Make Almost Stochastic Dominance really Almost

Guo, Xu and Wong, Wing-Keung and Zhu, Lixing

Hong Kong Baptist University, Hong Kong Baptist University, Hong Kong Baptist University

11 September 2013
Make Almost Stochastic Dominance really Almost

Xu Guo¹, Wing-keung Wong² and Lixing Zhu¹
¹Department of Mathematics, Hong Kong Baptist University, Hong Kong
²Department of Economics, Hong Kong Baptist University, Hong Kong

September 11, 2013

Abstract: Leshno and Levy (2002) extend stochastic dominance (SD) theory to almost stochastic dominance (ASD) for most decision makers. When comparing any two prospects, Guo, et al. (2013) find that there will be ASD relationship even there is only very little difference in mean, variance, skewness, or kurtosis. Investors may prefer to conclude ASD only if the dominance is nearly almost. Levy, et al. (2010) have provided two approaches to solve the problem. In this paper, we extend their work by first recommending an existing stochastic dominance test to handle the issue and thereafter developing a new test for the ASD which could detect dominance for any pre-determined small value. We also provide two approaches to obtain the critical values for our proposed test.

Keywords: stochastic dominance; almost stochastic dominance; risk aversion, stochastic dominance test, almost stochastic dominance test

JEL Classification : C0, D81, G11.

1 Introduction

Stochastic dominance (SD) theory has been well established, see, for example, Hanoch and Levy (1969), Hadar and Russell (1969), and Rothschild and Stiglitz (1970). Leshno and Levy (2002) extend it to the theory of almost stochastic dominance (ASD) for most decision makers. ASD has been widely used in many areas, especially in finance, see, for example, Levy (2006, 2009), Bali, et al. (2009), and Levy, et al. (2010).

Recently, Guo, et al. (2013) find that when comparing any two prospects, there will be ASD relationship even there is only very little difference in mean, variance, skewness, or kurtosis. Investors may prefer to conclude ASD only if the dominance is nearly almost. Levy, et al. (2010) have provided two approaches to solve the problem. In this paper, we extend their work by first recommending an existing stochastic dominance test to handle the issue and thereafter developing a new test for the ASD which could detect dominance for any pre-determined small value. We also provide two approaches to obtain the critical values for our proposed test.
2 Notations and Definitions

Suppose that random variables $X$ and $Y$ defined on the support $\Omega = [a, b]$ with means $\mu_X$ and $\mu_Y$ and standard deviations $\sigma_X$ and $\sigma_Y$ have the corresponding distribution functions $F$ and $G$, respectively.

$$H^{(1)} = H \quad \text{and} \quad H^{(n)}(x) = \int_a^x H^{(n-1)}(t) \, dt \quad \text{for} \quad H = F, G \quad \text{and} \quad n = 2, 3, 4 ;$$

$$||F^{(n)} - G^{(n)}|| = \int_a^b |F^{(n)}(x) - G^{(n)}(x)| \, dx , \quad \text{and} \quad (1)$$

$$S_n \equiv S_n(F, G) = \{ x \in [a, b] : G^{(n)}(x) < F^{(n)}(x) \} \quad \text{for} \quad n = 1, \ldots , 4 .$$

In this paper we modify the concept of the ASD by restricting the range of $\epsilon$ to be smaller than a predetermined value, say, for example, 5%, set by users, instead of the value of $1/2$ used in the ASD definition stated in Leshno and Levy (2002), modified by Tzeng et al. (2012), and Guo, et al. (2013, 2013a). To be precise, we present the definition as follows:

**Definition 1** Let $F$ and $G$ be the corresponding distribution functions of $X$ and $Y$. For any predetermined value positive value $\epsilon_0$ which is much smaller than $1/2$,

