

2008

# The sizes and powers of some stochastic dominance tests: A Monte Carlo study for correlated and heteroskedastic distributions

Hooi Hooi Lean

Wing Keung Wong

Hong Kong Baptist University, awong@hkbu.edu.hk

Xibin Zhang

This document is the authors' final version of the published article.

Link to published article: <http://dx.doi.org/10.1016/j.matcom.2007.09.002>

---

## Citation

Lean, Hooi Hooi, Wing Keung Wong, and Xibin Zhang. "The sizes and powers of some stochastic dominance tests: A Monte Carlo study for correlated and heteroskedastic distributions." *Mathematics and Computers in Simulation* 79.1 (2008): 30-48.

This Journal Article is brought to you for free and open access by the Department of Economics at HKBU Institutional Repository. It has been accepted for inclusion in Department of Economics Journal Articles by an authorized administrator of HKBU Institutional Repository. For more information, please contact [repository@hkbu.edu.hk](mailto:repository@hkbu.edu.hk).

# The Sizes and Powers of Some Stochastic Dominance Tests: A Monte Carlo Study for Correlated and Heteroskedastic Distributions

**Hooi-Hooi Lean**

School of Social Sciences  
Universiti Sains Malaysia, 11800 USM, Penang, Malaysia

**Wing-Keung Wong\***

Department of Economics, Hong Kong Baptist University, Hong Kong  
Department of Economics, National University of Singapore, Singapore 119260

**Xibin Zhang**

Department of Econometrics and Business Statistics  
Monash University, Caulfield East, Victoria 3145, Australia

**First submission: June 2006**

**Revised: July 2007**

**Abstract:** Testing for stochastic dominance among distributions is an important issue in the study of asset management, income inequality, and market efficiency. This paper conducts Monte Carlo simulations to examine the sizes and powers of several commonly used stochastic dominance tests when the underlying distributions are correlated or heteroskedastic. Our Monte Carlo study shows that the test developed by Davidson and Duclos [9] has better size and power performances than two alternative tests developed by Kaur et al. [18] and Anderson [1]. In addition, we find that when the underlying distributions are heteroskedastic, both the size and power of the test developed by Davidson and Duclos [9] are superior to those of the two alternative tests.

**Keywords:** stochastic dominance, correlated distributions, heteroskedasticity, grid points.

**JEL classification:** C12, D31, G11

## **Acknowledgments**

We are grateful to Professor Robert Beauwens and anonymous referees for their substantive comments that have significantly improved this manuscript. The second author would like to thank Professors Robert B. Miller and Howard E. Thompson for their continual guidance and encouragement. The research is partially supported by the National University of Singapore and Monash University.

---

\* Correspondence to: W-K Wong, Department of Economics, National University of Singapore, Singapore 119260. Email: [ecswwk@nus.edu.sg](mailto:ecswwk@nus.edu.sg).

# 1 Introduction

Investors make choices from among different investment opportunities to maximize their expected utilities or wealth. One of the necessary conditions for optimizing behavior in economic analysis is that an optimal portfolio cannot be inferior to another feasible portfolio. This can be characterized in terms of stochastic dominance (SD) preference, which in turn provides a framework for analyzing investors' behaviors under uncertainty. Although the SD methodology has been developed for more than four decades, powerful SD tests have been available only recently. In the literature, Levy and Sarnat [21, 22], Joy and Porter [17], Wingender and Groff [35], and Seyhun [27] discuss the use of SD rules empirically, but they have not provided any testing procedure for SD.

There are two broad classes of SD tests. One is based on the minimum/maximum values, and the other is based on a set of grid points on the distributions. Varian [34] first establishes statistical tests for the hypothesis of utility maximization when distributions of variables are known, while McFadden [25] is the first paper to develop an SD test using the minimum/maximum statistic, followed by Klecan et al. (KMM) [19], and Kaur et al. (KRS) [18]. On the other hand, Anderson [1], Dardanoni and Forcina [7, 8], and Davidson and Duclos (DD) [9]<sup>1</sup> are the commonly used SD tests that compare the underlying distributions at a finite number of grid points. Nonetheless, the literature is rather silent on the performance of SD tests. Recently, Tse and Zhang [33] present Monte Carlo studies to examine the sizes and powers of some SD tests when the underlying distributions are independent and homoskedastic. However, the size and power performances of SD tests have not been explored when the underlying distributions are correlated or heteroskedastic.

---

<sup>1</sup> Chow [5] has done some earlier work similar to DD's in his unpublished thesis.

Zheng and Cushing [41] show that the conventional inference procedure usually requires samples to be independently drawn. Nevertheless, most economic or finance issues involve studies of dependent data.<sup>2</sup> This paper examines the size and power performances of three commonly used SD tests, namely, the DD, Anderson, and KRS tests, when the underlying distributions are correlated. In addition, we study the influence of heteroskedasticity on the sizes and powers of the SD tests, as well as the issue of choosing the number of grids being used for the SD tests. Tse and Zhang [33] show that the DD test has superior size and power performances compared to both the Anderson and KRS tests when the underlying distributions are independent. In this paper, we find that the DD test also demonstrates better performance than both the Anderson and KRS tests in terms of both size and power when the underlying distributions are correlated. In addition, we find that heteroskedasticity significantly reduces the powers of all the SD tests, but in this situation, the DD test still performs well for any large sample as it obtains higher power for any large sample.

In addition, as most of the SD tests involve a choice of the number of grids in the testing procedure, we follow the suggestion of Tse and Zhang [33] to use 10-15 equally distanced major grids. However, we suggest employing at least 10 equally distanced minor grids between any two consecutive major grids but using critical values with only the number of major grids to be their degrees of freedom. This will ensure that important information between any two major grids is not omitted without violating the independence assumption of the grids for the test statistic. The simulation study confirms that our suggestion is better.

The rest of the paper is organized as follows. Section 2 describes briefly both the SD theory and some commonly used SD tests, which will be examined in this paper. Section 3

---

<sup>2</sup> See, for example, the momentum puzzle discussed by Jegadeesh and Titman [15, 16] in which the portfolio of winners is highly correlated with the portfolio of losers.

explains the procedure of Monte Carlo simulations, and we show the results and discuss our findings in Section 4. Section 5 concludes.

## 2 Stochastic Dominance and Its Tests

### 2.1 Stochastic Dominance

Let  $F_1 = F$  and  $G_1 = G$  be the cumulative distribution functions (CDFs) of two random variables,  $X$  and  $Y$ , with means,  $\mu_F$  and  $\mu_G$ , respectively, and with a common support of  $[a, b]$ , where  $a < b$ ; we define the  $i^{\text{th}}$  order distribution functions<sup>3</sup> to be:

$$F_i(x) = \int_a^x F_{i-1}(t) dt, \quad G_i(x) = \int_a^x G_{i-1}(t) dt \quad (1)$$

for  $i = 2, 3$  where  $a \leq x \leq b$ . The random variable  $X$  is said to stochastically dominate  $Y$  at the first (second, third) order, denoted by  $X \succeq_1 Y$  or  $F \succeq_1 G$  ( $X \succeq_2 Y$  or  $F \succeq_2 G$ ,  $X \succeq_3 Y$  or  $F \succeq_3 G$ ), if and only if  $F_1(x) \leq G_1(x)$  ( $F_2(x) \leq G_2(x), F_3(x) \leq G_3(x)$ ) for all possible returns  $x$  and  $\mu_F \geq \mu_G$ .<sup>4</sup>

### 2.2 Stochastic Dominance Tests

McFadden [25] introduces a statistical test on empirical distributions using SD statistical methodology by assuming that the paired observations are drawn from two independent distributions and that the observations are independent over time. KMM extend the McFadden

---

<sup>3</sup> We also call them  $i^{\text{th}}$  order integrals. Refer to Wong [36], Wong and Chan [37], and Wong and Ma [39] for discussions of the definitions.

<sup>4</sup> Refer to Li and Wong [23], Wong and Li [38], Anderson [1,2], and Wong [36] for more discussions of the properties of the stochastic dominance relationship.

test by relaxing the independence assumption on the underlying distributions. The first- and second-order KMM test statistics are, respectively,

$$d^* = \min\left(\max_x \left[\hat{F}_1(x) - \hat{G}_1(x)\right], \max_x \left[\hat{G}_1(x) - \hat{F}_1(x)\right]\right),$$

and

$$s^* = \min\left(\max_x \left[\hat{F}_2(x) - \hat{G}_2(x)\right], \max_x \left[\hat{G}_2(x) - \hat{F}_2(x)\right]\right)$$

where  $\hat{F}_2$  and  $\hat{G}_2$  are defined in (2) below.

Nevertheless, Shin [29] points out that the KMM test is not robust to tails. He claims that the KMM test may produce results accepted in the first-order null hypothesis but rejected in the second-order null hypothesis, especially when the min-max of the CDF difference (or integrated CDF difference) is attained in the left tail of the distribution. This is because a few negative extreme observations may result in a very small value of  $d^*$  and a large value of  $s^*$ . Recently, Linton et al. [24] find that the sub-sampling bootstrap technique is better than the traditional bootstrap method when computing critical values of the KMM test.

On the other hand, Barrett and Donald [3] extend McFadden's test by developing a Kolmogorov-Smirnov (KS) test applied to two independent samples with possibly unequal sample sizes. The test statistic is:

$$\hat{K}_s = \left(\frac{N^2}{2N}\right)^{1/2} \sup_z \left[\hat{F}_s(z) - \hat{G}_s(z)\right],$$

where

$$\hat{F}_s(z) = \frac{1}{N(s-1)!} \sum_{i=1}^N (z - x_i)_+^{s-1}, \text{ and } \hat{G}_s(z) = \frac{1}{N(s-1)!} \sum_{i=1}^N (z - y_i)_+^{s-1} \quad (2)$$

for  $s = 1, 2$  and  $3$ ; and  $N$  is the sample size.

They point out that the test based on comparisons at a fixed number of grid points is potentially inconsistent, since only a subset of the restrictions implied by SD is considered. Nevertheless, Barrett and Donald [3] claim that their test is consistent, since it compares the objects at all points. However, Anderson [2] points out that under smoothness assumptions, the inconsistency problem is not substantive. Owing to the computation complexity of these methods, we do not include it in our Monte Carlo experiment.

In this paper, we shall investigate the sizes and powers of several well-known SD statistics under the following null hypothesis:

$$H_0 : F_s = G_s,$$

against any of the following alternatives:

$$\begin{aligned} H_A : F_s \neq G_s \text{ but } F \not\prec_s G, G \not\prec_s F, \\ H_{A1} : F \succ_s G, \\ H_{A2} : G \succ_s F, \end{aligned} \tag{3}$$

for  $s = 1, 2$  or  $3$ .<sup>5</sup> Since some SD tests are based on only the alternative  $H_{A1}$  or  $H_{A2}$  without considering  $H_A$ , we set their corresponding nulls to be

$$\begin{aligned} H_{A1}^0 : F \not\prec_s G \text{ and} \\ H_{A2}^0 : G \not\prec_s F. \end{aligned} \tag{3a}$$

### 2.2.1 Kaur, Rao and Singh (1994)

The KRS test is a two-sample test comparing  $F_2$  and  $G_2$  to test for the second-order SD under the null hypothesis,  $H_{A1}^0$ , in (3a) that  $X$  does not dominate  $Y$  at the second order versus the

---

<sup>5</sup> The value of  $s$  can be any positive integer, but we discuss only the order up to 3 in this paper.

alternative hypothesis,  $H_{A1}$ , in (3) with  $s = 2$  that  $X$  dominates  $Y$  at the second order. A similar argument could be made for testing  $H_{A2}^0$  versus  $H_{A2}$ . The test statistics are

$$T_M^2 = \inf_{a \leq x \leq b} \{Z_n(x)\}, \text{ and } T_{M^*}^2 = \inf_{a \leq x \leq b} \{-Z_n(x)\}, \quad (4)$$

in which

$$Z_n(x) = \frac{\hat{F}_2(x) - \hat{G}_2(x)}{\sqrt{\frac{1}{n} S_{n,F}^2(x) + \frac{1}{n} S_{n,G}^2(x)}}, \text{ and}$$

$$S_{n,H}^2(x) = \frac{1}{n} \sum_{i=1}^n [(x - z_i)_+ - H_2(x)]^2, z = x, y; H = F, G;$$

where  $\hat{F}_2$  and  $\hat{G}_2$  are defined in (2), and  $n$  is the sample size. KRS define  $Z_n(x) = 0$  when  $x$  is less than the minimum value of the observations from the combined sample. The null hypothesis,  $H_{A2}^0$ , will not be rejected if  $T_{M^*}^s < Z_\alpha$ , where  $Z_\alpha$  is the  $(1 - \alpha)^{\text{th}}$  percentile of the standard normal distribution.

