Testing for stochastic dominance Using circular block methods

Hendrik Kläver
University of Cologne
Graduate School of Risk Management
Albertus-Magnus-Platz
D-50923 Köln
GERMANY
+49-(0)221-470-7704
klaever@wiso.uni-koeln.de

September 6, 2005

Abstract

This paper investigates the performance of various tests on stochastic dominance for dependent data. We consider the tests of Schmid and Trede (1997), of Xu, Fisher and Willson (1997) and of Linton, Maasoumi and Whang (2003). The dependence structures explored are the contemporaneous correlation between samples and conditional heteroskedasticity within samples. Simulations show that with the proposed bootstrap methods the tests perform rather poorly for small samples. We develop new circular bootstrap methods which make the tests of Schmid and Trede and of Linton, Maasoumi and Whang robust to the mentioned dependence structures. We determine the block lengths which make the tests keep the size and explore the power of the tests. The tests are applied to the daily returns of some German stocks.

Keywords: Stochastic dominance, test, conditional heteroskedasticity, block bootstrap. JEL Code C12, C15, C44, G10.

1 Introduction

In this paper, we investigate various tests on stochastic dominance when conditional heteroskedasticity is prevalent in the data. As financial data often feature this property, we have to consider it for the application of stochastic dominance tests. Various tests developed in the last years asymptotically capture the dependence structure very well, but we still do not know how these tests perform for finite samples. This paper analyzes this question and proposes some new tests which are asymptotically equivalent and perform better for finite samples.

Stochastic dominance is an important concept of decision theory. In particular, it has been applied in various branches of economics. In those fields, stochastic dominance of first and of second order are of interest. A real-valued random variable X (weakly) dominates a random variable Y

- in the sense of *First Order Stochastic Dominance* (SD1) if $E(u(X)) \geq E(u(Y))$ holds for every *nondecreasing* utility function u where E denotes the expected value
- in the sense of *Second Order Stochastic Dominance* (SD2) if $E(u(X)) \geq E(u(Y))$ holds for every *nondecreasing* and *concave* utility function u.

Stochastic dominance of X over Y in the sense of SDi ($i = 1, 2$) is denoted by $X \succ_i Y$. An agent who has a nondecreasing and concave utility function is called risk averse. If X dominates Y in the sense of SD1, X obviously also dominates in the sense of SD2.

If the distributions of X and Y are known, the analytic investigation of stochastic dominance is straightforward. Let F_X, F_Y be the distribution functions and Q_X, Q_Y the quantile functions of X and Y. $X \succ_1 Y$ is then equivalent to $F_X(x) \leq F_Y(x)$ for all $x \in \mathbb{R}$ and to $Q_X(x) \geq Q_Y(x)$ for all $x \in]0, 1[$. $X \succ_2 Y$ is equivalent to $\int_{-\infty}^{x} F_X(t)dt \leq \int_{-\infty}^{x} F_Y(t)dt$ for all $x \in \mathbb{R}$ and to $\int_0^x Q_X(t)dt \geq \int_0^x Q_Y(t)dt$ for all $x \in]0, 1[$. For a survey concerning stochastic dominance see e.g. Whitmore/ Findlay (1978) and Levy (1992).

In applications the distribution functions are usually unknown and have to be inferred from the observations of X and Y. In a descriptive approach one could compare the empirical distribution or quantile functions. The drawback of proceeding in this way is that the standard error interferes with the results. In particular SD1 is almost always rejected for larger samples. For investigations in terms of this topic see e.g. Nelson/ Pope (1991) and Stein/ Pfaffenberger/ Kumar (1983). Hence we need statistical tests for surveying a stochastic dominance relationship.

In the last two decades, many tests for stochastic dominance have been developed. Many of them have restrictive assumptions, in particular concerning the independence of the data. However, in general economic data do not satisfy these constraints. In particular, this is the case for financial data such as daily returns on assets or currencies, which show conditional heteroskedasticity. In other words, financial time series feature intertemporal dependence. In addition to this, for every time index t we have some contemporaneous dependence: usually X_t and Y_t are positively correlated.

This paper investigates the performance of various tests on stochastic dominance when conditional heteroskedasticity and contemporaneous correlation are prevalent in the data. We consider the tests of Schmid/ Trede (1997), Xu/ Fisher/ Willson (1997) and Linton/ Maasoumi/ Whang (2003). Hereafter, we denote the tests by ST, XFW and LMW. We investigate these tests because they asymptotically capture a dependence structure which is suitable for financial data. The ST test investigates whether X dominates Y in the sense of SD2. The XFW and LMW tests address SD1 as well as SD2. Schmid/ Trede take X_t and Y_t as matched pairs for each t; their test is based on permutation. Hence they consider the correlation of X_t and Y_t, however, no intertemporal dependence. The XFW and LMW tests use block methods for capturing the dependence structure within each time series and the correlation between them: Xu/ Fisher/ Willson use the moving block bootstrap, Linton/ Maasoumi/ Whang use a subsampling approach. They both propose that the tests perform well asymptotically if the data are generated by strongly mixing processes. In particular, GARCH processes are strongly mixing.