$\epsilon_0$-ASD$_1$: $X$ is said to dominate $Y$ by $\epsilon_0$-FSD$_1$, denoted by $X \succeq_{1}^{\epsilon_0} Y$ or $F \succeq_{1}^{\epsilon_0} G$, if and only if

$$\int_{S_1} [F(x) - G(x)] \, dx \leq \epsilon_0 ||F - G||$$

$\epsilon_0$-ASD$_2$: $X$ is said to dominate $Y$ by $\epsilon_0$-ASD$_2$, denoted by $X \succeq_{2}^{\epsilon_0} Y$ or $F \succeq_{2}^{\epsilon_0} G$, if and only if

$$\int_{S_2} [F^{(2)}(x) - G^{(2)}(x)] \, dx \leq \epsilon_0 ||F^{(2)} - G^{(2)}|| \quad \text{and} \quad \mu_X \geq \mu_Y ;$$

$\epsilon_0$-ASD$_3$: $X$ is said to dominate $Y$ by $\epsilon_0$-ASD$_3$, denoted by $X \succeq_{3}^{\epsilon_0} Y$ or $F \succeq_{3}^{\epsilon_0} G$, if and only if

$$\int_{S_3} [F^{(3)}(x) - G^{(3)}(x)] \, dx \leq \epsilon_0 ||F^{(3)} - G^{(3)}|| \quad \text{and} \quad G^{(n)}(b) \geq F^{(n)}(b) \quad \text{for} \quad n = 2, 3 .$$

$\epsilon_0$-ASD$_4$: $X$ is said to dominate $Y$ by $\epsilon_0$-ASD$_4$, denoted by $X \succ_{4}^{\epsilon_0} Y$ or $F \succ_{4}^{\epsilon_0} G$, if and only if

$$\int_{S_4} [F^{(4)}(x) - G^{(4)}(x)] \, dx \leq \epsilon_0 ||F^{(4)} - G^{(4)}|| \quad \text{and} \quad G^{(n)}(b) \geq F^{(n)}(b) \quad \text{for} \quad n = 2, 3, 4 ,$$

2
where $\epsilon_0$-ASD$_n$ is the n-order ASD for $n = 1, \ldots, 4$.

From Definition 1, we can see clearly that the proportions of the nagging negative areas in which $G^{(n)}(x)$ below $F^{(n)}(x)$ over the total absolute area differences are bounded by a predetermined value $\epsilon_0$ which is much smaller than $1/2$. To explain the advantage of using our modification, we define the following utility functions:

**Definition 2** For $n = 1, \ldots, 4$,

$$
U_n = \{ u : (-1)^i u^{(i)}(x) \leq 0, \ i = 1, \ldots, n \},
$$

$$
U_n^*(\epsilon_0) = \{ u \in U_n : (-1)^{n+1} u^{(n)}(x) \leq \inf \{ (-1)^{n+1} u^{(n)}(x) \} [1/\epsilon_0 - 1] \forall x \},
$$

in which $\epsilon_0$ is much smaller than $1/2$.

We note that since $\epsilon_0$ is a predetermined value which is much smaller than $1/2$, $U_n^*(\epsilon_0)$ can be much closer to $U_n$ such that most of the investors in $U_n$ will be in $U_n^*(\epsilon_0)$. Thus, the advantage of using $\epsilon_0$-ASD – the modification of ASD – in Definition 1 is that when one confirms $X$ is preferred to $Y$ in the sense of $\epsilon_0$-ASD$_n$, one could conclude that all investors with $u$ in $U_n^*(\epsilon_0)$ will prefer $X$ is preferred to $Y$ which, in turn, could implies that most of the investors with $u$ in $U_n$ will prefer $X$ is preferred to $Y$. Hence, we claim that by using our modified ASD could draw preference for most investors if $\epsilon_0$ is small.

### 3 The Theory

Recently, Guo, et al. (2013) find that when comparing any two prospects, there will be ASD relationship even there is only very little difference in mean, variance, skewness, or kurtosis. Investors may prefer to conclude ASD only if the dominance is nearly almost. Levy, et al. (2010) have provided a good solution. They suggest two approaches. We modify their suggestion as follows:

The first approach is to check the actual area violation $\epsilon$ that is (significantly) smaller than $1/2$. The second approach is to find for a given group of subjects what is the allowed area violation by each investor and whether for all subjects belonging to this group the allowed area violation is greater than the actual area violation.

