pak, Ferreira, and Colson (2016) investigate the obesity inequality among US adults over time using an SD test. Because people care about their weights relative to their peers, obesity inequality plays an important role in subjective well-being. Using the National Health and Nutritional Examination Survey (NHANES) data, they find that the BMI (Body Mass Index) distribution of each NHANES study first-order stochastically dominates that of the previous wave from 1971–74 to 2003–06, while more recent comparisons fail to reject the null hypothesis of nondominance. Madden (2012) and Sahn (2009) also investigate the obesity issue using SD methods.

1.2.7 Agricultural Economics

Langyintuo et al. (2005) use the SD approach to assess risk efficiency of yields and returns to farmers' household resources in rice production across different production systems. An improved (short-duration cover crop) fallow system is compared with two alternative fallow systems (the traditional natural bush fallow and continuous rice-cropping systems). The analysis assumes that farmers try to maximize both food self-sufficiency (rice grain yield) and cash income (monetary returns to farm household resources). Using agronomic data from Northern Ghana and employing the two-sample Kolmogorov–Smirnov test, they conclude that the yield and income distributions of the improved fallow system stochastically dominate those of the alternative systems. See Mahoney et al. (2004), Smith, Clapperton, and Blackshaw (2004), Ribera, Hons, and Richardson (2004), and Lien et al. (2006) for related studies.

It is important to prevent land degradation in the form of soil erosion and nutrient depletion in order to ensure food security and sustainability of agricultural production in many developing countries. Kassie et al. (2008) examine the impact of stone bunds on the value of crop production in the areas of the Ethiopian highlands using cross-sectional data. They find that, in the regions with low rainfall, the yield distributions with conservation first-order stochastically dominate those without conservation, while the relation is reversed and is not significant in the regions with heavy precipitation. For related studies, see Kassie et al. (2009) and Bekele (2005).

1.3 Outline of Subsequent Chapters

Chapter 2 introduces the basic ideas of standard tests of stochastic dominance, frequently used in the literature. The tests are classified into different categories depending on hypotheses of interest and types of test statistics. Various types of test statistics are discussed and compared. Two empirical examples are also given to illustrate how the tests can be used in practical applications.

Chapter 3 introduces methods to improve power performance of some SD tests. Many of the existing tests of stochastic dominance consider the least favorable case (LFC) of the null hypothesis to compute critical values. However, they may be too conservative in practice, because their asymptotic distributions depend only on the binding part (or so-called "contact set") of the inequality restrictions. This chapter discusses various approaches to improve power performance by utilizing information about the binding restrictions. This chapter also introduces applications of stochastic dominance to program evaluations. In particular, inference methods for distributional treatment effects and counterfactual policy effects are discussed. Some other issues of stochastic dominance tests, such as the problem of unbounded supports, classification rules for SD relations, and large deviation approximation of the distribution of the SD test statistics, are also discussed. This chapter provides empirical examples to evaluate distributional treatment effects and returns to schooling.

Chapter 4 provides the main ideas of how to test stochastic dominance when there are covariates. Consideration of the covariates is important in many economic applications because stochastic dominance relations might hold only for subpopulations defined by observed covariates. In particular, this chapter introduces the notions of stochastic monotonicity and conditional stochastic dominance. It also illustrates empirical applications to evaluate distributional conditional treatment effects using a real data set on academic achievements and to test a strong leverage hypothesis in financial markets.

Chapter 5 introduces nonparametric testing procedures for various extensions of stochastic dominance, including multivariate stochastic dominance, Lorenz dominance, poverty dominance, initial dominance, marginal conditional stochastic dominance, stochastic dominance efficiency, convex stochastic dominance, almost stochastic dominance, approximate stochastic dominance, and infinite-order stochastic dominance, as well as some related concepts such as density ratio ordering, uniform stochastic ordering, positive quadrant dependence, expectation dependence dominance, central dominance, and spatial dominance.

Chapter 6 discusses some further topics recently studied in the literature. They include inference on a distributional overlap measure, testing for generalized functional inequalities, SD tests under measurement errors, conditional SD tests with many covariates, and robust forecasting comparisons.

Chapter 7 concludes the book and suggests some future directions for econometric research on stochastic dominance.

Finally, the appendices provide the basic technical tools and the MATLAB code for some of the SD tests discussed in the main text.