The KRS test is consistent and has an upper bound  $\alpha$  on the asymptotic size. Allowing sample sizes to be unequal, it is evaluated over the full support of the population rather than at certain grid points. KRS develop the test only for  $s = 2$ . Tse and Zhang [33] extend the test to the situations when  $s = 1$  and 3. We follow KRS for  $s = 2$  and follow Tse and Zhang [33] for  $s = 1$  and 3 in our study.<sup>6</sup>

---

<sup>6</sup> One may refer to Tse and Zhang [33] for the KRS test statistic as well as their algorithms of finding minimum values when  $s = 1$  and 3.



### 2.2.2 Anderson (1996)

Anderson applies a trapezoidal rule to approximate the required integrals to estimate  $H_i$  in (1) for  $H = F$  or  $G$ . He proposes a non-parametric test<sup>7</sup> for the first three orders of SD with an omnibus test for the differences in two distributions. To use the test, we first partition the data into  $k$  mutually exclusive categories. Let  $p_{ij}$  be the probability of finding an observation in the  $i^{\text{th}}$  category of the  $j^{\text{th}}$  population for  $j = X, Y$ . Assume that  $x$  and  $y$  are the empirical frequency vectors with size  $n_X$  and  $n_Y$  drawn from populations  $X$  and  $Y$ , respectively. Then,

$$v(i) = \frac{x}{n_X} - \frac{y}{n_Y} \xrightarrow{d} N(0, m\Omega) \quad \text{and} \quad v'(m\Omega)^{-1}v \xrightarrow{d} \chi^2(k-1)$$

where  $m = \frac{(n_X + n_Y)/n}{n_X n_Y}$ .

$$\text{Define } I_f = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \cdot & 1 & 1 & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \quad \text{and} \quad I_F = 0.5 \begin{pmatrix} d_1 & 0 & 0 & \dots & 0 \\ d_1 + d_2 & d_2 & 0 & \dots & 0 \\ d_1 + d_2 & d_2 + d_3 & d_3 & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ d_1 + d_2 & d_2 + d_3 & d_3 + d_4 & \dots & d_k \end{pmatrix},$$

where  $d_j$  is the length of the  $j^{\text{th}}$  interval. The Anderson test is used to examine the hypotheses in (3) where  $F_s(x_i) - G_s(x_i)$  is replaced by  $I_f(p_X - p_Y)$ ,  $I_F I_f(p_X - p_Y)$  and  $I_F I_F I_f(p_X - p_Y)$  for  $s = 1, 2$  and  $3$ , respectively. For  $i = 1, 2, \dots, k$ , the test statistic is

$$A_i^s = \frac{H_s v(i)}{\sqrt{m H_s \Omega^* H_s'(i, i)}} \xrightarrow{d} N(0, 1), \quad (5)$$

<sup>7</sup> Note that the test assumes independent sampling.

where  $H_s = I_f$ ,  $I_F I_f$  or  $I_F I_F I_f$ , represent the first-, second- and third-order SD test statistics, respectively, for  $s = 1, 2$ , and  $3$ , and  $\Omega^*$  is the estimate of  $\Omega$  such that

$$\Omega = n \begin{pmatrix} p_1(1-p_1) & -p_1 p_2 & \cdot & \cdot & -p_1 p_k \\ -p_2 p_1 & p_2(1-p_2) & \cdot & \cdot & -p_2 p_k \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -p_k p_1 & -p_k p_2 & \cdot & \cdot & p_k(1-p_k) \end{pmatrix},$$

in which  $p$  is estimated by  $p^* = \frac{x+y}{n_X + n_Y}$ .

### 2.2.3 Davidson and Duclos (2000)

DD derive the asymptotic sampling distribution of various estimators to test any order of SD for any two random variables,  $X$  and  $Y$ , with CDFs,  $F$  and  $G$ , based on conditional moments of distributions, in which the size is well controlled by a studentized maximum modulus (SMM). A variety of different hypotheses shown in (3) are defined with respect to a finite set of values for  $z_i$ . Suppose that we have  $N$  pairs of observations,  $x_i$  and  $y_i$ , from the two random variables,  $X$  and  $Y$ , with their  $s$ -order distribution functions,  $F_s$  and  $G_s$ , defined in (1). For  $s = 1, 2$ , and  $3$ , the corresponding DD test statistic is

$$T_s(z) = \frac{\hat{F}_s(z) - \hat{G}_s(z)}{\sqrt{\hat{V}_s(z)}}, \quad (6)$$

where  $\hat{F}_s(z)$  and  $\hat{G}_s(z)$  are defined in (2), respectively, and

$$\begin{aligned}\hat{V}_s(z) &= \hat{V}_{F_s}(z) + \hat{V}_{G_s}(z) - 2\hat{V}_{FG_s}(z); \\ \hat{V}_{H_s}(z) &= \frac{1}{N} \left[ \frac{1}{N((s-1)!)^2} \sum_{i=1}^N (z-h_i)_+^{2(s-1)} - \hat{H}_s(z)^2 \right], H = F, G; h = x, y; \\ \hat{V}_{FG_s}(z) &= \frac{1}{N} \left[ \frac{1}{N((s-1)!)^2} \sum_{i=1}^N (z-x_i)_+^{s-1} (z-y_i)_+^{s-1} - \hat{F}_s(z)\hat{G}_s(z) \right].\end{aligned}$$

To test for SD, the distributions should be examined for the full support, but this is empirically impossible. Thus,  $H_0$  is tested only on a selected finite number of values for  $z$  in the distribution. Bishop et al. [4] suggest applying a union-intersection test for multiple comparisons. One advantage of using the multiple comparison approach is that it is easier to deal with dependent samples (Barrett and Donald, [3]). The significance of the test is determined asymptotically by the critical value of the SMM distribution with  $k$  and  $\infty$  degrees of freedom and the following decision rules will be adopted:

$$\begin{aligned}\text{If } |T_s(z_i)| &< M_{\infty, \alpha}^k \text{ for } i=1, \dots, k, \text{ accept } H_0; \\ \text{If } T_s(z_i) &< M_{\infty, \alpha}^k \text{ for all } i \text{ and } -T_s(z_i) > M_{\infty, \alpha}^k \text{ for some } i, \text{ accept } H_{A1}; \\ \text{If } -T_s(z_i) &< M_{\infty, \alpha}^k \text{ for all } i \text{ and } T_s(z_i) > M_{\infty, \alpha}^k \text{ for some } i, \text{ accept } H_{A2}; \text{ and} \\ \text{If } T_s(z_i) &> M_{\infty, \alpha}^k \text{ for some } i \text{ and } -T_s(z_i) > M_{\infty, \alpha}^k \text{ for some } i, \text{ accept } H_A.\end{aligned}\tag{7}$$

### 3 Monte Carlo Simulation

In this section, we employ the Monte Carlo simulation to study the sizes and powers of the DD, Anderson, and KRS tests. The simulation is carried out according to the following linear return-generating process<sup>8</sup>:

$$r_{jt} = \ln \left( \frac{p_{j,t}}{p_{j,t-1}} \right) = \alpha_j + \beta_j r_{pt} + e_{jt}, \quad r_{pt} \text{ iid } N(0,1),\tag{8}$$

---

<sup>8</sup> Chow [6] uses a similar process in his simulation study and indicates that SD does not rely on the linear return-generating process.

where  $p_{j,t}$  and  $r_{jt}$  are the price and the corresponding return for the  $j^{th}$  stock or portfolio at time  $t$ ,  $e_{jt}$  is the error term, and  $r_{pt}$  is the return from the market index, say, for example, the Dow Jones Industrial Index. This is a simple and well-known single-factor market model (Sharpe [28]). In this paper, without loss of generality, we assume  $r_{pt}$  *iid* (independent and identically distributed as)  $N(0,1)$  and  $corr(e_{jt}, r_{pt}) = 0$ . We examine the sizes and powers of several SD tests under both homoskedastic and heteroskedastic situations, in which

$$e_{jt} \text{ iid } N(0,1) \tag{9a}$$

or

$$e_{jt} = \sigma_t \gamma_t, \tag{9b}$$

where  $\gamma_t \text{ iid } N(0,1)$ ,  $\ln \sigma_t^2 = \frac{1}{2} \theta \ln \sigma_{t-1}^2 + \tau_t$ , and  $\tau_t \text{ iid } N(0,1)$ . Specifically, we consider the following situations for the returns  $r_{it}$  and  $r_{jt}$  generated by the linear return-generating process in (8) such that:

- (i)  $e_i$  and  $e_j$  are independent; and
- (ii)  $e_i$  and  $e_j$  are dependent.

In the first case,  $e_i$  follows (9a) if it is homoskedastic and follows (9b) if it is heteroskedastic. In the second case,

$$e_i = e_j \rho + a \sqrt{1 - \rho^2}, \tag{10}$$

where  $a \text{ iid } N(0,1)$ ,  $corr(a, e_j) = 0$ ,  $\rho = corr(e_i, e_j)$ , and both  $e_i$  and  $e_j$  can be either homoskedastic or heteroskedastic.

The returns,  $r_i$  and  $r_j$ , generated from the return-generating process in (8), are dependent when both  $\beta_i$  and  $\beta_j$  are non-zero and/or  $e_i$  and  $e_j$  are dependent. They are independent when both  $\beta_i$  and  $\beta_j$  are zero and the error terms,  $e_i$  and  $e_j$ , are independent. Since the independent case with a homoskedastic process has been studied by Tse and Zhang [33], in this paper, we report only the situations in which the returns,  $r_i$  and  $r_j$ , are dependent and either homoskedastic or heteroskedastic.<sup>9</sup>

It is well-known that (i) if  $\alpha_i > \alpha_j$  and  $\beta_i = \beta_j$ , then  $r_i \succ_1 r_j$ ; and (ii) if  $\alpha_i = \alpha_j$  and  $\beta_i < \beta_j$ , then  $r_i \succ_2 r_j$ .<sup>10</sup> In this paper, we test for the SD relationship between the series with different values of  $\alpha$  and  $\beta$  to cover the first two orders of SD. With the hierarchical relationship of SD, the existence of the second-order SD implies the existence of the third-order SD. Thus, we note that the third-order SD relationship could be obtained by using (ii) in (10) and therefore it is not necessary to include another parameter for constructing the third-order SD relationship. Figures 1.1 to 2.3 depict the different order integrals defined in (1) for the returns,  $r_i$  and  $r_j$ , defined in (8) with different values of  $\alpha$  and  $\beta$ .

Figure 1.1 (1.2, 1.3) exhibits the first- (second-, third-) order integrals for the returns,  $r_i$  and  $r_j$ , generated by the linear return-generating process in (8) with  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$  and  $\beta_i = \beta_j = 1.0$ . These figures show that  $r_i$  stochastically dominates  $r_j$  at all of the first three orders. On the other hand, Figure 2.1 (2.2, 2.3) depicts the first- (second-, third-) order integrals when  $\alpha_i = \alpha_j = 0.0$ ,  $\beta_i = 0.0$  and  $\beta_j = 2.0$ . Figure 2.1 reveals no first-order SD between the two

---

<sup>9</sup> We have also conducted simulations for the independent situation with heteroskedastic processes. Since the results draw conclusions similar to those for the dependent situation, we skip reporting the results, which are available on request.

<sup>10</sup> One may refer to Hadar and Russell [13], Tesfatsion [32], and Li and Wong [23] for the proofs.

series, but Figure 2.2 and Figure 2.3 show that  $r_i$  stochastically dominates  $r_j$  at the second and third orders, respectively. The greater the difference between  $\alpha_i$  and  $\alpha_j$  and between  $\beta_i$  and  $\beta_j$ , the more significant the domination between the two series. We note that the values of  $\alpha$  and  $\beta$  are chosen from the range of the estimated means and the estimated standard deviations of all the industrial indices from the US market, so that all  $r_i$  and  $r_j$  studied in this paper could represent the returns of some empirical portfolios. In addition, our study covers both homoskedastic and heteroskedastic processes and hence our simulation could cover most, if not all, of the empirical situations.

We report only the results obtained when  $e_i$  and  $e_j$  are dependent. The value of  $\rho$  used in our simulation is 0.8, which is the empirical correlation coefficient between portfolios of winners and losers in the US stock market during the period from 1965 to 2000. We also use other values of  $\rho$  and find that the results are not sensitive to different values of  $\rho$ . Thus, we report only the dependent situations with  $\rho = 0.8$ .