---

1We note that the theory can be extended to satisfy utilities defined to be non-differentiable and/or non-expected utility functions, readers may refer to Wong and Ma (2008) and the references therein for more information.
In this paper, we extend their work by first recommending an existing stochastic dominance test to handle the issue and thereafter developing a new test for the ASD which could detect dominance for any pre-determined small value. We first recommend an existing stochastic dominance test to handle the issue in next subsection.

3.1 Initial Stochastic Dominance Test

There are several SD tests. Among them, there are two broad classes of SD tests. One is the minimum/maximum statistic, while the other is based on distribution values computed on a set of grid points. McFadden (1989) first develops a SD test using the minimum/maximum statistic. Later on, Barrett and Donald (2003) develop a Kolmogorov-Smirnov-type test and Linton et al. (2005) extend their work to relax the iid assumption. On the other hand, the SD tests developed by Anderson (1996) and Davidson and Duclos (DD, 2000) compare the underlying distributions at a finite number of grid points whereas Bai, et al. (2011) extend their work by deriving the limiting distributions of the test statistics to be stochastic processes, proposing a bootstrap method to decide the critical points of the tests and proving the consistency of the bootstrap tests. The advantages of the DD test are (1) it has been examined to be one of the most powerful approaches and yet less conservative in size (see Tse and Zhang, 2004; Lean et al., 2008), and (2) it compares the underlying distributions at a finite number of grid points which, in turn, tells us the percentage of one prospect dominating another and vice verse. This information is useful for decision makers to determine the ASD relationship among the prospects and thus we first recommend decision makers to apply the DD test to determine the ASD relationship among the prospects. Readers may refer to Davidson and Duclos (2000) and Bai, et al. (2011) for the test statistics or refer to Fong, et al. (2005, 2008), Gasbarro, et al. (2007, 2012), and Chan, et al. (2012) for the use of the SD tests in applications. We discuss the procedure briefly in this section.

To check whether there is any SD between $F$ and $G$, we could test the following hypothesis, for $n = 1, \cdots, 4$, $H_0 : F_n \equiv G_n$, against three alternatives

$$H_1 : F \neq_n G , \quad H_{1l} : F \succ_n G , \quad \text{and} \quad H_{1r} : F \prec_n G . \quad (2)$$

Let $\{f_i\}$ $(i = 1, 2, \cdots, N_f)$ and $\{g_i\}$ $(i = 1, 2, \cdots, N_g)$ be observations drawn from the independent random variables $Y$ and $Z$ with distribution functions $F$ and $G$, respectively.
The integrals $F^{(n)}$ and $G^{(n)}$ for $F$ and $G$ are defined in (1) for $n = 1, \cdots, 4$. For a grid of pre-selected points $\{x_k, k = 1, \cdots, K\}$, we propose to use the following modified $n$-order DD test statistic, $T_n(x)$ ($n = 1, \cdots, 4$) to test for $H_1$, $H_1l$, and $H_1r$:

$$T_n(x) = \frac{\hat{F}^{(n)}(x) - \hat{G}^{(n)}(x)}{\sqrt{\hat{V}_n(x)}},$$

(3)

where

$$\hat{V}_n(x) = \hat{V}_{F^{(n)}}(x) + \hat{V}_{G^{(n)}}(x), \quad \hat{H}_n(x) = \frac{1}{N_h(n-1)!} \sum_{i=1}^{N_h} (x - h_i)^{n-1},$$

$$\hat{V}_{H_n}(x) = \frac{1}{N_h} \left[ \frac{1}{N_h((n-1)!)^2} \sum_{i=1}^{N_h} (x - h_i)^{2(n-1)} - \hat{H}^{(n)}(x)^2 \right], \quad H = F, G; \quad h = f, g.$$

The modified DD test compares distributions at a finite number of grid points. Various studies examine the choice of grid points. For example, Tse and Zhang (2004) show that an appropriate choice of $K$ for reasonably large samples ranges from 6 to 15. Too few grids will miss information about the distributions. To solve this problem, Lean, et al. (2007), Wong, et al. (2008), Qiao, et al. (2012), and others suggest to use the 10 major partitions with 10 minor partitions within any two consecutive major partitions in each comparison and draw statistical inference. We recommend to use the simulated critical value suggested by Bai, et al. (2011) to make inference.