Simulations show that all of these tests do not perform very well for finite samples when the data are generated by a GARCH(1,1) process whose parameters are close to one in summation. A remedy is found in other blocking methods: the circular block bootstrap, its subsampling equivalent and the block permutation. From Lahiri (1999) we know that the circular block bootstrap performs as well as the moving block bootstrap asymptotically. We show analytically that the asymptotic result of Linton/ Maasoumi/ Whang for usual subsampling also holds for circular subsampling. Further simulations indicate that, for a finite sample, circular subsampling performs better than the usual subsampling of Linton/ Maasoumi/ Whang and block permutation performs better than the permutation test of Schmid/ Trede, whereas circular block bootstrap does not improve the performance of the test developed by Xu/ Fisher/ Willson. The choice of the block length is crucial for the modified versions of the tests of Schmid/ Trede and of Linton/ Maasoumi/ Whang. For various values of sample size n, in each case with optimal block length, we explore the power of the tests. The main drawback in the investigation is the complexity of the alternative. We apply the modified tests of Linton/ Maasoumi/ Whang on SD1 and of Schmid/ Trede on SD2 to the daily returns of stocks of the German stock index DAX. In the empirical investigation we consider a 1-year and a 10-year period.

In this paper, we proceed as follows. Section 2 presents the tests of Schmid/ Trede, Xu/ Fisher/ Willson and Linton/ Maasoumi/ Whang which use various resampling methods. In section 3 we establish the performance of these tests using a simulation study. In section 4 we develop some modified tests based on circular block methods. The simulation results for these tests are presented in section 5. We examine the power of the tests in section 6. The tests are applied on empirical data in section 7. Section 8 concludes.

2 Tests Based on Resampling Methods

We begin this section by introducing some notation. For a real random variable X let F_X be the distribution function and Q_X the quantile function. Further let $F_X^{(1)} = F_X$ and $F_X^{(k+1)}(x) = \int_{-\infty}^{x} F_X^{(k)}(t)dt$ and $Q_X^{(k+1)}(q) = \int_0^q Q_X^{(k)}(t)dt$ for all $x \in \mathbb{R}$ and $k \in \mathbb{N}$. Let U_1 denote the set of all nondecreasing utility functions and U_2 the set of all nondecreasing and concave utility functions. For random variables X and Y the following statements are equivalent ($i \in \{1, 2\}$):

1. $X \succ_i Y$, i.e. $E(u(X)) \geq E(u(Y))$ for all $u \in U_i$,

2. $F_X^{(i)}(x) \leq F_Y^{(i)}(x)$ for all $x \in \mathbb{R}$,

3. $Q_X^{(i)}(x) \geq Q_Y^{(i)}(x)$ for all $x \in \mathbb{R}$.

Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ denote the observations of X and Y and $x_{(i)}$, $y_{(i)}$ the i-th order statistics of the samples. Further let $\hat{F}_{X,n}(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{x \geq x_i\}}$ be the empirical distribution function and let $\hat{Q}_{X,n}(q) = x_{(\lceil nq \rceil)}$ be the empirical quantile function, where $\lceil z \rceil$ denotes the smallest integer equal or larger than z. Define $\hat{F}_{X,n}^{(2)}(x) = \int_{-\infty}^{x} \hat{F}_{X,n}(t)dt = \frac{1}{n}\sum_{i=1}^{n}(x - x_i)1_{\{x \geq x_i\}}$ and $\hat{Q}_{X,n}^{(2)}(q) = \int_0^q \hat{Q}_{X,n}(t)dt = \frac{1}{n}\sum_{i=1}^{\lceil nq \rceil} x_{(i)} - (\frac{\lceil nq \rceil}{n} - q)x_{(\lceil nq \rceil)}$. The notation regarding Y is analogous.

Statistical test procedures concerning stochastic dominance are based on statement 2 or 3 and use their empirical equivalents $\hat{F}_n^{(i)}$ and $\hat{Q}_n^{(i)}$. This is, for instance, the case for the tests of Schmid/ Trede, Xu/ Fisher/ Willson and Linton/ Maasoumi/ Whang which we will illustrate in this section.

2.1 A Permutation Test from Matched Pairs

Schmid/ Trede test the null hypothesis $H_0 : (X \succ_2 Y)$ against the alternative $H_1 :$ (not H_0) and $H_0^* : (X \succ_2 Y$ or $Y \succ_2 X)$ vs. $H_1^* :$ (not H_0^*). We confine our investigation to the first testing problem, which can also be written as

$$\begin{aligned} H_0 \quad &: \quad \text{For all } x \in \mathbb{R} : \; F_X^{(2)}(x) \leq F_Y^{(2)}(x) \\ \text{vs. } H_1 \quad &: \quad \text{There exists } x' \in \mathbb{R} : \; F_X^{(2)}(x') > F_Y^{(2)}(x'). \end{aligned}$$

Schmid/ Trede use the test statistic

$$T = \sup_{t \in \mathbb{R}} (\hat{F}_{X,n}^{(2)}(t) - \hat{F}_{Y,n}^{(2)}(t)) = \max_i (\hat{F}_{X,n}^{(2)}(z_{(i)}) - \hat{F}_{Y,n}^{(2)}(z_{(i)}))$$

where $z_{(i)}$ denotes the i-th order statistic of the combined sample $(z_1, \ldots, z_{2n}) = (x_1, \ldots, x_n, y_1, \ldots, y_n)$.

H_0 is rejected if $T \geq c$ where the critical value c is determined by permutations. There are 2^n possibilities of permuting x_i and y_i in the paired sample $(x_1, y_1), \ldots, (x_n, y_n)$. The corresponding values of the test statistics can be ordered according to size: $T^{(1)} \leq \cdots \leq T^{(2^n)}$. The critical value c is determined by $c = T^{((1-\alpha)2^n)}$. Under $F_X = F_Y$ the probability of wrongly rejecting H_0 is approximately α. As the number of permutations becomes large very quickly with increasing n, Schmid/ Trede take only M permutations at random and determine the critical value by $c = T^{((1-\alpha)M)}$. They show in a Monte Carlo study that under the assumption of a bivariate normal distribution with intertemporal independence $M = 500$ permutations are sufficient. Schmid/ Trede do not give any advice on how to decide if there is a tie, i.e. $T = T^{(k)}$ for some $k < (1-\alpha)M$ and some $k \geq (1-\alpha)M$. [1]

In section 3 we will investigate the performance of the test by means of simulation for the case that conditional heteroskedasticity is prevalent in the data.