Since there is no definite solution theoretically for the optimal number of grid points to maximize the power of an SD test, the number of grid points used in empirical studies is often decided by a rule of thumb. Zheng et al. [42] recommend using 10 to 15 ordinates in empirical applications for the SD tests. More information on the distributions would be revealed by using more ordinates, but the independence assumption of the grid statistics would then be violated by using excessive grid points (Stoline and Ury, [31]). Moreover, as pointed out by Sidak [30] and Hochberg [14], the SMM test is conservative if these statistics are not independent. On the other hand, too sparse a number of grid points would miss the detection of SD behavior for the distributions between any two consecutive grids. Ideally, we would choose a sufficiently large number of grids to reveal more information without compromising the grid points' independence

property. In view of this, we follow Tse and Zhang's [33] recommendation to use  $k$  major intervals, where  $k = 6, 10,$  and  $15,$  but we suggest that each major interval be further partitioned into 10 (or more) equally distanced minor intervals to be examined.<sup>11</sup> Examining these 10 equally distanced minor intervals between any two consecutive major grids ensures non-omission of important information between any consecutive major grid points.<sup>12</sup> Based on our computing experience, distributions with minor intervals possess better sizes and powers than distributions without minor intervals.<sup>13</sup> Hence, without the imposition of minor grids, information within the minor intervals could be omitted, which may, in turn, alter the decision. Nevertheless, we cannot use the critical values with their degrees of freedom based on  $10k$  grids (total grid points include the end points of both major and minor intervals), since this will violate the independence assumption of different grids.<sup>14</sup> Thus, we suggest using the critical values in Stoline and Ury [31] with degrees of freedom based only on  $k$  (instead of  $10k$ ) major grids for all the major and minor grids to satisfy the independence assumption of the grids.

## 4 Results and Discussions

We conduct 10,000 iterations for each simulation scheme, where the sample sizes are, respectively, 50, 100, 500, and 1000. The Monte Carlo simulation is conducted on dependent

---

<sup>11</sup> Here, we construct a simple example to show that checking minor grids is necessary. Let  $X=i$  for  $i=2, 4, \dots, 20$  with equal probability  $1/10$  on each point and  $Y=i$  for  $i=1, 3, \dots, 19$  with equal probability  $1/10$ . A choice of 10 major grids on  $\{2, 4, \dots, 20\}$  will never distinguish  $X$  from  $Y$ , whereas one will always be able to distinguish  $X$  from  $Y$  if any other minor grid is being examined.

<sup>12</sup> In the simulation, the two generated series are combined and sorted in ascending order. This combined sample is used to determine the grid points. The range of the combined sample is divided into  $k$  major intervals, and each major interval is then further partitioned into 10 minor equally distanced intervals.

<sup>13</sup> The results are available on request.

<sup>14</sup> Let's say  $k=10$ , then  $10k$  becomes 100. If we choose 100 grids, we cannot use  $M_{\infty, \alpha}^{100}$  (see (7)) as a critical value because this will violate the independence assumption of the grids for the test statistic, since any two consecutive grids are then too close to be independent (Stoline and Ury [31]; Richmond [26]). One may refer to Fong et al. [11], Lean et al. [20], Gasbarro et al. [12], and Wong et al. [40] for the empirical studies using both major and minor grids in their analyses.

returns,  $r_{it}$  and  $r_{jt}$ , in (8) generated by setting different values of  $\alpha_i$  and  $\alpha_j$  and/or different values of  $\beta_i$  and  $\beta_j$ ; and by adopting innovations,  $e_{it}$  and  $e_{jt}$ , which are either homoskedastic (see (9a)) or heteroskedastic (see (9b)), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).<sup>15</sup> We first investigate the size performance of each SD test, then investigate the powers of these tests with different values of  $\alpha_i$  and  $\alpha_j$  and/or different values of  $\beta_i$  and  $\beta_j$ .

#### 4.1 Homoskedastic Errors

Table 1 reports the empirical sizes<sup>16</sup> of the DD, Anderson, and KRS tests under the null hypothesis,  $H_0$ , in (3) for the dependent returns,  $r_{it}$  and  $r_{jt}$ , in (8) generated by setting  $\alpha_i = \alpha_j = 0.0$  and  $\beta_i = \beta_j = 1.0$  with their innovations,  $e_{it}$  and  $e_{jt}$ , following the homoskedastic process shown in (9a) where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10). Our simulation results show that all three tests are conservative, since their empirical sizes are less than 0.05.<sup>17</sup> From the table, we find that the empirical size of the DD test is closer to the critical values, whereas the empirical sizes of both the KRS and Anderson tests are close to zero. These results are consistent with the results reported in Tse and Zhang [33] for independent returns generated by a homoskedastic process.

In addition, from the table, we find that the size of the DD test decreases as  $k$  increases from 6 to 15, and the size of the DD test decreases from the first order to the third order (except

---

<sup>15</sup> We report only the results with  $\rho = 0.8$  in this paper, since we find that the simulated results are not sensitive to different values of  $\rho$ . The results for different values of  $\rho$  are available on request.

<sup>16</sup> The empirical size is estimated by calculating the percentage of rejections of the null hypothesis when the null hypothesis is true. The DD test follows the SMM distribution, while both the Anderson and KRS tests follow the standardized normal distribution. At the nominal size of 0.05, the SMM critical values of  $k$  for 6, 10, and 15 are 2.928, 3.254 and 3.48,7 respectively, while the  $Z_{0.025}$  is 1.96 for the KRS test.

<sup>17</sup> Except the cases for the DD test with  $s = 1$ ,  $k = 6$ , and  $N = 100, 500$  and 1000.



for sample size 50 at  $k = 15$ ). Overall, the DD test's biggest (smallest) size is 0.1021 (0.0009) in our simulation. Comparing the DD test with the Anderson and KRS tests, the size of the DD test is, in general, the closest to the nominal size of 0.05, especially when  $N = 1000$ ,  $k = 10$ , and  $s = 1$ . The effect of sample size on the DD test's size performance is inconsistent. For example, the size performance increases when sample size becomes larger for the first-order SD. However, the size performance decreases and then increases as samples become larger for the second-order SD. For the KRS test, its biggest (smallest) size is 0.0059 (0.0000) when the sample size is 1000 (50). In contrast to the DD test, the KRS test's size increases from the first order to the third order. With exception of the case for  $s = 1$  and  $N = 50$ , in which the size of the Anderson test increases as  $k$  increases, all the other sizes of the Anderson test in our simulation are either zero or very close to zero. Thus, we conclude that the DD test performs better than the Anderson and KRS tests in size performance, since the empirical size of the DD test is the closest to the nominal critical values.

To examine the power performance<sup>18</sup> of the SD tests, we first investigate two return series,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$  and  $\beta_i = \beta_j = 1.0$ , respectively. In this case,  $r_i$  stochastically dominates  $r_j$  at all orders. We summarize the corresponding empirical powers for the DD and Anderson tests in Table 2a and summarize that for the KRS test in Panel A of Table 3 for the homoskedastic process. In this situation, the DD test attains the highest power among all the three SD tests being studied. As expected, the power is smaller for small sample size like 50 and 100, but it increases as the sample size increases for all the tests. However, as the sample size increases, the power of the DD test increases faster than those of the

---

<sup>18</sup> The empirical power is estimated by the probability of rejecting the null hypothesis when the null hypothesis of no dominance is false or, equivalently, estimated by the probability of failing to reject the alternative hypothesis when the alternative hypothesis is true.

Anderson and KRS tests. Davidson-Duclos [10] note that one cannot accept the alternative hypothesis in a statistical test when the null is rejected unless the alternative includes everything not included in the null. There is much outside  $H_0$  that is not included in  $H_{A1}$ .<sup>19</sup>

We then examine the case in which the two return series,  $r_{it}$  and  $r_{jt}$ , are generated by setting  $\alpha_i = \alpha_j = 0.0$ ,  $\beta_i = 0.0$ ,  $\beta_j = 2.0$ , respectively. In this situation, there is no first-order SD between  $r_i$  and  $r_j$  but  $r_i$  stochastically dominates  $r_j$  at the second and third orders. Under this situation, we depict the corresponding empirical powers of the DD and Anderson tests in Table 2b and display the corresponding empirical power of the KRS test in Panel B of Table 3 for the homoskedastic process.<sup>20</sup>

We first examine the power for  $s = 1$  in which the first-order SD does not exist. Table 2b shows that the empirical powers for all the SD tests being studied in our paper are very high with nearly all having values of one, especially the large sample. This strongly rejects the hypothesis that the cumulative distribution functions of  $r_i$  and  $r_j$  are equal. Panel B of Table 3 shows that the empirical relative frequencies that the hypothesis  $H_{A2}^0$  is not rejected for  $s = 1$  are one for all samples, implying that  $r_j$  does not stochastically dominate  $r_i$  at the first order. On the other hand, the empirical relative frequencies that the hypothesis  $H_{A1}$  is not rejected for  $s = 1$  are zero for all samples, implying that  $r_j$  does not stochastically dominate  $r_j$ . Table 2b reveals that the empirical power for the DD test is slightly larger than that for the Anderson test. We note that we

---

<sup>19</sup> We would like to show our appreciation to an anonymous referee for pointing out this issue.

<sup>20</sup> In Table 2b, we compute the empirical power for  $s=1$  to be the estimated probability of failing to reject the alternative hypothesis  $H_A$ , since FSD does not exist. For  $s=2$  and  $s=3$ , we compute their empirical powers to be the estimated probability of failing to reject the alternative hypothesis  $H_{A1}$ , since both second- and third-order SD exist.

cannot compare the powers between the KRS test and the DD test or Anderson test because the former uses  $H_{A2}^0$  as the null, whereas the latter uses  $H_0$  as the null and  $H_A$  as the alternative.<sup>21</sup>

We then examine the empirical powers of the SD tests for  $s = 2$  and 3 in which both second- and third-order SD exist. From Table 2b and Panel B of Table 3, we find that the DD test has higher power compared to both the Anderson and KRS tests in small sample sizes of 50 and 100. The powers of all the tests attain one in sample sizes of 500 and 1000. In both situations, as expected, the power increases significantly as the sample size increases. Nonetheless, it is interesting to note from Table 2b that the powers of both the DD and Anderson tests decrease as  $k$  increases, i.e., the power is actually higher for smaller  $k$ . These results could possibly be due to the violation of the independence assumption of grids for larger  $k$ . In addition, our simulation shows that smaller order yields bigger power.

Overall, our simulation results conclude that for the homoskedastic process, the DD test is superior to both the KRS and Anderson tests because (1) the size of the DD test is closer to the nominal one and (2) the DD test possesses a higher power.

## 4.2 Heteroskedastic Errors

In order to understand the size and power performances of the SD tests when the series being examined are heteroskedastic, we conduct simulation schemes by allowing the errors in (8) to be heteroskedastic. Our purpose is to check whether the size and the power have changed when the assumption of homoskedasticity is violated.

---

<sup>21</sup> We note that, for  $s=1$ , in Table 3, the empirical relative frequencies of not rejecting the hypothesis  $H_{A1}$  are all zero, implying that return  $r_{it}$  does not dominate  $r_{jt}$  at the first order.

Table 4 presents the empirical sizes of the DD, Anderson, and KRS tests for the heteroskedastic process. Similar to the results for the homoskedastic process, our simulation results show that both the Anderson and KRS tests in all orders and the DD test in the second and third orders are conservative under the heteroskedastic process, since their sizes are far less than the nominal size of 0.05 under all situations. The empirical size of the Anderson test is zero for the second- and third-order and close to zero for the first-order SD, while the size of the KRS test is zero for the first-order and is close to zero for the second- and third-order SD. The DD test is not conservative for the first-order SD, since its empirical size is greater than the nominal size of 0.05 in seven out of twelve situations, and the average size for the first-order SD is 0.0786. The size of the DD test decreases as  $k$  increases and as the order ascends from the first to the third.

Compared with the results for the homoskedastic process in Table 1, the simulation results in Table 4 show that heteroskedasticity has no impact on the size of the Anderson test for all orders and no impact on the size of the KRS test for the first-order SD, since its size is zero or close to zero.<sup>22</sup> Nevertheless, we find that heteroskedasticity causes the empirical size to be larger for the DD test in all orders, and similarly for the KRS test in the second and third orders, especially for large sample sizes of 500 and 1000. Nonetheless, this effect is not serious and does not affect the performance of the tests, since their sizes are still much smaller than the nominated significance level. Heteroskedasticity creates the problem of under-rejection of  $H_0$  in the first-order SD of the DD test, especially for the case with 6 or 10 grids in the large sample, since the empirical size becomes as large as 0.2376 for  $N = 1000$  and  $k = 6$ . Table 4 shows that a choice of grids of either 10 or 15 yields better size in the simulation. Table 4 shows that the best simulated

---

<sup>22</sup> In fact, we cannot conclude that heteroskedasticity has no impact on the sizes of the Anderson and KRS tests because their sizes are zero. If these tests could be improved so that their sizes are close to the nominated one, the impact of the heteroskedastic effect on them could then become meaningful.

size is 0.0551 for the DD test when  $N = 500$  and  $k = 15$  in which its simulated size is the closest to the nominated size of 0.05.<sup>23</sup>

Tables 5a and b summarize the empirical powers of the DD and Anderson tests, while Table 6 exhibits the power of the KRS test for the heteroskedastic process. Our simulation results show that the DD test obtains the highest power among the three SD tests under all situations. As a result, we claim that the DD test is superior to both the Anderson and KRS tests in the power performance for the heteroskedastic process. When the first-, second- and third-order SD exists (Table 5a and Panel A of Table 6), the Anderson test has the lowest power, which can be ignored in small sample sizes like 50 and 100. On the other hand, when the second- and third-order SD exists in the absence of the first-order SD<sup>24</sup> (Table 5b and Panel B of Table 6), the KRS test has the lowest power. Hence, the Anderson and KRS tests do not dominate each other in terms of power performance. In contrast, the DD test performs very well for all orders for the sample size above 100, especially when the sample size increases to 500 or 1000, with its power approaching one. As expected, the power for the three SD tests increases significantly as the sample size increases. The power of both the DD and Anderson tests decreases as  $k$  increases. A possible explanation of this phenomenon is that it could be due to the violation of the independence assumption of grids for large  $k$ . In addition, their powers are bigger for smaller orders.