We note that by applying the DD test $T_n(x)$ in (3), one will know how many percent $F \succ_n G$ and how many percent $G \succ_n F$ significantly for $n = 1, \cdots, 4$. If $F \succ_n G$ is more than 50% while $G \succ_n F$ is less than the predetermined small value $\epsilon_0$ which is much smaller than $1/2$, then we conclude that $F \succ_n^\text{almost}(\epsilon_0) G$ for $n = 1, \cdots, 4$.

### 3.2 New ASD Test

The advantage of applying the modified DD test $T_n(x)$ in (3) is that one will know how many percent $F \succ_n G$ and how many percent $G \succ_n F$ significantly. This information could then be used to determine whether $F \succ_n^\text{almost}(\epsilon_0) G$ or $G \succ_n^\text{almost}(\epsilon_0) F$ for $n = 1, \cdots, 4$. However, the disadvantage of applying the DD test $T_n(x)$ in (3) is that it does not fit in the requirement of ASD in Definition 1. To circumvent this limitation, in this paper we propose to use another test for testing ASD. We will develop the test in this section to test the following hypothesis:

$$H_0 : \int_{S_n} \left| F^{(n)}(x) - G^{(n)}(x) \right| dx \leq \epsilon_0 \left\| F^{(n)} - G^{(n)} \right\|, \quad n = 1, \cdots, 4.$$
Similarly, we can have 

\[ \int_{S_n} [F^{(n)}(x) - G^{(n)}(x)] \, dx := I_1 = \int \max(F^{(n)}(x) - G^{(n)}(x), 0) \, dx. \]

Similarly, we can have 

\[ \int_{S_n} [F^{(n)}(x) - G^{(n)}(x)] \, dx := I_2 = \int \min(F^{(n)}(x) - G^{(n)}(x), 0) \, dx. \]

Through some computations, we can find that the null hypothesis is equivalent to 

\[ H_0 : (1 - \epsilon_0)I_1 + \epsilon_0 I_2 \leq 0, \quad n = 1, \cdots, 4. \]

As for \( F^{(n)}(x) \), note that 

\[ F^{(n)}(x) = \frac{1}{(n-1)!} \int_x^\infty (x-t)^{n-1} dF(t) = \frac{1}{(n-1)!} E(x - X)^{n-1} \]

here the function \( t \mapsto (t)_+ = \max(0,t) \).

Suppose now \( \{X_i\}_{i=1}^N \) and \( \{Y_i\}_{i=1}^M \) are independent random samples from distributions with \( F \) and \( G \), respectively. We also assume \( M/(N + M) = \hat{\lambda} \to \lambda \in (0,1) \). Then, we can estimate \( F^{(n)}(x) \) and \( G^{(n)}(x) \) by the following equations:

\[ \hat{F}^{(n)}(x) = \frac{1}{(n-1)!} \frac{1}{N} \sum_{i=1}^N (x-X_i)^{n-1} I(X_i \leq x), \quad \hat{G}^{(n)}(x) = \frac{1}{(n-1)!} \frac{1}{M} \sum_{i=1}^M (x-Y_i)^{n-1} I(Y_i \leq x). \]

When the estimates of \( F^{(n)}(x) \) and \( G^{(n)}(x) \) are obtained, we can estimate \( I_1 \) and \( I_2 \) by using the following:

\[ \hat{I}_1 = \int \max(\hat{F}^{(n)}(x) - \hat{G}^{(n)}(x), 0) \, dx, \quad \hat{I}_2 = \int \min(\hat{F}^{(n)}(x) - \hat{G}^{(n)}(x), 0) \, dx. \]