[1]Indeed, one can ignore this problem for the original ST test because in our simulations there is no tie for any replication of the test. However, this problem becomes important for the modification of the ST test which we will present in section 4.

2.2 Tests Using Moving Block Methods

Xu/ Fisher/ Willson test $H_0^i : (X \succ_i Y)$ vs. $H_1^i :$ (not H_0) for $i = 1, 2$ which can be written as

$$\begin{aligned} H_0^i &: \text{ For all } x \in \mathbb{R} : \; Q_X^{(i)}(x) \geq Q_Y^{(i)}(x) \\ \text{vs. } H_1^i &: \text{ There exists } x' \in \mathbb{R} : \; Q_X^{(i)}(x') < Q_Y^{(i)}(x'). \end{aligned}$$

They compute the difference of the empirical quantile functions $\hat{Q}_n^{(i)}$ $(i = 1, 2)$ at various grid points $p_1, \ldots, p_n$ satisfying $0 < p_1 < \cdots < p_n < 1$ and define $\hat{Q}_n^{(i)}(P) = [\hat{Q}_n^{(i)}(p_1), \ldots, \hat{Q}_n^{(i)}(p_n)]'$.

The test statistic is given by

$$T_i = \max_{q \geq 0} [(\hat{Q}_{X,n}^{(i)} - \hat{Q}_{Y,n}^{(i)} - q)' \hat{\Lambda}^{-1} (\hat{Q}_{X,n}^{(i)} - \hat{Q}_{Y,n}^{(i)} - q)]$$

where $\hat{\Lambda}$ is a consistent estimate of the covariance matrix Λ of $\hat{Q}_{X,n}^{(i)} - \hat{Q}_{Y,n}^{(i)}$. T_i is asymptotically distributed as a weighted sum of χ^2-variates. The weights are determined by Monte Carlo simulation using nonlinear programming.

In this procedure, the estimation of Λ is crucial. Xu/ Fisher/ Willson propose that moving block bootstrap (MBB) captures the dependence structure if the processes $(X_t)_{t \in \mathbb{Z}}$ and $(Y_t)_{t \in \mathbb{Z}}$ are strongly mixing (or α-mixing). The strong mixing coefficient of two sigma fields $\mathcal{A}$ and $\mathcal{B}$ is defined by

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{A}, B \in \mathcal{B}\}.$$

A sequence $(Z_k)_{k \in \mathbb{Z}}$ of random variables is strongly mixing if for the generated sigma fields $\mathcal{F}_a^b := \sigma(Z_k : a \leq k < b)$ the following holds:

$$\alpha(m) := \sup\{\alpha(\mathcal{F}_{-\infty}^k, \mathcal{F}_{k+m}^\infty) : k \in \mathbb{Z}\} \underset{m \to \infty}{\longrightarrow} 0.$$

The strong mixing coefficient $\alpha(m)$ of $(Z_k)_{k \in \mathbb{Z}}$ is defined for $m \in \mathbb{N}$. As stated in the previous section, a stationary GARCH process is strongly mixing.

MBB was developed by Künsch (1989) and Liu/ Singh (1992). In the last years a plethora of bootstrap methods has been developed which are constructed to capture the dependence structures emerging in time series; see e.g. Härdle/ Horowitz/ Kreiss (2003).

In contrast to the usual bootstrap introduced by Efron (1979) MBB does not resample single observations, but whole blocks of a fixed length b. For a sample of observations $(z_1, \ldots, z_n)$ denote the moving blocks as $B_1, \ldots, B_{n-b+1}$, where $B_j = (x_j, x_{j+1}, \ldots, x_{j+b-1})$ stands for the block consisting of b observations starting from x_j. One bootstrap resample consists of $k = \lfloor \frac{n}{b} \rfloor$ randomly resampled moving blocks where $\lfloor x \rfloor$ denotes the largest integer equal to or smaller than x.

The MBB estimate is consistent if $b(n)$ and $k(n)$ approach infinity with n approaching infinity. For a finite sample the choice of b is vital: on the one hand, a large value of b is necessary to capture strong dependence, while on the other hand, the number of blocks should also be large enough to reproduce the variability of the original sample.

Xu/ Fisher/ Willson proceed as follows: The observations of X and Y are resampled M times by MBB. For every resample Xu/ Fisher/ Willson compute the differences of the empirical quantile functions at the grid points. The empirical covariance matrix $\hat{\Lambda}$ of these vectors is taken as an estimator for Λ. In an empirical example, Xu/ Fisher/ Willson choose $M = 500$. We follow suit in our investigation.