Compared with the results in the homoskedastic situations shown in Table 2, the simulation results for the heteroskedastic series in Table 5 show that the power of the DD test

---

<sup>23</sup> We note that it could be a drawback for the DD test that some of its simulated sizes are greater than the nominated size of 0.05. However, unlike the situations in both the Anderson and KRS tests, in which their sizes are always far below the nominal significance level (actually they are zero or close to zero), for each  $N$ , one could still be able to find some values of  $k$  such that the simulated size of the DD test is closer to the nominated size. Thus, one could still claim that the DD test performs relatively better, though there are rooms for the DD test to be improved.

<sup>24</sup> Since the conclusion drawn for  $s = 1$  under the heteroskedastic process is the same as that drawn under the homoskedastic process as discussed in Section 4.1, we skip the discussion for the situation of  $s = 1$  under the heteroskedastic process.

decreases significantly under heteroskedasticity when the first three orders of SD exist, as well as when the first-order SD is absent but the second- and third-order SD exists.

Figures 3 and 4 produce evidence that heteroskedasticity reduces the power of the DD tests for different values of  $\alpha$  and  $\beta$  defined in (8). The figures for the Anderson and KRS tests under the homoskedastic and heteroskedastic processes for different values of  $\alpha$  and  $\beta$  lead us to draw the same conclusion for the Anderson and KRS tests.<sup>25</sup> Nevertheless, the DD test still attains good power to detect the dominance alternative for sample sizes larger than 500. On the other hand, the power for the Anderson test is inconsistent, especially for small sample sizes when SD exists in all three orders. If we compare the results in Table 3 with those in Table 6, the power of the KRS test reduces significantly under the heteroskedastic process. Thus, we conclude that heteroskedasticity also reduces the power of the KRS test.<sup>26</sup>

We also examine the power for different values of  $\alpha$  and  $\beta$ . By fixing  $\beta_i = \beta_j = 1$  and  $\alpha_i = 0$ ,<sup>27</sup> we vary the values of  $\alpha_j$  from -0.2 to 0.2. The relative values of  $\alpha_i$  and  $\alpha_j$  will affect the two distributions in terms of the FSD. The wider the difference between  $\alpha_i$  and  $\alpha_j$ , the more significant SD becomes. The power first decreases and then increases as  $\alpha_j$  goes from -0.2 to 0.2. Figure 5 clearly shows that the DD test has the best power for all values of  $\alpha_j$ . In a separate exercise, we fix  $\alpha_i = \alpha_j = 0$  and  $\beta_j = 1$ <sup>28</sup> and vary the values of  $\beta_i$  from 0 to 2. The relative values of  $\beta_i$  and  $\beta_j$  will alter the second-order SD (hence, altering the third-order SD)

---

<sup>25</sup> As the shapes of the figures are similar to those of Figures 3 and 4, we skip reporting the figures, which are available on request.

<sup>26</sup> The plots of the Anderson and KRS tests are available on request.

<sup>27</sup> We skip reporting the plots of other values of  $\alpha_i$ ,  $\beta_i$  and  $\beta_j$ , since all the plots will lead us to draw the same conclusion.

<sup>28</sup> We skip reporting the plots of other values of  $\alpha_i$ ,  $\alpha_j$ , and  $\beta_i$ , since all the plots will lead us to draw the same conclusion.

relationship between the distributions. We find that the wider the difference between  $\beta_i$  and  $\beta_j$ , the bigger the respective gap of the second-order SD (and the third-order SD), with its power first increasing slightly, then decreasing, then increasing again as  $\beta_i$  goes from 0 to 2. Figure 6 clearly shows that the DD test obtains the best power for all the values of  $\beta_i$ .

## 5 Conclusion

This paper investigates the sizes and powers of three commonly used SD tests, namely, the DD, Anderson, and KRS tests, when the underlying distributions are correlated or heteroskedastic. Our findings add to the literature by offering a better understanding of the sizes and powers of several SD tests. Although all three tests are conservative, the DD test is found to be the best among the three in our simulation in terms of both size and power performances for correlated distributions. While heteroskedasticity significantly reduces the powers of all the SD tests in our studies, the power of the DD test is still reasonably good for large samples.

Our Monte Carlo study shows that when the sample size is at least 500 and the number of major grid points is around 10 to 15, all three SD tests have reasonably good power performance, with the DD test being the best compared to the other two. To ensure that important information is not omitted in the underlying distributions, we suggest partitioning 10 equally distanced grids for each major interval and applying the DD test on all the grids but using the critical values with the number of major grids to be their degrees of freedom. This will ensure that important information between any two major grids is not omitted without violating the independence assumption of the grids for the test statistic. Our simulation study supports our suggestion.

Finally, we note that the poor performance of the KRS test could be due to the fact that sample points are rare in the tails of the distribution. Therefore, there is not enough information to get a significant test statistic. If the infimum is computed over the whole pooled empirical distribution, the infimum is necessarily small, with the result that rejections are very unlikely. To circumvent this problem, one may consider modifying the KRS test by adopting the approach in Davidson and Duclos [10], who suggest testing for restricted stochastic dominance, for instance, by computing the infimum for the modified KRS test over an interval strictly inside the support of the pooled empirical distribution.<sup>29</sup> The performance of the modified KRS test could then be improved in the simulations.<sup>30</sup> We also note that the results of our simulation have been widely used to compare different assets' performance or to explain some financial anomalies.<sup>31</sup>

---

<sup>29</sup> We would like to show our appreciation to an anonymous referee for pointing out this issue.

<sup>30</sup> Since developing a modified KRS test is beyond the scope of this paper, the significance of its performance should clearly be scrutinized carefully in a separate paper when the modified KRS test is available.

<sup>31</sup> See, for example, Fong et al. [11], Lean et al. [20], Gasbarro et al. [12], and Wong et al. [40] for more information.



## References

- [1] G. Anderson, Nonparametric tests of stochastic dominance in income distributions, *Econometrica* 64 (1996) 1183 – 1193.
- [2] G. Anderson, Making inferences about the polarization, welfare and poverty of nations: A study of 101 countries 1970-1995, *Journal of Applied Econometrics* 19(5) (2004) 537-550.
- [3] G.F. Barrett, S.G. Donald, Consistent test for stochastic dominance, *Econometrica* 71(1) (2003) 71-104.
- [4] J.A. Bishop, J.P. Formby, P.D. Thistle, Convergence of the South and non-South income distributions, *American Economic Review* 82(1) (1992) 262-272.
- [5] K.V. Chow, Statistical inference for stochastic dominance: a distribution free approach, Ph.D. Thesis, University of Alabama (1989).
- [6] K.V. Chow, Marginal conditional stochastic dominance, statistical inference, and measuring portfolio performance, *Journal of Financial Research* 24(2) (2001) 289 – 307.
- [7] V. Dardanoni, A. Forcina, A unified approach to likelihood inference on stochastic orderings in a nonparametric context, *Journal of the American Statistical Association* 93 (1998) 1112 – 1123.
- [8] V. Dardanoni, A. Forcina, Inference for Lorenz curve orderings, *Econometrics Journal* 2 (1999) 49 – 75.
- [9] R. Davidson, J.Y. Duclos, Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica* 68(6) (2000) 1435-1464.
- [10] R. Davidson, J.-Y. Duclos, Testing for restricted stochastic dominance, mimeo, McGill University (2006).
- [11] W.M. Fong, W.K. Wong, H.H. Lean, International momentum strategies: A stochastic dominance approach, *Journal of Financial Markets*, 8 (2005) 89-109.
- [12] D. Gasbarro, W.K. Wong, J.K. Zumwalt, Stochastic dominance analysis of iShares, *European Journal of Finance*, 13 (2007) 89-101.
- [13] J. Hadar, W.R. Russell, Stochastic dominance and diversification, *Journal of Economic Theory* 3 (1971) 288-305.
- [14] Y. Hochberg, Some generalizations of the T-method in simultaneous inference, *Journal of Multivariate Analysis* 4 (1974) 224 - 234.

- [15] N. Jegadeesh, S. Titman, Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance* 48 (1993) 65 - 91.
- [16] N. Jegadeesh, S. Titman, Profitability of momentum strategies: an evaluation of alternative explanations, *Journal of Finance* 56 (2001) 699 - 720.
- [17] O.M. Joy, R.B. Porter, Stochastic dominance and mutual fund performance, *Journal of Financial and Quantitative Analysis* 9 (1974) 24-31.
- [18] A. Kaur, B.L.S.P. Rao, H. Singh, Testing for second order stochastic dominance of two distributions, *Econometric Theory* 10 (1994) 849-866.
- [19] L. Klecan, R. McFadden, D. McFadden, A robust test for stochastic dominance, Working Paper, MIT & Cornerstone Research (1991).
- [20] H.H. Lean, R. Smyth, W.K.Wong, Revisiting calendar anomalies in Asian stock markets using a stochastic dominance approach, *Journal of Multinational Financial Management* 17(2) (2007) 125–141.
- [21] H. Levy, M. Sarnat, Alternative efficiency criteria: an empirical analysis, *Journal of Finance* 25 (1970) 1153-1158.
- [22] H. Levy, M. Sarnat, Investment performance in an imperfect securities market, *Financial Analyst Journal* 28 (1972) 78-81.
- [23] C.K. Li, W.K. Wong, Extension of stochastic dominance theory to random variables, *RAIRO Recherche Opérationnelle* 33(4) (1999) 509-524.
- [24] O. Linton, E. Maasoumi, Y.J. Whang, Consistent testing for stochastic dominance under general sampling schemes, *Review of Economic Studies* 72 (2005) 735-765.
- [25] D. McFadden, Testing for stochastic dominance, in: Fomby, T.B., Seo, T.K., (Eds.), *Studies in the Economics of Uncertainty*, Springer Verlag, New York (1989) 113-132.
- [26] J. Richmond, A general method for constructing simultaneous confidence intervals, *Journal of the American Statistical Association* 77 (1982) 455-460.
- [27] H.N. Seyhun, Can omitted risk factors explain the January effects? A stochastic dominance approach, *Journal of Financial and Quantitative Analysis* 28(2) (1993) 195-212.
- [28] W. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19 (1964) 425-442.
- [29] S-H. Shin, Essays on asset return predictability (stock market, foreign exchange rates), Unpublished Ph.D. Thesis, Massachusetts Institute of Technology (1994).

- [30] Z., Sidak, Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* 62 (1967) 626-633.
- [31] M.R. Stoline, H.K. Ury, Tables of the studentized maximum modulus distribution and an application to multiple comparisons among means, *Technometrics* 21 (1979) 87-93.
- [32] L. Tesfatsion, Stochastic dominance and maximization of expected utility, *Review of Economic Studies* 43 (1976) 301-315.
- [33] Y.K. Tse, X. Zhang, A Monte Carlo investigation of some tests for stochastic dominance, *Journal of Statistical Computation and Simulation* 74 (2004) 361-378.
- [34] H. Varian, Nonparametric tests of models of investor behavior, *Journal of Financial and Quantitative Analysis* 18 (1983) 269-278.
- [35] J. Wingender, J.E. Groff, On stochastic dominance analysis of day-of-the-week return patterns, *Journal of Financial Research* 12(1) (1989) 51-55.
- [36] W.K. Wong, Stochastic dominance and mean-variance measures of profit and loss for business planning and investment, *European Journal of Operational Research* 182 (2007) 829-843.
- [37] W.K. Wong, R. Chan, Markowitz and prospect stochastic dominances, *Annals of Finance* (2007) forthcoming.
- [38] W.K. Wong, C.K. Li, A note on convex stochastic dominance theory, *Economics Letters* 62 (1999) 293-300.
- [39] W.K. Wong, C. Ma, Preferences over location-scale family, *Economic Theory* (2007) forthcoming.
- [40] W.K. Wong, K.F. Phoon, H.H. Lean, Performance of Asian hedge funds: stochastic dominance and mean-variance approaches, *Pacific-Basin Finance Journal* (2007) forthcoming.
- [41] B. Zheng, B.J. Cushing, Statistical inference for testing inequality indices with dependent samples, *Journal of Econometrics* 101 (2001) 315-335.
- [42] B. Zheng, J.P. Formby, W.J. Smith, V.K. Chow, Inequality orderings, normalized stochastic dominance, and statistical inference, *Journal of Business and Economic Statistics* 18(4) (2000) 479-488.