Then, the test statistic can be defined as follows:

\[ T_{N,M} = \sqrt{\frac{NM}{N + M}} [(1 - \epsilon_0)\hat{I}_1 + \epsilon_0 \hat{I}_2]. \quad (4) \]

When the value of \( T_{N,M} \) is very large, we reject the null hypothesis. The decision rule is that “rejecting \( H_0 \) if \( T_{N,M} > c_n \)”, where \( c_n \) is the critical value that will be discussed in next section. To state the properties of the test, we first introduce the following notations:

\[ T_1(x) = \sqrt{\frac{NM}{N + M}} \left[ (\hat{F}^{(n)}(x) - F^{(n)}(x)) - (\hat{G}^{(n)}(x) - G^{(n)}(x)) \right], \]
\[ \overline{T}_1(x) = \frac{1}{(n-1)!} \int_x^\infty (x-t)^{n-1} d \left( \sqrt{\lambda}B(F(t)) - \sqrt{1-\lambda}B(G(t)) \right), \]
\[ \overline{T} = \int \max(\overline{T}_1(x), 0) \, dx \]

where \( B(\cdot) \) is the standardized Brownian bridge on the interval \([0,1]\).

The following result characterizes the properties of our proposed test.
Theorem 1 Assume that $c_n$ is a positive finite constant, we have:

(A) if $H_0$ is true, then

$$\lim_{N,M \to \infty} P(\text{reject } H_0) \leq P(\overline{T} > c_n) \equiv \alpha(c_n);$$

(B) otherwise,

$$\lim_{N,M \to \infty} P(\text{reject } H_0) = 1.$$

4 Determination of critical values

4.1 Multiplier Methods

In this section, we propose to use the multiplier method which is also known as the non-parametric Monte Carlo method. As discussed by Van der Vaart and Wellner (1996), this method aims to simulate a process that is identical and independent to the original limiting asymptotic process. We discuss the Monte Carlo test procedure for determining the $p$-value as follows:

Step 1. Generate random variables $e_i^F(i = 1, 2, \cdots, N)$ and $e_i^G(i = 1, 2, \cdots, M)$ independently with zero mean and unit variance. Let $E_N^F := (e_1^F, \cdots, e_N^F)$, $E_N^G := (e_1^G, \cdots, e_M^G)$ and define the conditional counterpart of $T_1(x)$ as

$$T_1(E_N^F, E_M^G) = \hat{\lambda}^{1/2} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[ \frac{1}{(n-1)!}(x - X_i)^{n-1}I(X_i \leq x) - \hat{F}^{(n)}(x) \right] e_i^F$$

$$- (1 - \hat{\lambda})^{1/2} \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \left[ \frac{1}{(n-1)!}(x - Y_i)^{n-1}I(Y_i \leq x) - \hat{G}^{(n)}(x) \right] e_i^G.$$

The resultant conditional test statistic is

$$\overline{T}(E_N^F, E_M^G) = \int \max(T_1(E_N^F, E_M^G, x), 0)dx.$$

Step 2. Generate $m$ sets of $E_N^F, E_M^G$, say $E_N^{(i),F}, E_M^{(i),G}, i = 1, \cdots, m$ and get $m$ values of $\overline{T}(E_N^F, E_M^G)$, say $\overline{T}(E_N^{(i),F}, E_M^{(i),G}), i = 1, \cdots, m$.

Step 3. The $p$-value is estimated by $\hat{p}_k = n_k/(m+1)$, where $n_k$ is the number of $\overline{T}(E_N^{(i),F}, E_M^{(i),G})$ which is larger than or equal to $\overline{T}_{N,M}$. Reject $H_0$ when $\hat{p}_k \leq \alpha$ for a designed level $\alpha$.