Linton/ Maasoumi/ Whang test for stochastic maximality of a set of prospects. A set is stochastically maximal if no prospect is stochastically dominated by another prospect in the set. The test can be easily modified such that it also tests for stochastic dominance. The test problem is $H_0^i : (X \succ_i Y)$ vs. $H_1^i :$ (not H_0) for $i = 1, 2$ as in Xu/ Fisher/ Willson, but in contrast Linton/ Maasoumi/ Whang use the test statistic

$$T_{n,i} = \sup_{x \in \mathbb{R}} \sqrt{n}(\hat{F}_{X,n}^{(i)}(x) - \hat{F}_{Y,n}^{(i)}(x)).$$

In the study of Linton/ Maasoumi/ Whang (X_t) and (Y_t) are errors in a linear regression model. For the investigation of stochastic dominance some regularity conditions have to be satisfied. If we do not assume a regression model, the only persisting regularity condition is that $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are strongly mixing with $\alpha(m) = O(m^{-3})$. If $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are generated by a strictly stationary GARCH process with innovations ε_t satisfying $E|\varepsilon_t| < \infty$ for some $\delta > 0$, then they are strongly mixing with a geometric rate, i.e. $\alpha(m) = O(a^m)$ for some $a \in (0,1)$ (Davis/ Mikosch/ Basrak, 1999); thus $\alpha(m) = O(m^{-3})$ holds.

For the approximation of the distribution of $T_{n,i}$ under H_0^i Linton/ Maasoumi/ Whang use a subsampling method developed by Politis/ Romano (1994). An overview of resampling methods for various situations, e.g. stationary observations, is given by Politis/ Romano/ Wolf (1999). We outline the procedure for the test of SD1. Let

$$d_n(W_1, \ldots, W_n) = \frac{1}{\sqrt{n}} T_{n,1}$$

and $d_{n,b,k} = d_b(W_k, W_{k+1}, \ldots, W_{k+b-1})$ for $k = 1, \ldots, n-b+1$ be the transformed test statistic for the subsample $(W_k, W_{k+1}, \ldots, W_{k+b-1})$ of size b. Further let $g_{n,b}$ be the empirical quantile function of $\{\sqrt{b} d_{n,b,k} : k = 1, \ldots, n-b+1\}$ and g the quantile function of the asymptotic distribution of $T_{n,1}$ under H_0^1. Assume that $b(n) \underset{n\to\infty}{\longrightarrow} \infty$ and $\frac{b(n)}{n} \underset{n\to\infty}{\longrightarrow} 0$ and that the mixing condition stated above holds. For example, this will be the case for a stationary GARCH process.

Then under the subcase $F_X = F_Y$ of H_0^1 we have $g_{n,b}(1-\alpha) \xrightarrow{p} g(1-\alpha)$ and

$$P(T_{n,1} > g_{n,b}(1-\alpha)) \underset{n\to\infty}{\longrightarrow} \alpha.$$

Under H_1^1 the test is consistent, i.e.

$$P(T_{n,1} > g_{n,b}(1-\alpha)) \underset{n\to\infty}{\longrightarrow} 1.$$

The result concerning SD2 is analogous.

The described tests in this section are robust to contemporaneous correlation between the processes. Moreover, the XFW and the LMW tests are asymptotically robust to intertemporal dependence within the processes if the processes are strongly mixing. An important theoretical topic and problem for applications is the performance of these tests for finite samples if the data are dependent, in particular if they are conditionally heteroskedastic. This will be investigated in the next section.

3 Simulation Results for the Conventional Tests

By simulation we investigate the effect of some dependence structures on the size of the tests described in the previous section. The nominal size in each test is $\alpha = 0.05$. Unless stated differently, the sample size is $n = 1000$ and the number of replications is $R = 500$ for the ST test and $R = 1000$ for the XFW test and the LMW test. For our research concerning contemporaneous dependence we determine the size of the tests for various values of the correlation coefficient ρ. In exploring the effects of intertemporal dependence we confine ourselves to GARCH(1,1), a kind of conditional heteroskedasticity.

Though this is a special approach to modelling conditional heteroskedasticity, it is commonly believed that this model is well suited to be used on financial data. Akgiray (1989) concludes in an empirical study of the temporal behavior of daily stock market returns: "The conditional heteroskedastic processes … fit to data very satisfactorily. More importantly , they provide improved forecasts of volatility. Within the class of such models, GARCH(1,1) processes show the best fit and forecast accuracy."

We investigate the following situations:

- For every t the vector (X_t, Y_t) has a bivariate normal distribution with mean $(0 \quad 0)$ and covariance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. The process $(X_t, Y_t)_{t\in\mathbb{Z}}$ is serially independent.

- For every t the random variables X_t and Y_t are independent. Both random variables follow a GARCH(1,1) process which we define here as follows: Let

$(\varepsilon_t)_{t\in\mathbb{Z}}$ denote a sequence of independent and identically $\mathcal{N}(0,1)$-distributed random variables. Let $X_t = \sigma_t\varepsilon_t$ where

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

for $t \in \mathbb{Z}$ and $\alpha_0 > 0$, $\alpha_1, \beta_1 \geq 0$. Then (X_t) is called a GARCH(1,1) process.

- $(X_t, Y_t)_{t\in\mathbb{Z}}$ is a bivariate GARCH(1,1) process: Let $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ be independent and identically $\mathcal{N}(0, I_2)$-distributed random vectors for all $t \in \mathbb{Z}$; here I_n is the n-dimensional identity matrix. For $i = 1, 2$ define

$$h_{ii,t} = \alpha_0 + \alpha_1 Z_{i,t-1}^2 + \beta_1 h_{ii,t-1}$$

and

$$h_{12,t} = \rho\alpha_0 + \alpha_1 Z_{1,t-1}Z_{2,t-1} + \beta_1 h_{12,t-1}$$

and $H_t := \begin{pmatrix} h_{11,t} & h_{12,t} \\ h_{12,t} & h_{22,t} \end{pmatrix}$ where $Z_{1,t} = X_t$, $Z_{2,} = Y_t$. H_t is positive definite for all $t \in \mathbb{Z}$. Let $(X_t, Y_t)' = C_t\varepsilon_t$, where C_t is the positive definite matrix with $C_t^2 = H_t$ (the root of H_t). Then $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ follow a GARCH(1,1) process as described above. For every t the unconditional correlation between X_t and Y_t is ρ. For a more general approach and details concerning multivariate GARCH, see Engle/Kroner (1995).