**Table 1:** Empirical Sizes of the DD, Anderson, and KRS Tests for Homoskedastic Process

Parameters of Distributions : $\alpha_i = 0.0, \alpha_j = 0.0, \beta_i = 1.0, \beta_j = 1.0$										
		$s = 1$			$s = 2$			$s = 3$		
$N$	$k$	DD	A	KRS	DD	A	KRS	DD	A	KRS
50	6	0.0260	0.0002	0.0000	0.0155	0.0000	0.0000	0.0078	0.0000	0.0000
	10	0.0111	0.0021		0.0075	0.0000		0.0034	0.0000	
	15	0.0036	0.0023		0.0039	0.0000		0.0014	0.0000	
100	6	0.0778	0.0001	0.0000	0.0138	0.0000	0.0007	0.0071	0.0000	0.0012
	10	0.0221	0.0000		0.0059	0.0000		0.0026	0.0000	
	15	0.0129	0.0000		0.0024	0.0000		0.0010	0.0000	
500	6	0.0883	0.0001	0.0000	0.0177	0.0000	0.0004	0.0105	0.0000	0.0048
	10	0.0425	0.0000		0.0072	0.0000		0.0031	0.0000	
	15	0.0216	0.0000		0.0020	0.0000		0.0010	0.0000	
1000	6	0.1021	0.0003	0.0000	0.0206	0.0000	0.0017	0.0099	0.0000	0.0059
	10	0.0474	0.0000		0.0061	0.0000		0.0049	0.0000	
	15	0.0255	0.0000		0.0037	0.0000		0.0009	0.0000	
Average		0.0401	0.0004	0.0000	0.0089	0.0000	0.0007	0.0045	0.0000	0.0030

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , in (8) generated by setting  $\alpha_i = \alpha_j = 0.0$  and  $\beta_i = \beta_j = 1.0$  with their innovations,  $e_{it}$  and  $e_{jt}$ , following the homoskedastic process shown in (9a), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures are the empirical sizes of the Davidson and Duclos (DD), Anderson (A) and Kaur, Rao and Singh (KRS) tests. The empirical size is estimated by calculating the percentage of rejections of the null hypothesis,  $H_0$ , when the null hypothesis,  $H_0$ , is true.

$H_0$  is defined in (3),  $N$  is the sample size,  $k$  is the number of major grid points, and  $s$  is the order of SD.

**Table 2a:** Empirical Powers of the DD and Anderson Tests for Homoskedastic Process

Parameters of Distributions	$N$	$k$	$s = 1$		$s = 2$		$s = 3$	
			DD	A	DD	A	DD	A
$\alpha_i = 0.2$ $\alpha_j = -0.2$ $\beta_i = 1.0$ $\beta_j = 1.0$	50	6	0.6931	0.0246	0.9536	0.0007	0.8776	0.0003
		10	0.5171	0.0047	0.9094	0.0000	0.8047	0.0000
		15	0.3518	0.0040	0.8643	0.0000	0.7372	0.0000
	100	6	0.9887	0.1259	0.9999	0.0103	0.9984	0.0059
		10	0.9455	0.0374	0.9990	0.0013	0.9956	0.0003
		15	0.8962	0.0138	0.9977	0.0002	0.9931	0.0002
	500	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		10	1.0000	0.9990	1.0000	0.9998	1.0000	0.9993
		15	1.0000	0.9910	1.0000	0.9983	1.0000	0.9937
1000	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
Average			0.8660	0.5167	0.9770	0.5009	0.9506	0.5000

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , (see, (8)) generated by setting  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$ , and  $\beta_i = \beta_j = 1.0$  and by adopting homoskedastic innovations,  $e_{it}$  and  $e_{jt}$ , (see (9a)), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures are the empirical powers of the Davidson and Duclos (DD) and Anderson (A) tests. The empirical power is estimated by the probability of rejecting the null hypothesis,  $H_0$ , when the null hypothesis,  $H_0$ , of no dominance is false or, equivalently, estimated by the probability of failing to reject the alternative hypothesis,  $H_{A1}$ , when the alternative hypothesis,  $H_{A1}$ , is true. Since there is SD between  $X$  and  $Y$  for  $s = 1, 2, 3$ ,  $H_{A1}$  should not be rejected for  $s = 1, 2, 3$ .

$H_0$  and  $H_{A1}$  are defined in (3),  $N$  is the sample size,  $k$  is the number of major grid points, and  $s$  is the order of SD.

**Table 2b:** Empirical Powers of the DD and Anderson Tests for Homoskedastic Process

Parameters of Distributions	$N$	$k$	$s = 1$		$s = 2$		$s = 3$	
			DD	A	DD	A	DD	A
$\alpha_i = 0.0$ $\alpha_j = 0.0$ $\beta_i = 0.0$ $\beta_j = 2.0$	50	6	0.9616	0.8982	0.6465	0.4220	0.4308	0.2279
		10	0.8931	0.7328	0.4471	0.2084	0.2345	0.0798
		15	0.8126	0.5424	0.3180	0.1043	0.1435	0.0269
	100	6	0.9999	0.9996	0.9854	0.9589	0.9475	0.8807
		10	0.9997	0.9983	0.9517	0.8830	0.8519	0.7098
		15	0.9986	0.9918	0.9128	0.7862	0.7469	0.5430
	500	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1000	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
Average			0.9721	0.9303	0.8551	0.7802	0.7796	0.7057

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , (see, (8)) generated by setting  $\alpha_i = \alpha_j = 0.0$ ,  $\beta_i = 0.0$  and  $\beta_j = 2.0$  and by adopting homoskedastic innovations,  $e_{it}$  and  $e_{jt}$ , (see (9a)) where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures are the empirical powers of the Davidson and Duclos (DD) and Anderson (A) tests. The empirical power is estimated by the probability of rejecting the null hypothesis,  $H_0$ , when the null hypothesis,  $H_0$ , of no dominance is false or, equivalently, estimated by the probability of failing to reject the alternative hypothesis when the alternative hypothesis is true. Since SD exists between  $X$  and  $Y$  for  $s = 2$  and  $3$  but there is no first-order SD between  $X$  and  $Y$ , we compute the empirical power for  $s=1$  to be the estimated probability of failing to reject the alternative hypothesis,  $H_A$ , for  $s=1$  and compute their empirical power to be the estimated probability of failing to reject the alternative hypothesis,  $H_{A1}$ , for  $s=2$  and  $s=3$ .

$H_0$ ,  $H_A$  and  $H_{A1}$  are defined in (3),  $N$  is the sample size,  $k$  is the number of major grid points, and  $s$  is the order of SD.

**Table 3:** Empirical Powers of the KRS Tests for Homoskedastic Process

	A	s = 1		s = 2		s = 3	
	N	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$
$\alpha_i = 0.2$ $\alpha_j = -0.2$ $\beta_i = \beta_j = 1.0$	50	0.0001	1.0000	0.1040	1.0000	0.0398	1.0000
	100	0.0048	1.0000	0.4003	1.0000	0.2147	1.0000
	500	0.7894	1.0000	0.9673	1.0000	0.8238	1.0000
	1000	0.9925	1.0000	0.9996	1.0000	0.9704	1.0000
	Average		0.4467	1.0000	0.6178	1.0000	0.5122
	B	s = 1		s = 2		s = 3	
	N	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$
$\alpha_i = \alpha_j = 0.0$ $\beta_i = 0.0$ $\beta_j = 2.0$	50	0.0000	1.0000	0.0800	1.0000	0.1303	1.0000
	100	0.0000	1.0000	0.1552	1.0000	0.6719	1.0000
	500	0.0000	1.0000	0.3437	1.0000	1.0000	1.0000
	1000	0.0000	1.0000	0.5103	1.0000	1.0000	1.0000
	Average		0.0000	1.0000	0.2723	1.0000	0.7006

Notes: The Monte Carlo simulation is conducted on dependent returns  $r_{it}$  and  $r_{jt}$  (see, (8)) generated by setting different values of  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$  and  $\beta_j$  and by adopting homoskedastic innovations,  $e_{it}$  and  $e_{jt}$ , (see (9a)) where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures in Panels A and B are the empirical relative frequencies that the hypothesis  $H_{A1}$  or  $H_{A2}^0$  is NOT rejected for  $s = 1, 2, 3$ .

The values in the columns under  $H_{A1}$  are the empirical powers of the KRS test. The empirical power is estimated by the probability of failing to reject the alternative hypothesis,  $H_{A1}$ , when the alternative hypothesis,  $H_{A1}$ , is true.

For Panel A in which  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$ , and  $\beta_i = \beta_j = 1.0$ , there exists SD between  $X$  and  $Y$  for  $s = 1, 2, 3$  and thus  $H_{A1}$  should not be rejected for  $s = 1, 2, 3$ . The empirical powers of the correspondence columns under  $H_{A1}$  confirm this argument for  $s = 1, 2, 3$ .

For Panel B in which  $\alpha_i = \alpha_j = 0.0$ ,  $\beta_i = 0.0$  and  $\beta_j = 2.0$ , SD exists between  $X$  and  $Y$  for  $s = 2$  and 3 but there is no first-order SD between  $X$  and  $Y$ . Thus,  $H_{A1}$  should not be rejected for  $s = 2$  and 3, but it should be rejected for  $s = 1$ . The empirical powers of the correspondence columns under  $H_{A1}$  for  $s = 2, 3$  confirm that  $r_{it}$  dominates  $r_{jt}$  at the second and third orders, whereas the empirical powers of the corresponding column under  $H_{A1}$  for  $s = 1$  confirm that  $r_{it}$  does not dominate  $r_{jt}$  at the first order.

$H_{A1}$  or  $H_{A2}^0$  are defined in (3) and (3a) respectively,  $N$  is the sample size, and  $s$  is the order of SD.

**Table 4:** Empirical Sizes of the DD, Anderson, and KRS Tests for Heteroskedastic Process

Parameters of Distributions : $\alpha_i = 0.0, \alpha_j = 0.0, \beta_i = 1.0, \beta_j = 1.0$										
$N$	$k$	$s = 1$			$s = 2$			$s = 3$		
		DD	A	KRS	DD	A	KRS	DD	A	KRS
50	6	0.0356	0.0002	0.0000	0.0150	0.0000	0.0000	0.0058	0.0000	0.0000
	10	0.0108	0.0008		0.0046	0.0000		0.0022	0.0000	
	15	0.0041	0.0030		0.0023	0.0000		0.0005	0.0000	
100	6	0.0845	0.0002	0.0000	0.0174	0.0000	0.0016	0.0062	0.0000	0.0009
	10	0.0270	0.0000		0.0057	0.0000		0.0020	0.0000	
	15	0.0142	0.0000		0.0028	0.0000		0.0009	0.0000	
500	6	0.1505	0.0004	0.0000	0.0354	0.0000	0.0022	0.0113	0.0000	0.0101
	10	0.0850	0.0001		0.0119	0.0000		0.0032	0.0000	
	15	0.0551	0.0001		0.0054	0.0000		0.0009	0.0000	
1000	6	0.2376	0.0004	0.0000	0.0683	0.0000	0.0017	0.0202	0.0000	0.0095
	10	0.1449	0.0000		0.0275	0.0000		0.0059	0.0000	
	15	0.0938	0.0000		0.0112	0.0000		0.0027	0.0000	
Average		0.0786	0.0004	0.0000	0.0173	0.0000	0.0014	0.0052	0.0000	0.0051

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , (see (8)) generated by setting  $\alpha_i = \alpha_j = 0.0$  and  $\beta_i = \beta_j = 1.0$  with their innovations,  $e_{it}$  and  $e_{jt}$ , following the heteroskedastic process as shown in (9b), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures are the empirical sizes of the Davidson and Duclos (DD), Anderson (A) and Kaur, Rao and Singh (KRS) tests. The empirical size is estimated by calculating the percentage of rejections of the null hypothesis,  $H_0$ , when the null hypothesis,  $H_0$ , is true.

$H_0$  is defined in (3),  $N$  is the sample size,  $k$  is the number of major grid points, and  $s$  is the order of SD.