Using the above multiplier method, we obtain the following theorem:

Theorem 2 Applying the multiplier method described above, we have
(A) if $H_0$ is true, then
\[ \lim_{N,M \to \infty} P(\text{reject } H_0) \leq \alpha ; \]

(B) otherwise,
\[ \lim_{N,M \to \infty} P(\text{reject } H_0) = 1 . \]

In order to compute the $p$-values in practice, we must deal with the fact that the integrals that define the relevant random variables must be calculated. Notice that $I_1 = E\left( \max(F^{(n)}(U) - G^{(n)}(U), 0) \right)(b - a)$, Here, $U$ follows a uniform distribution on $[a, b]$. Now, we generate $K$ independent random samples $\{U_j\}_{j=1}^K$ from $U(a, b)$, we then approximate $\hat{I}_1$ by
\[ \frac{1}{K} \sum_{j=1}^K \max(\hat{F}^{(n)}(U_j) - \hat{G}^{(n)}(U_j), 0)(b - a). \]
Similarly, we can approximate $T(E_F^N, E_G^M) = \int \max(T_1(E_F^N, E_G^M, x), 0)dx$ in the above algorithm by
\[ \frac{1}{K} \sum_{j=1}^K \max(T_1(E_F^N, E_G^M, U_j), 0)(b - a). \]
By using this procedure, one can make the approximation as accurate as one wants subject to one’s time and constraints on their computers.

4.2 Bootstrap Methods

Another approach to obtain the $p$-value simulation is to use bootstrap technique as described below:

**Step 1.** Draw a sample $\{X^*_i, i = 1, \cdots, N\}$ from the pooled sample $\{X_i, Y_j : i = 1, 2, \cdots, N; j = 1, 2, \cdots, M\}$ with replacement and draw another sample $\{Y^*_i, i = 1, \cdots, M\}$ in the same way.

**Step 2.** Compute $T_{N,M}^*$ in the same way as $T_{N,M}$ but with the bootstrapped samples $\{X^*_i, i = 1, \cdots, N\}$ and $\{Y^*_i, i = 1, \cdots, M\}$. Repeat Step 2 $m$ times to get $m$ $T_{N,M}^*$’s, denoted by $T_{N,M,i}^*(i = 1, 2, \cdots, m)$.

**Step 3.** The $p$-value is estimated by $\hat{p}_k = n_k/(m + 1)$, where $n_k$ is the number of $T_{N,M,i}^*$ which is larger than or equal to $T_{N,M}$. Reject $H_0$ when $\hat{p}_k \leq \alpha$ for a designed level $\alpha$.

Using the above bootstrap method, we obtain the following theorem:
**Theorem 3** Applying the bootstrap method described above, we have

(A) if $H_0$ is true, then
\[
\lim_{N,M \to \infty} P(\text{reject } H_0) \leq \alpha;
\]

(B) otherwise,
\[
\lim_{N,M \to \infty} P(\text{reject } H_0) = 1.
\]

At last, we link the relationship of the two tests we discussed in this paper in the following theorem:

**Theorem 4** For any $n = 1, \ldots, 4$, if we there are more than 50% of $T_n(x)$ in (3) confirms that $F \succeq_n G$ significantly for more than 50% while $F \prec_n G$ significantly for $\epsilon_0$ with $\epsilon_0$ much smaller than $1/2$, then using $T_{N,M}$ in (4) to test ASD will be satisfied for $\epsilon_0$.

5 Concluding Remarks

Leshno and Levy (2002) extend it to the stochastic dominance (SD) theory of almost stochastic dominance (ASD) for most decision makers. When comparing any two prospects, Guo, et al. (2013) find that there will be ASD relationship even there is only very little difference in mean, variance, skewness, or kurtosis. Investors may prefer to conclude ASD only if the dominance is nearly almost. Levy, et al. (2010) have provided two approaches to solve the problem. In this paper, we extend their work by first recommending an existing stochastic dominance test to handle the issue and thereafter developing a new test for the ASD which could detect dominance for any pre-determined small value. We also provide two approaches to obtain the critical values for our proposed test statistic.
References