First we investigate the ST test. On the recommendation of Schmid/ Trede the number of permutations is $M = 500$. As expected, the test is robust to contemporaneous correlation. We simulate two samples $(X_t : t = 1, \ldots, n)$ and $(Y_t : t = 1, \ldots, n)$. Within each sample, these data are independent; for every time t the observations (X_t, Y_t) follow a bivariate normal distribution with correlation coefficient ρ. The true size ranges from $\alpha = 0.05$ to $\alpha = 0.04$ if the correlation varies from $\rho = -0.5$ to $\rho = 0.5$ (see table 1). Hence the correlation does not have a significant effect on the size of the test. This was already asserted by Schmid and Trede.

ρ	-0.5	0	0.5
size	0.05	0.05	0.04

Table 1: Rejection probability of the ST test for the nominal value $\alpha = 0.05$. The processes are intertemporally independent and contemporaneously correlated with coefficient ρ. The sample size is $n = 1000$, the number of Monte Carlo replications is $R = 500$.

α_1	β_1	$\alpha_1 + \beta_1$	size
0.1	0.8	0.9	0.07
0	0.99	0.99	0.04
0.1	0.89	0.99	0.30
0.14	0.85	0.99	0.30

Table 2: Rejection probability of the ST test for the nominal value $\alpha = 0.05$. The processes are generated by contemporaneously independent GARCH(1,1) processes with parameters $\alpha_0 = 0.1$, α_1 and β_1. The sample size is $n = 1000$, the number of Monte Carlo replications is $R = 500$.

Concerning conditional heteroskedasticity within the samples things are much more involved. Table 2 presents the results for various values of the GARCH(1,1) parameters α_1 and β_1. The size of the test increases slowly in α_1 and β_1; however, the increase becomes faster for larger values of α_1 and β_1. If $\alpha_1 + \beta_1$ is close to 1, the true size of the test is much larger than the nominal size of the test. For $\alpha_1 = 0.14$ and $\beta_1 = 0.85$, which is an appropriate choice for financial data, the size is $\alpha = 0.30$. Hence the ST test should not be used if the data follow a GARCH(1,1) process with large parameters.

In our investigation of the XFW test we choose the grid points $(\frac{1}{K}, \ldots, \frac{K-1}{K})$ with $K = 10$ and $K = 20$. The performance depends on the length b of the blocks, whereas the number of grid points K does not have a significant effect.

ρ	-0.8	-0.5	0	0.5	0.8
size	0.06	0.06	0.06	0.07	0.05

Table 3: Rejection probability of the XFW test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and block length $b = 1$. The processes are intertemporally independent and contemporaneously correlated with parameter ρ. The number of grid points is $K - 1 = 9$, the number of Monte Carlo replications is $R = 1000$.

Hence we only report the results for $K = 10$. In the case of intertemporal independence within each sample the test works very well, even for $b = 1$, i.e. usual bootstrap. For various values of contemporaneous correlation ρ the size ranges between $\alpha = 0.05$ and $\alpha = 0.07$ (see table 3).

However, if the observations are generated by a GARCH(1,1) process, there is a notable effect on the size of the test. We focus on the parameter constellation $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. For $K = 10$ and $b = 1$ we get $\alpha = 0.62$, the size decreases for increasing b as can be seen ín table 4. However, for larger values of b the size increases, thus, for $b = 200$ the size is larger than for $b = 100$.

If the data are generated by a bivariate GARCH(1,1) process, the results are similar. For $\alpha_0 = 0.1$, $\alpha_1 = 0.14$, $\beta_1 = 0.85$ and $\rho = \pm 0.5$ we choose various values of the block length b; the results are also reported in table 4. As in the case of independent GARCH(1,1) processes, the size first decreases with increasing block length, but increases after the block length has exceeded a critical border. We observe the following results: For sample size $n = 1000$ we cannot manage to keep the nominal rejection probability α just by adjusting the size of the block length. Hence the asymptotic result of Xu/ Fisher/ Willson is not useful if the number of observations is equal to 1000 or even less!

For the LMW test the results are similar. Even for small block length the test performs very well in the case of intertemporal independence. The block length $b = 1$ is, by construction, no reasonable choice. Table 5 shows the results

ρ / b	-0.5	0	0.5
1	0.38	0.62	0.47
10	0.29	0.44	0.34
50	0.18	0.21	0.15
100	0.14	0.13	0.17
200	0.20	0.21	0.24

Table 4: Rejection probability of the XFW test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and various values of block length b. The processes are generated by bivariate GARCH(1,1) processes with correlation parameter ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of grid points is $K - 1 = 9$. The number of Monte Carlo replications is $R = 1000$.

for various block lengths and correlation parameters. For $b = 10$ the size lies between $\alpha = 0.06$ and $\alpha = 0.09$ for various values of the correlation coefficient ρ. If we choose $b = 40$, the nominal size is kept well: the size ranges from $\alpha = 0.04$ to $\alpha = 0.06$.

However, if the data are generated by a GARCH(1,1) process, there is a significant effect on the size of the test. Table 6 shows the same effect as in the ST test and the XFW test: If $\alpha_1 + \beta_1$ approaches 1, the nominal size is not kept any more.