**Table 5a:** Empirical Powers of the DD and Anderson Tests for Heteroskedastic Process

Parameters of Distributions	$N$	$k$	$s = 1$		$s = 2$		$s = 3$	
			DD	A	DD	A	DD	A
$\alpha_i = 0.2$ $\alpha_j = -0.2$ $\beta_i = 1.0$ $\beta_j = 1.0$	50	6	0.5588	0.0190	0.7714	0.0009	0.5858	0.0004
		10	0.3830	0.0054	0.6708	0.0002	0.4745	0.0000
		15	0.2717	0.0041	0.5873	0.0000	0.3924	0.0000
	100	6	0.9254	0.0820	0.9711	0.0116	0.9068	0.0052
		10	0.8280	0.0282	0.9441	0.0016	0.8523	0.0008
		15	0.7410	0.0084	0.9176	0.0003	0.7990	0.0000
	500	6	1.0000	0.9959	1.0000	0.9872	1.0000	0.9159
		10	1.0000	0.9749	1.0000	0.9173	1.0000	0.7663
		15	1.0000	0.9307	1.0000	0.7958	1.0000	0.5907
1000	6	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9986	
	15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9986	
Average			0.8090	0.5041	0.9052	0.4762	0.8342	0.4397

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$ , and  $\beta_i = \beta_j = 1.0$  and by adopting heteroskedastic innovations,  $e_{it}$  and  $e_{jt}$  (see (9b)), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures are the empirical powers of the Davidson and Duclos (DD) and Anderson (A) tests. The empirical power is estimated by the probability of rejecting the null hypothesis,  $H_0$ , when the null hypothesis,  $H_0$ , of no dominance is false or, equivalently, estimated by the probability of failing to reject the alternative hypothesis,  $H_{A1}$ , when the alternative hypothesis,  $H_{A1}$ , is true. Since there is SD between  $X$  and  $Y$  for  $s = 1, 2, 3$ ,  $H_{A1}$  should not be rejected for  $s = 1, 2, 3$ . The terms  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $H_0$  and  $H_{A1}$  are defined in (3);  $N$  is the sample size;  $k$  is the number of major grid points; and  $s$  is the order of SD.

**Table 5b:** Empirical Powers of the DD and Anderson Tests for Heteroskedastic Process

Parameters of Distributions	$N$	$k$	$s = 1$		$s = 2$		$s = 3$	
			DD	A	DD	A	DD	A
$\alpha_i = 0.0$ $\alpha_j = 0.0$ $\beta_i = 0.0$ $\beta_j = 2.0$	50	6	0.8803	0.6643	0.4940	0.2086	0.2872	0.0949
		10	0.7827	0.4723	0.3309	0.0952	0.1585	0.0307
		15	0.6822	0.3135	0.2349	0.0492	0.0886	0.0103
	100	6	0.9971	0.9727	0.8889	0.6428	0.7299	0.4389
		10	0.9874	0.9205	0.8018	0.4744	0.5880	0.2693
		15	0.9743	0.8466	0.7248	0.3560	0.4621	0.1692
	500	6	1.0000	1.0000	1.0000	1.0000	0.9997	0.9875
		10	1.0000	1.0000	1.0000	1.0000	0.9993	0.9818
		15	1.0000	1.0000	1.0000	1.0000	0.9990	0.9750
1000	6	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9982	
	15	1.0000	1.0000	1.0000	1.0000	0.9999	0.9985	
Average			0.9420	0.8492	0.7896	0.6522	0.6927	0.5795

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = \alpha_j = 0.0$ ,  $\beta_i = 0.0$  and  $\beta_j = 2.0$  and by adopting heteroskedastic innovations,  $e_{it}$  and  $e_{jt}$  (see (9b)), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

The figures are the empirical powers of the Davidson and Duclos (DD) and Anderson (A) tests. The empirical power is estimated by the probability of rejecting the null hypothesis,  $H_0$ , when the null hypothesis,  $H_0$ , of no dominance is false or, equivalently, estimated by the probability of failing to reject the alternative hypothesis when the alternative hypothesis is true. Since SD exists between  $X$  and  $Y$  for  $s = 2$  and  $3$  but there is no first-order SD between  $X$  and  $Y$ , we compute the empirical power to be the estimated probability of failing to reject the alternative hypothesis,  $H_A$ , for  $s=1$  and compute their empirical power to be the estimated probability of failing to reject the alternative hypothesis,  $H_{A1}$ , for  $s=2$  and  $s=3$ .

The terms  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $H_0$ ,  $H_A$  and  $H_{A1}$  are defined in (3);  $N$  is the sample size;  $k$  is the number of major grid points; and  $s$  is the order of SD.

**Table 6:** Empirical Power of the KRS Test for Heteroskedastic Process

	A	s = 1		s = 2		s = 3	
	N	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$
$\alpha_i = 0.2$ $\alpha_j = -0.2$ $\beta_i = \beta_j = 1.0$	50	0.0002	1.0000	0.0543	1.0000	0.0178	1.0000
	100	0.0012	1.0000	0.1512	1.0000	0.0711	1.0000
	500	0.2506	1.0000	0.4808	1.0000	0.1193	1.0000
	1000	0.7111	1.0000	0.7333	1.0000	0.1332	1.0000
	Average		0.2408	1.0000	0.3549	1.0000	0.0854
$\alpha_i = \alpha_j = 0.0$ $\beta_i = 0.0$ $\beta_j = 2.0$	B	s = 1		s = 2		s = 3	
	N	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$	$H_{A1}$	$H_{A2}^0$
	50	0.0000	1.0000	0.0455	1.0000	0.0704	1.0000
	100	0.0000	1.0000	0.0933	1.0000	0.3431	1.0000
	500	0.0000	1.0000	0.2223	1.0000	0.8013	1.0000
1000	0.0000	1.0000	0.3378	1.0000	0.9034	1.0000	
Average		0.0000	1.0000	0.1747	1.0000	0.5296	1.0000

Notes: The Monte Carlo simulation is conducted on dependent returns,  $r_{it}$  and  $r_{jt}$ , generated by setting different values of  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  and by adopting heteroskedastic innovations,  $e_{it}$  and  $e_{jt}$  (see (9b)), where  $e_{it}$  and  $e_{jt}$  are correlated as shown in (10).

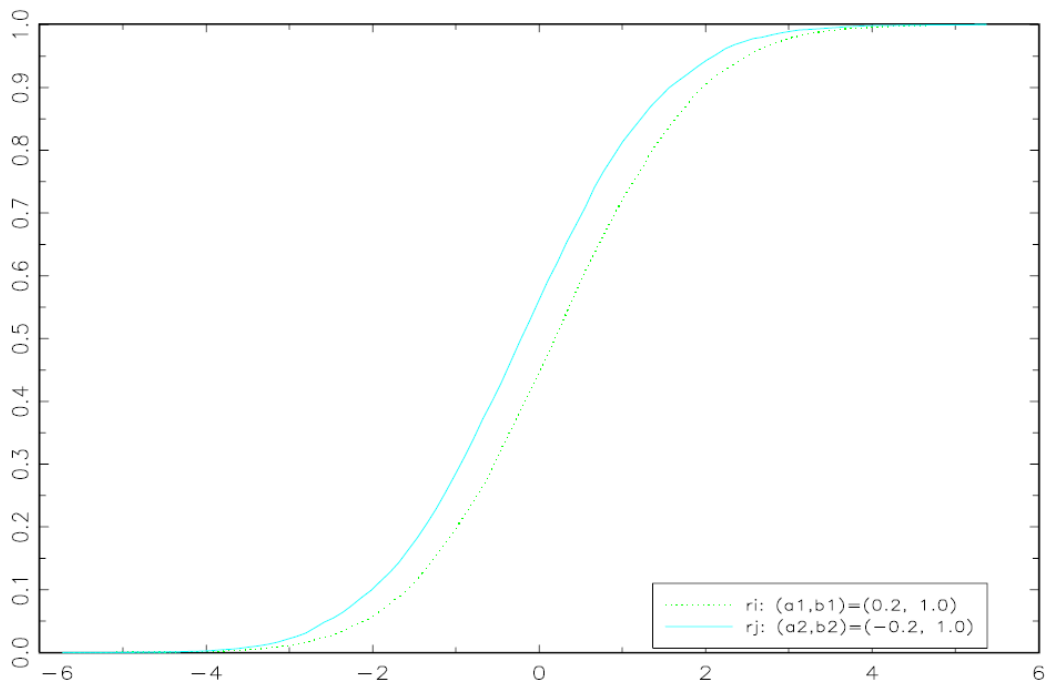
The figures in Panels A and B are the empirical relative frequencies that the hypothesis  $H_{A1}$  or  $H_{A2}^0$  is NOT rejected for  $s = 1, 2, 3$ .

For Panel A in which  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$  and  $\beta_i = \beta_j = 1.0$ , SD exists between  $X$  and  $Y$  for  $s = 1, 2, 3$  and thus  $H_{A1}$  should not be rejected for  $s = 1, 2, 3$ .

For Panel B in which  $\alpha_i = \alpha_j = 0.0$ ,  $\beta_i = 0.0$  and  $\beta_j = 2.0$ , SD exists between  $X$  and  $Y$  for  $s = 2$  and  $3$  but there is no first-order SD between  $X$  and  $Y$ . Thus,  $H_{A1}$  should not be rejected for  $s = 2$  and  $3$ , but it should be rejected for  $s = 1$ .

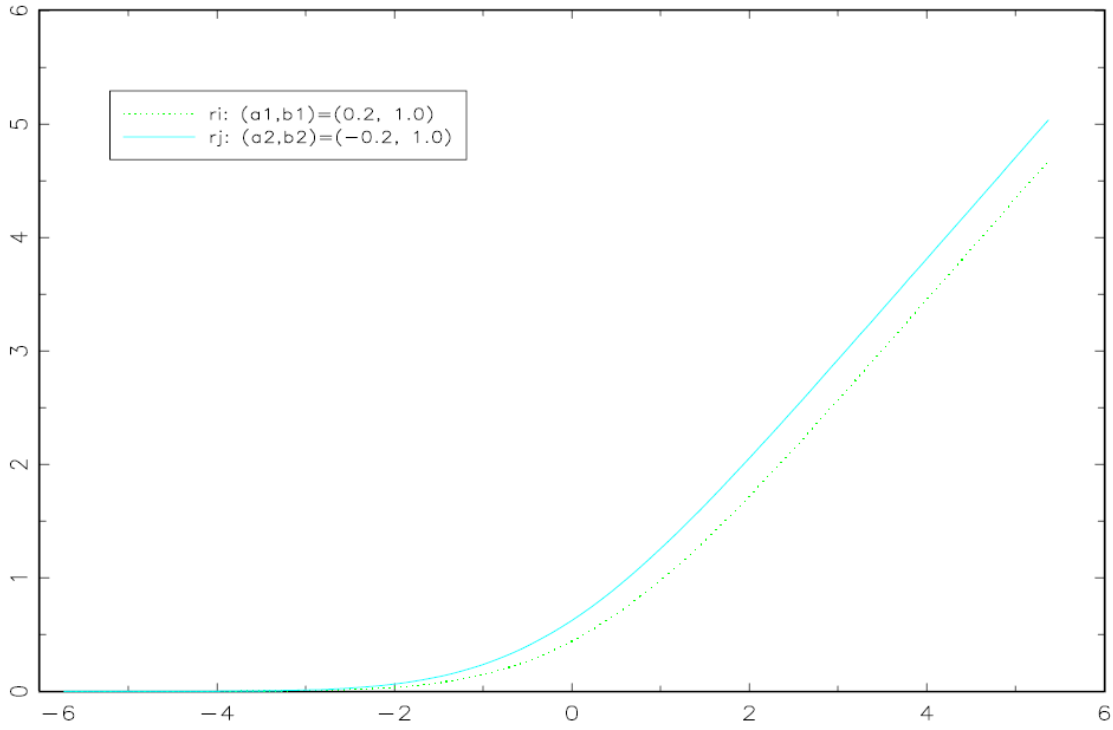
The terms  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $H_{A1}$  or  $H_{A2}^0$  are defined in (3) and (3a), respectively;  $N$  is the sample size; and  $s$  is the order of SD.

Figure 1.1: Plot of the First-Order Integrals for Different Alpha



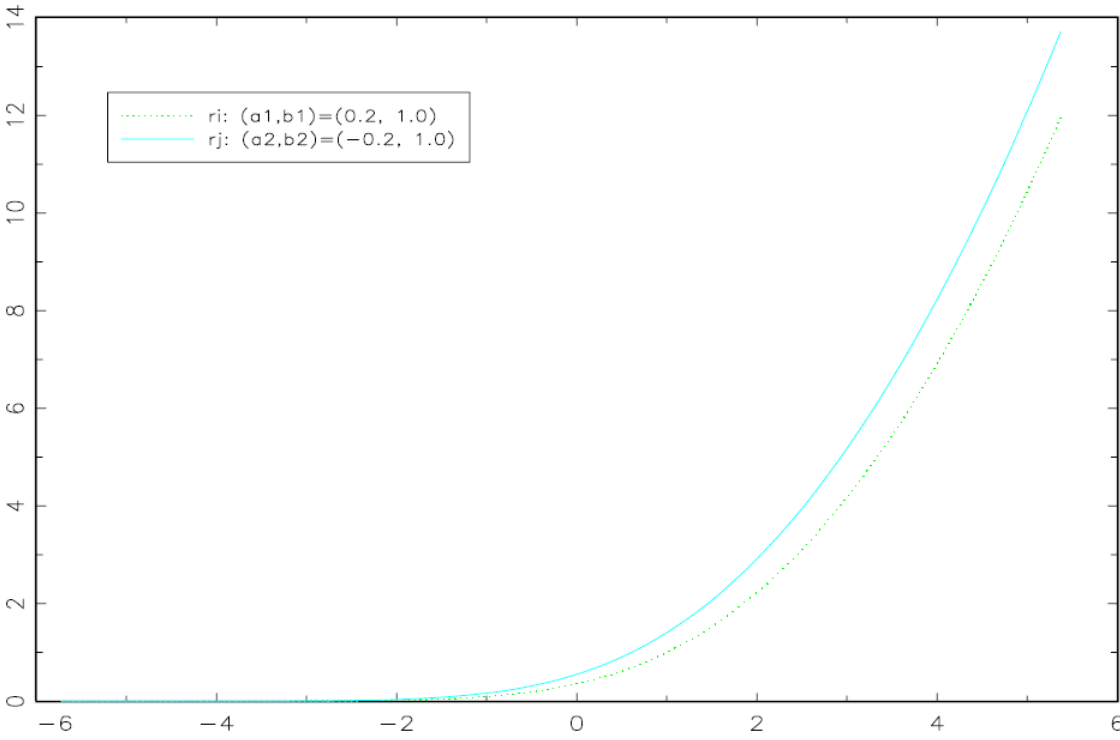
Note: The plots are the first-order integrals,  $F_1$  and  $G_1$  (see (1)) for the returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$ , and  $\beta_i = \beta_j = 1$ . Refer to (8) for the definitions of  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$ .

Figure 1.2: Plot of the Second-Order Integrals for Different Alpha



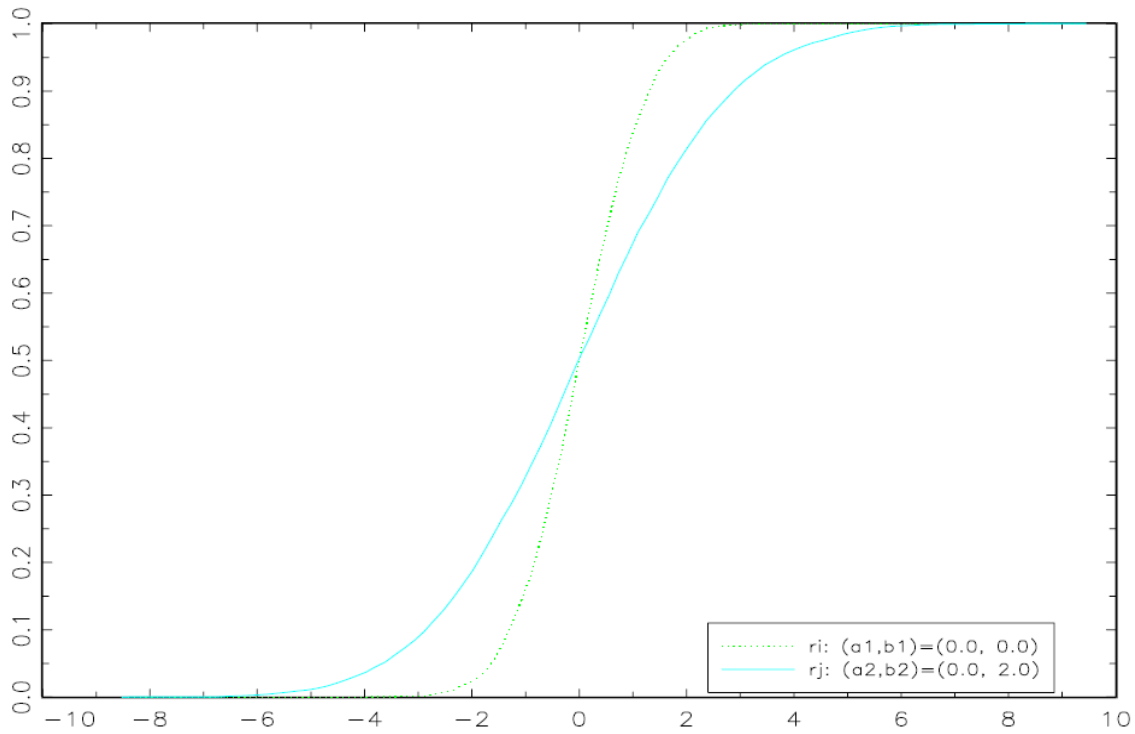
Note: The plots are the second-order integrals,  $F_2$  and  $G_2$  (see (1)) for the returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$ , and  $\beta_i = \beta_j = 1$ . Refer to (8) for the definitions of  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$ .

Figures 1.3: Plot of the Third-Order Integrals for Different Alpha



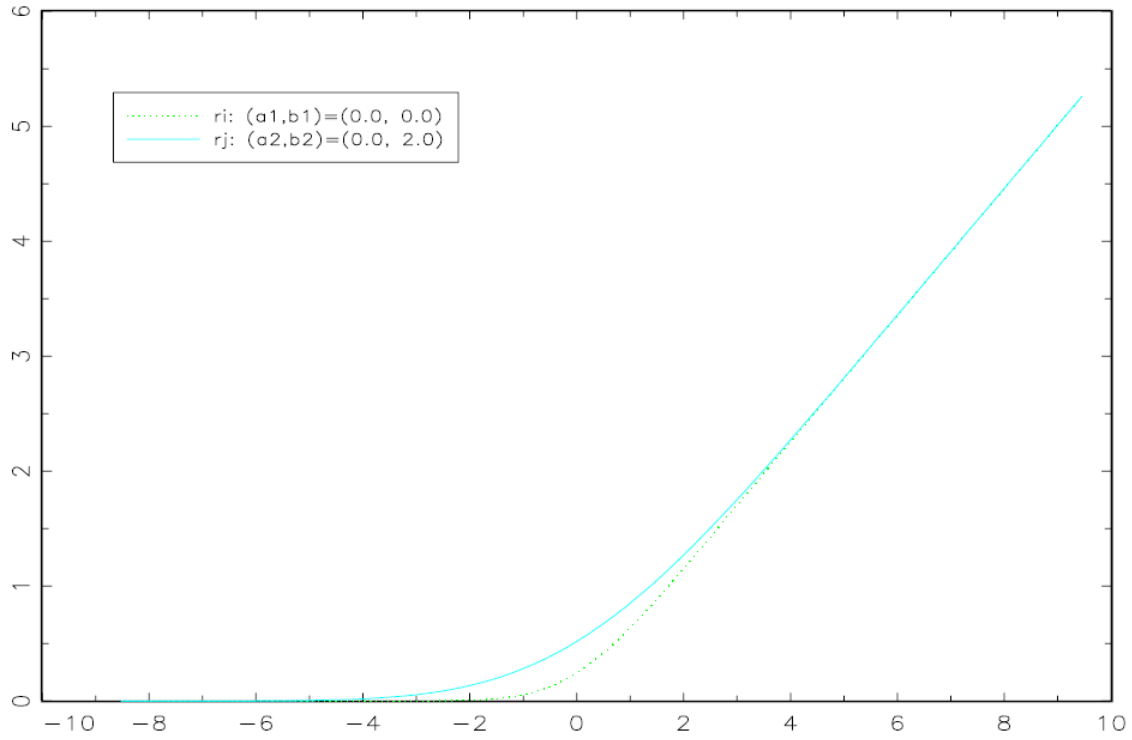
Note: The plots are the third-order integrals,  $F_3$  and  $G_3$  (see (1)) for the returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = 0.2$ ,  $\alpha_j = -0.2$ , and  $\beta_i = \beta_j = 1$ . Refer to (8) for the definitions of  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$ .

Figure 2.1: Plot of the First-Order Integrals for Different Beta



Note: The plots are the first-order integrals,  $F_1$  and  $G_1$  (see (1)) for the returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = \alpha_j = 0$ ,  $\beta_i = 0$ , and  $\beta_j = 2$ . Refer to (8) for the definitions of  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$ .

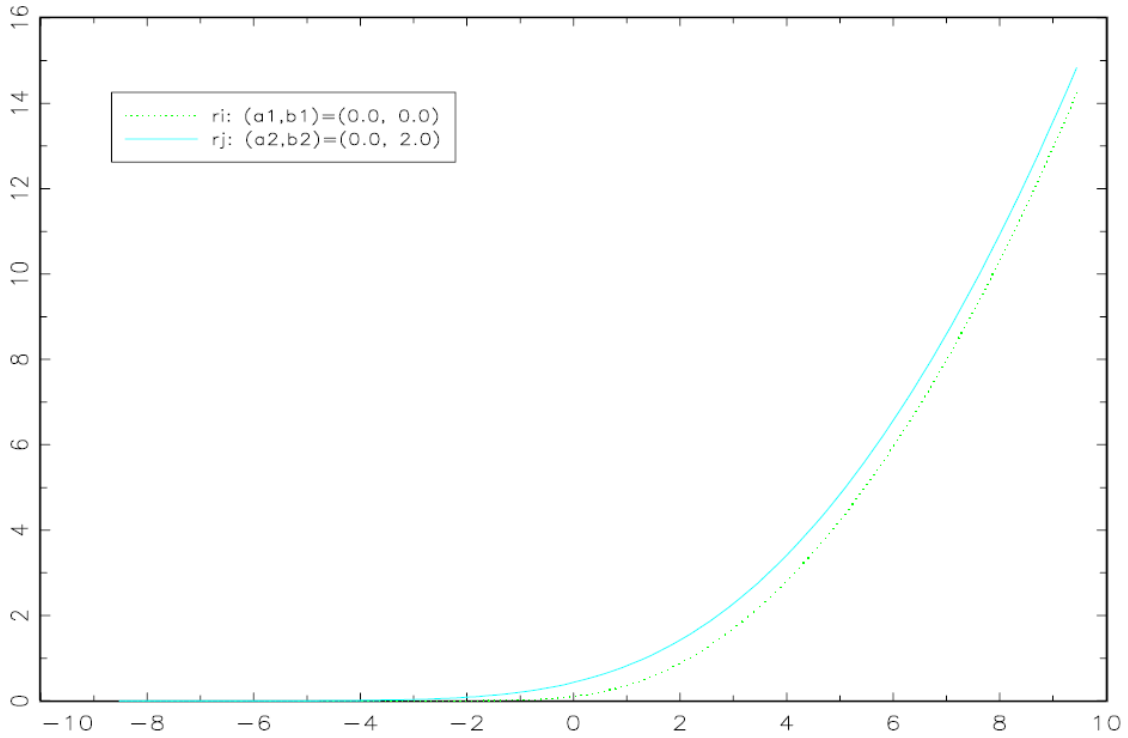
Figure 2.2: Plot of the Second-Order Integrals for Different Beta



Note: The plots are the second-order integrals,  $F_2$  and  $G_2$  (see (1)) for the returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = \alpha_j = 0$ ,  $\beta_i = 0$ , and  $\beta_j = 2$ . Refer to (8) for the definitions of  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$ .

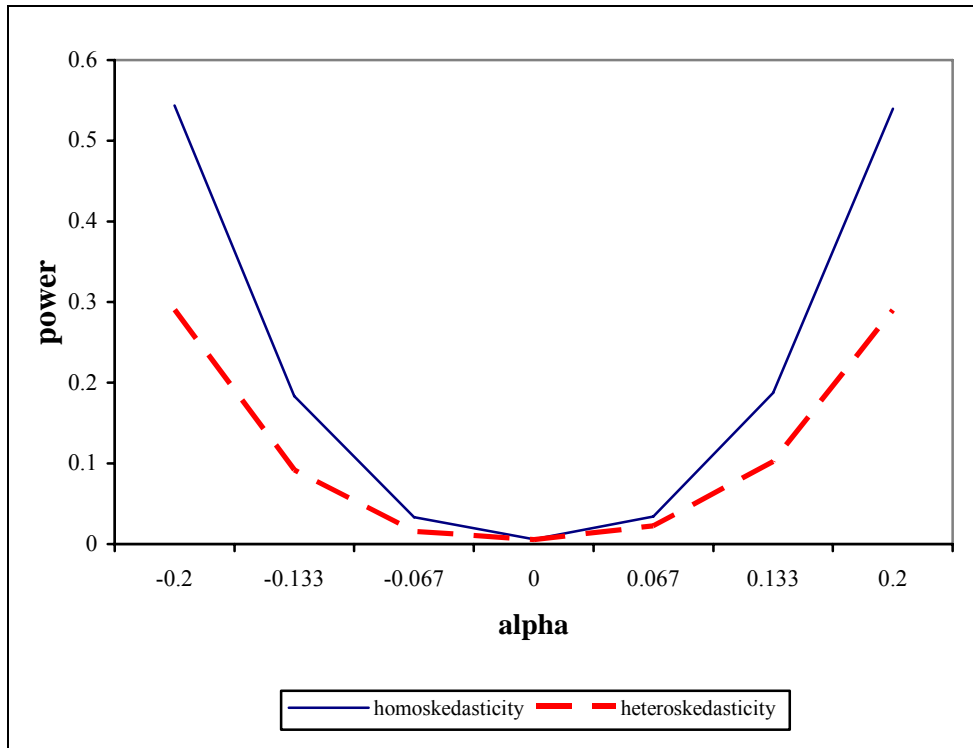


Figures 2.3: Plot of the Third-Order Integrals for Different Beta



Note: The plots are the third-order integrals,  $F_3$  and  $G_3$  (see (1)) for the returns,  $r_{it}$  and  $r_{jt}$ , generated by setting  $\alpha_i = \alpha_j = 0$ ,  $\beta_i = 0$ , and  $\beta_j = 2$ . Refer to (8) for the definitions of  $r_{it}$ ,  $r_{jt}$ ,  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$ .