Is the variation of the block length a remedy? No! As we see in table 7, for the GARCH parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$ and correlation coefficient $\rho = 0$ or $\rho = \pm 0.5$, the size first decreases with increasing block length, but then increases again. We do not find any block length such that the nominal size is kept. The LMW test should not be used if conditional heteroskedasticity is prevalent in the data and the number of observations is 1000 or less.

In many applications in finance a GARCH(1,1) process with large parame-

b \ ρ	-0.8	-0.5	0	0.5	0.8
10	0.07	0.06	0.08	0.09	0.06
20	0.07	0.06	0.04	0.06	0.06
40	0.04	0.04	0.04	0.04	0.06

Table 5: Rejection probability of the LMW test for the nominal value $\alpha = 0.05$, intertemporal independence, contemporaneous correlation ρ, sample size $n = 1000$ and block length b. The number of Monte Carlo replications is $R = 1000$.

α_1	β_1	$\alpha_1 + \beta_1$	size
0	0.99	0.99	0.05
0.1	0.8	0.9	0.08
0.1	0.89	0.99	0.35
0.14	0.85	0.99	0.28

Table 6: Rejection probability of the LMW test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and block length $b = 20$. The processes are generated by contemporaneously independent GARCH(1,1) processes with parameters $\alpha_0 = 0.1$, α_1 and β_1. The number of Monte Carlo replications is $R = 1000$.

b \ ρ	-0.5	0	0.5
10	0.31	0.38	0.37
20	0.25	0.28	0.30
40	0.16	0.20	0.19
100	0.10	0.10	0.11
200	0.07	0.09	0.09
250	0.07	0.07	0.07
300	0.09	0.07	0.08
500	0.08	0.13	0.11

Table 7: Rejection probability of the LMW test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and various values of block length b. The processes are generated by bivariate GARCH(1,1) processes with correlation parameter ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of Monte Carlo replications is $R = 1000$.

ters, i.e. $\alpha_1 + \beta_1$ close to 1, is a good fit of the time series. The simulations indicate that the dominance tests of Schmid/ Trede, Xu/ Fisher/ Willson and Linton/ Maasoumi/ Whang are not useful in this case if there are less than 1000 observations.

4 Circular Block Methods as an Alternative Concept

Due to the moderate success of the moving block methods we propose another kind of block methods: circular block methods. The circular block bootstrap (CBB) was developed by Politis/ Romano (1992). As the MBB method, CBB resamples overlapping blocks of observations which are of a fixed length b. One problem of MBB is that the observations at the beginning and the end of the time series are considered less. CBB solves this problem as follows. The collection of blocks from which it is resampled consists of the blocks $B_1, \ldots, B_{n-b+1}$ of the MBB and additionally of the blocks $B_{n-b+2}, \ldots, B_n$ of the form $B_k = (x_k, \ldots, x_n, x_1, \ldots, x_{k+b-n-1})$. Lahiri (1999) investigated the asymptotic behavior of some block bootstrap methods and found that MBB and CBB are asymptotically equivalent. We apply CBB to the XFW test and investigate by simulation whether this improves the size of the test.

The modification of the subsampling method of Linton/ Maasoumi/ Whang is analogous. The distribution of $T_{n,i}$ under H_0^i is approximated by $\sqrt{b}d_{n,b,k}$ where

$$d_{n,b,k} = \begin{cases} d_b(W_k, W_{k+1}, \ldots, W_{k+b-1}) & \text{for} \quad k = 1, \ldots, n-b+1, \\ d_b(W_k, \ldots, W_n, W_1, \ldots, W_{k+b-n-1}) & \text{for} \quad k = n-b+2, \ldots, n. \end{cases}$$

By some modification of the proofs we can show that the main results of Linton/ Maasoumi/ Whang still hold if we use the modified subsampling method. The theorem is shown in the appendix.

Theorem 1. *Let* $(X_t)_{t\in\mathbb{Z}}$ *and* $(Y_t)_{t\in\mathbb{Z}}$ *be strongly mixing with* $\alpha(m) = O(m^{-3})$*. Assume* $b(n) \to \infty$ *and* $\frac{b(n)}{n} \to 0$ *as* $n \to \infty$*. Let* $\alpha \in (0,1)$*,* $g_{n,b}$ *be the empirical quantile function of* $\{\sqrt{b}d_{n,b,k} : k = 1,\ldots,n\}$ *and* g *the quantile function of the asymptotic distribution of* $T_{n,1}$ *under* H_0^1*. Then:*

1. *Under the subcase* $F_X = F_Y$ *of* H_0^1 *we have* $g_{n,b}(1-\alpha) \xrightarrow{p} g(1-\alpha)$ *and* $P(T_{n,1} > g_{n,b}(1-\alpha)) \underset{n\to\infty}{\longrightarrow} \alpha$*, i.e. asymptotically the test keeps the size* α*.*

2. *Under* H_1^1 *we have* $P(T_{n,1} > g_{n,b}(1-\alpha)) \underset{n\to\infty}{\longrightarrow} 1$*, i.e. the test is consistent.*

By simulation we investigate if the modified subsampling method improves the performance of the LMW test for finite samples.

In contrast to the XFW and the LMW test, the ST test does not consider any intertemporal dependence at all. In this paper, we modify the permutation test to a block permutation test. As the block bootstrap and subsampling methods, the block permutation reproduces the dependence structure of the observations.

The block permutation method is performed as follows. We consider the random variable $U = |\{i \in \{1,\ldots,n\} : X_i \text{ and } Y_i \text{ are transposed}\}|$. In the permutations of Schmid/ Trede, U follows a binomial distribution with parameters n and $\frac{1}{2}$. Therefore, in the modified test we first generate for every permutation the number u of the transposed pairs. If the given block length is b, we choose by chance $\lfloor \frac{u}{b} \rfloor$ blocks of length b and one block of length $u - b\lfloor \frac{u}{b} \rfloor$ for which X_i and Y_i are transposed. The test is further performed as described in Schmid/ Trede.

In the modified test the tie described in section 2.1 sometimes arises. One could attain the size α exactly by randomization, but this could be unacceptable for applications. Hence we decide as follows in the case of a tie. Let $k_1 := \max\{k \in \mathbb{Z} : T = T^{(k)}\}$ and $k_2 := \min\{k \in \mathbb{Z} : T = T^{(k)}\}$ and reject H_0 if and only if $\frac{1}{2}(k_1 + k_2) > (1-\alpha)M$.

We investigate by simulation whether this block permutation improves the ST test for finite samples with conditional heteroskedasticity.

In this section, we develop some modifications of the tests of Schmid/ Trede, Xu/ Fisher/ Willson and Linton/ Maasoumi/ Whang. We find that the circular block methods are asymptotically equivalent to their moving block counterparts. For finite samples, there are two opposed effects. On the one hand, the observations at the beginning and at the end of the time series are considered as much as the observations in the middle. This is an improvement to moving block bootstrap and subsampling. On the other hand, some blocks we build are no reasonable construction in terms of reproduction of the dependence structure. In strongly mixing processes the observations with a large time lag are nearly independent. Therefore a block consisting of the first and the last observations does not seem to make sense. But we can reply to this objection that a block of the first k and the last $b-k$ observations is just a combination of two blocks with a strong dependence structure. In other words, the resample consists of blocks with different lengths and a strong dependence within each block.

On the basis of these considerations, one can expect that the circular block methods improve the performance of the tests. In the next section, we investigate this by means of simulation.

5 Simulation Results Using Circular Block Methods

In this section we report on the simulation results of the modified tests. We refer to the modified tests as STm test, XFWm test and LMWm test.

The modification described in the previous section is a remedy for the ST test. As can be seen in table 8, the modification does not destroy the good result for data which are serially independent, but contemporaneously correlated. The nominal size $\alpha = 0.05$ is kept well, the test is just a bit too conservative.

But the modification is a real improvement. In section 3 we stated that the ST

ρ	-0.5	0	0.5
size	0.03	0.03	0.03

Table 8: Rejection probability of the STm test for the nominal value $\alpha = 0.05$ and sample size $n = 1000$. The processes are intertemporally independent and contemporaneously correlated with coefficient ρ. The block length is $b = 100$, the number of Monte Carlo replications is $R = 500$.

ρ / b	-0.5	0	0.5
100	0.09	0.14	0.13
200	0.07	0.08	0.08
300	0.04	0.05	0.06
500	0.00	0.01	0.00

Table 9: Rejection probability of the STm test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and various values of block length b. The processes are generated by bivariate GARCH(1,1) with correlation coefficient ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of Monte Carlo replications is $R = 500$.

test does not keep the nominal size if the data are generated by a GARCH(1,1) process with parameters whose sum is close to 1. Table 9 shows that the STm test keeps the nomimal size if the block length is chosen appropriately. This also holds for the bivariate GARCH(1,1) process with correlation coefficient ρ. The choice $b = 200$ is not sufficient to keep the level; for $b = 500$ the test is too conservative. According to table 9, $b = 300$ seems to be a good choice.

The XFW test cannot be improved significantly by modifying it with circular block bootstrap. At least, like the original test, the modified version still keeps the size if the observations are serially independent, but contemporaneously cor-

ρ	-0.8	-0.5	0	0.5	0.8
size	0.06	0.06	0.05	0.07	0.05

Table 10: Rejection probability of the XFWm test for the nominal value $\alpha = 0.05$ and sample size $n = 1000$. The processes are intertemporally independent and contemporaneously correlated with parameter ρ. The number of grid points is $K - 1 = 9$, the block length $b = 1$. The number of Monte Carlo replications is $R = 1000$.

related. Table 10 presents the simulation results for $K = 10$, block length $b = 1$ and various values of the correlation coefficient ρ.

However, if the data are generated by a GARCH(1,1) process, the circular block bootstrap is no remedy for the XFW test. The simulation results (table 11) show that there is no block length for which the nominal size is kept. As for the original versions of the XFW and LMW test, for small block length b the size decreases with increasing b, but is always significantly higher than the nominal size. For $b > 100$ the size increases with increasing b. This holds for both independent GARCH(1,1) processes and bivariate GARCH(1,1) processes with contemporaneous correlation.

The LMWm test keeps the size for appropriate block length. Table 12 shows the simulation results for serially independent time series and various values of contemporaneous correlation ρ. The block length is $b = 40$. As for the original LMW test, the nominal size $\alpha = 0.05$ is kept well. Therefore, the performance of the modified LMW test at least is not worse than that of the original test.

But the modified version is even better. With the appropriate block length it keeps the nominal size. The simulation results for a bivariate GARCH(1,1) process with parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$ and correlation $\rho = 0$ and $\rho = \pm 0.5$ are reported in table 13. In contrast to the original LMW test, the size decreases monotonically with increasing b. If we choose block length $b = 300$,

b \ ρ	-0.5	0	0.5
10	0.28	0.44	0.34
50	0.17	0.21	0.17
100	0.16	0.17	0.17
200	0.19	0.19	0.22
500	0.42	0.49	0.52

Table 11: Rejection probability of the XFWm test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and various values of block length b. The processes are generated by bivariate GARCH(1,1) with correlation parameter ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of grid points is $K - 1 = 9$. The number of Monte Carlo replications is $R = 1000$.