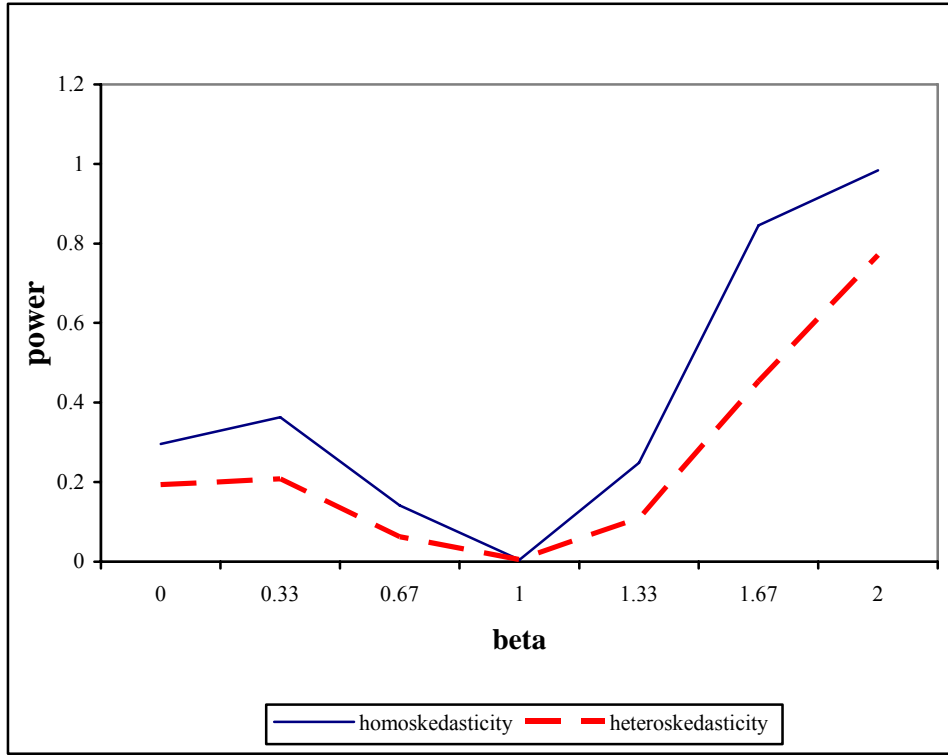
Figure 3: Power Graph for the DD Test for Different Values of Alpha under Homoskedastic and Heteroskedastic Processes



Note: This figure shows the power of the Davidson and Duclos (DD) test for the returns  $r$  generated by different values of alpha ( $\alpha$ ) under both the homoskedastic and heteroskedastic processes.<sup>32</sup> Refer to (8) for the definitions of  $r$  and  $\alpha$  and refer to (9) and (10) for the definitions of the homoskedastic and heteroskedastic processes.

<sup>32</sup> We report only the plot of powers for different values of  $\alpha_j$  when  $n = 100$ ,  $k = 10$ ,  $s = 2$ ,  $\alpha_i = 0.0$  and  $\beta_i = \beta_j = 1.0$ . We skip reporting the plots of other values of  $n$ ,  $k$ ,  $s$ ,  $\alpha_i$ ,  $\beta_i$  and  $\beta_j$ , since all the plots will lead us to draw the same conclusion. The terms  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $N$  is the sample size;  $k$  is the number of major grid points; and  $s$  is the order of SD.

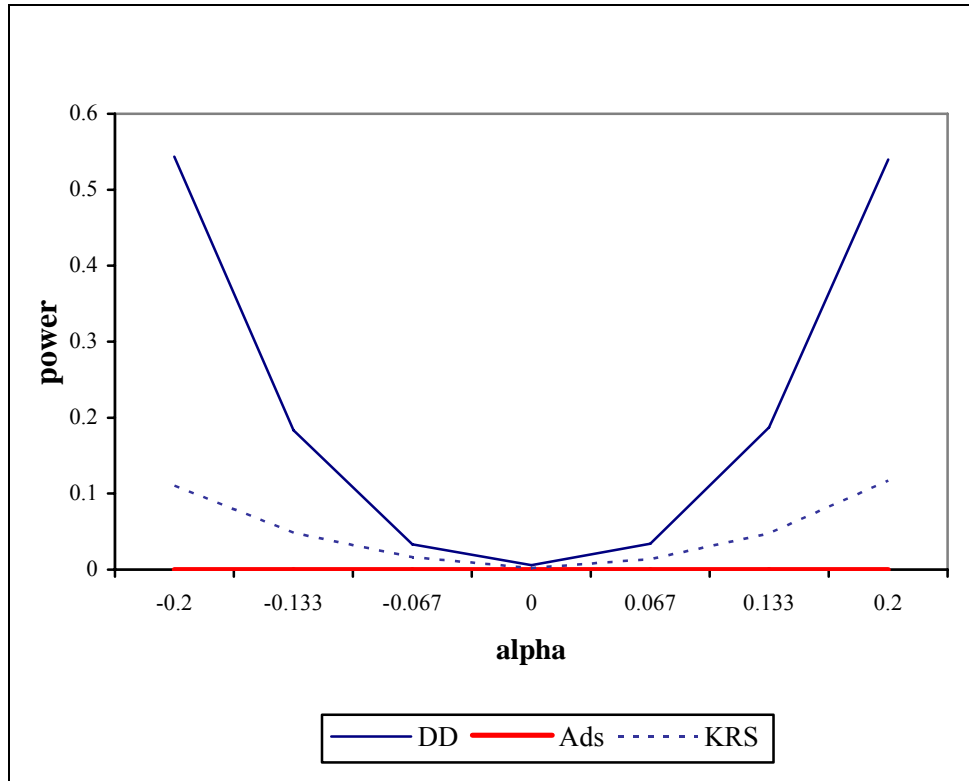
Figure 4: Power Graph for the DD Test for Different Values of Beta under Homoskedastic and Heteroskedastic Processes



Note: This figure shows the power of the Davidson and Duclos (DD) test for the returns  $r$  generated by different values of beta ( $\beta$ ) under both the homoskedastic and heteroskedastic processes.<sup>33</sup> Refer to (8) for the definitions of  $r$  and  $\beta$  and refer to (9) and (10) for the definitions of the homoskedastic and heteroskedastic processes.

<sup>33</sup> We report only the plot of powers for different values of  $\beta_j$  when  $n = 100$ ,  $k = 10$ ,  $s = 2$ ,  $\alpha_i = \alpha_j = 0.0$  and  $\beta_i = 1.0$ . We skip reporting the plots of other values of  $n$ ,  $k$ ,  $s$ ,  $\alpha_i$ ,  $\alpha_j$ , and  $\beta_i$ , since all the plots will lead us to the same conclusion. The terms  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $N$  is the sample size;  $k$  is the number of major grid points; and  $s$  is the order of SD.

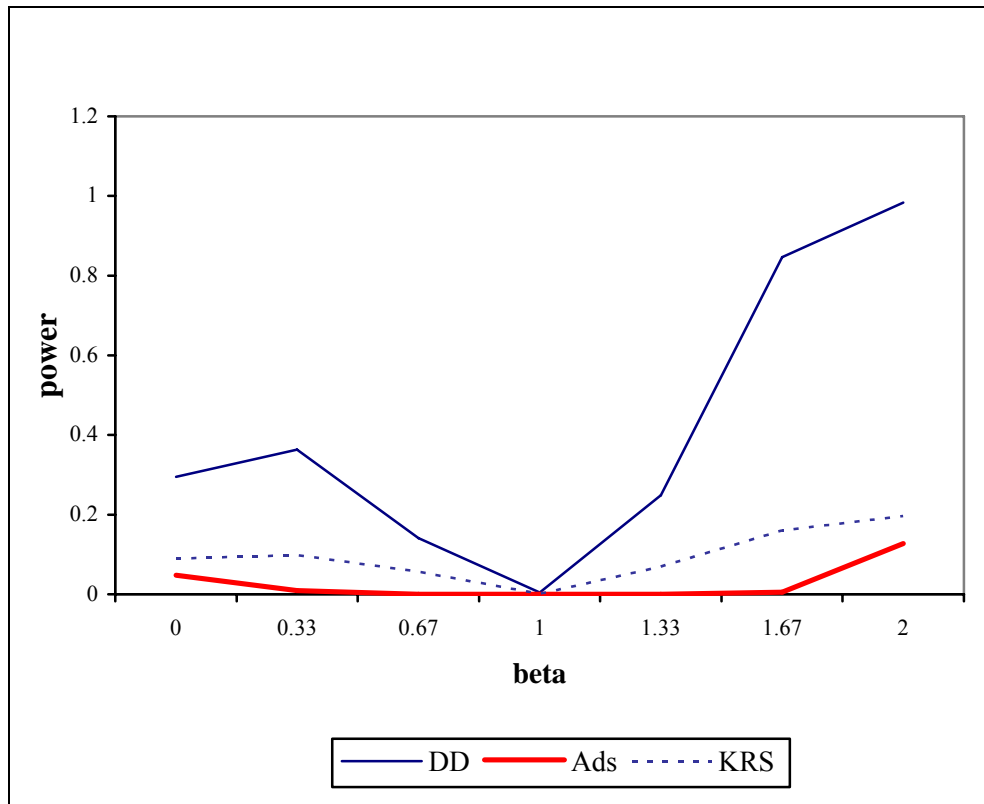
Figure 5: Power Graph for the DD, Anderson, KRS Tests with Different Values of  $\alpha$  under Heteroskedastic Process



Note: This figure shows the powers of the Davidson and Duclos (DD) test, Anderson (Ads) test and Kaur, Rao and Singh (KRS) test for the returns  $r$  generated by different values of  $\alpha$  under heteroskedastic process.<sup>34</sup> Refer to (8) for the definitions of  $r$  and  $\alpha$  and refer to (9) and (10) for the definitions of the homoskedastic and heteroskedastic processes.

<sup>34</sup> We report only the plot of powers for different values of  $\alpha_j$  when  $n = 100$ ,  $k = 10$ ,  $s = 2$ ,  $\alpha_i = 0.0$  and  $\beta_i = \beta_j = 1.0$ . We skip reporting the plots of other values of  $n$ ,  $k$ ,  $s$ ,  $\alpha_i$ ,  $\beta_i$ , and  $\beta_j$ , since all the plots will lead us to the same conclusion. The terms  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $N$  is the sample size;  $k$  is the number of major grid points; and  $s$  is the order of SD.

Figure 6: Power Graph for the DD, Anderson, KRS Tests with Different Values of  $\beta$  under Heteroskedastic Process



Note: This figure shows the powers of the Davidson and Duclos (DD) test, Anderson (Ads) test and Kaur, Rao and Singh (KRS) test for the returns  $r$  generated by different values of beta ( $\beta$ ) under heteroskedastic process.<sup>35</sup> Refer to (8) for the definitions of  $r$  and  $\beta$  and refer to (9) and (10) for the definitions of the homoskedastic and heteroskedastic processes.

<sup>35</sup> We report only the plot of powers for different values of  $\beta_j$  when  $n = 100$ ,  $k = 10$ ,  $s = 2$ ,  $\alpha_i = \alpha_j = 0.0$  and  $\beta_i = 1.0$ . We skip reporting the plots of other values of  $n$ ,  $k$ ,  $s$ ,  $\alpha_i$ ,  $\alpha_j$ , and  $\beta_i$ , since all the plots will lead us to the same conclusion. The terms  $\alpha_i$ ,  $\alpha_j$ ,  $\beta_i$ , and  $\beta_j$  are defined in (8);  $N$  is the sample size;  $k$  is the number of major grid points; and  $s$  is the order of SD.