ρ	-0.8	-0.5	0	0.5	0.8
size	0.04	0.04	0.04	0.05	0.06

Table 12: Rejection probability of the LMWm test for the nominal value $\alpha = 0.05$, sample size $n = 1000$ and block length $b = 40$. The processes are intertemporally independent and contemporaneously correlated with parameter ρ. The number of Monte Carlo replications is $R = 1000$.

b \ ρ	-0.5	0	0.5
100	0.09	0.11	0.09
200	0.05	0.07	0.06
300	0.05	0.05	0.05
500	0.03	0.03	0.04

Table 13: Rejection probability of the LMWm test for the nominal value $\alpha = 0.05$ and various values of block length b. The processes are generated by bivariate GARCH(1,1) with correlation parameter ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of Monte Carlo replications is $R = 1000$.

the size is kept well, whereas for $b = 500$ the test is too conservative. As for the ST test, $b = 300$ seems to be a good choice.

We have seen that for sample size $n = 1000$ the STm and LMWm tests keep the size if we choose the appropriate block length. If the sample size varies, which block length is the best choice? We explore this question with the help of some further simulations. Tables 14 and 15 show the corresponding results for the STm and the LMWm test.

For the STm test, the optimal block length seems to increase with rate $\sqrt{n}$. The block length $b = 150$ is a good choice for $n = 250$ whereas the size cannot be kept for $b = 100$ or $b = 200$. With increasing sample size the block length has to increase. The block length $b = 300$ yields a bad result for $n = 2500$, but the performance is much better for $b = 500$. This implies an optimal block length of approximately $b(n) \approx 10\sqrt{n}$.

We see that for the sample size $n = 4000$ the LMWm test performs well if we choose the block length $b = 300$. This result suggests that for the considered dependence structure and a sample size larger than $n = 1000$ the increase of the optimal block length is very slow. Furthermore, the range of block lengths with

ρ		-0.5	0	0.5
n	b			
250	100	0.14	0.12	0.18
250	150	0.06	0.03	0.03
250	200	0.12	0.17	0.20
2500	300	0.12	0.12	0.11
2500	500	0.05	0.05	0.04

Table 14: Rejection probability of the STm test for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The processes are generated by bivariate GARCH(1,1) with correlation coefficient ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of Monte Carlo replications is $R = 500$.

good performance becomes larger with increasing sample size: $b = 600$ is still a reasonable choice for $n = 4000$. However, for smaller block lengths the choice of the block length is more critical. For $n = 250$ the block length $b = 150$ is a good choice whereas $b = 100$ and $b = 200$ yield rather poor results. These results imply that at least for smaller sample sizes the optimal block length increases with the rate $\sqrt{n}$, the optimal block length is approximately $b(n) \approx 10\sqrt{n}$, as for the STm test.

Summing up the results of this section, the modifications of the considered tests are successful for the tests of Schmid/ Trede and of Linton/ Maasoumi/ Whang, whereas it does not improve the performance of the test of Xu/ Fisher/ Willson. This test cannot be improved significantly by moving block bootstrap, but block permutation makes the permutation test of Schmid/ Trede robust to conditional heteroskedasticity, and circular subsampling improves the performance of the test of Linton/ Maasoumi/ Whang for finite samples. The choice of the appropriate block length is crucial. For both the modified ST and LMW

n	b ρ	-0.5	0	0.5
250	100	0.09	0.08	0.10
250	150	0.04	0.04	0.05
250	200	0.01	0.01	0.01
500	220	0.04	0.06	0.06
4000	300	0.04	0.05	0.05
4000	600	0.04	0.03	0.05

Table 15: Rejection probability of the LMWm test for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The processes are generated by bivariate GARCH(1,1) with correlation parameter ρ and parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.14$ and $\beta_1 = 0.85$. The number of Monte Carlo replications is $R = 1000$.

tests the optimal block length is approximately $b(n) = 10\sqrt{n}$.

6 Power Investigation

In the previous sections we modified the tests of Schmid/ Trede on SD2 and Linton/ Maasoumi/ Whang on SD1 successfully. We investigated the optimal block lengths for various sample sizes.

In this section we explore the power of the tests. The main problem is the shape of the alternative H_1. There are many combinations of distributions F_X, F_Y such that $F_X(x) > F_Y(x)$ for at least one $x \in \mathbb{R}$ (analogous for $F^{(2)}$). Hence the alternative $H_1 : X \not\succeq Y$ is very complex.

We begin with the investigation of some scale alternatives. For both tests we consider the alternative

$$H_1(\sigma_X) : F_X = \mathcal{N}(0, \sigma_X), F_Y = \mathcal{N}(0, 1)$$

and vary σ_X from 1.1 to 1.5 in 0.1 steps. Further, we analyze the alternative

$$H_1(\sigma_Y) : F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(0,\sigma_Y)$$

with $\sigma_Y = 1.1, 1.2, \ldots, 1.5$ only for the LMWm test on SD1. Note that $F_X = \mathcal{N}(0,1)$, $F_Y = \mathcal{N}(0,\sigma_Y)$ with $\sigma_Y > 1$ is in the null hypothesis for SD2. For both tests we consider the sample sizes $n = 250, 1000, 2500$ and the values of the block length b which are found to be optimal in the previous section. The samples are generated by contemporaneously and intertemporally independent processes. The number of Monte Carlo replications is $R = 100$. This small number causes a weak significance of the simulation results, but we can at least take them as a general tendency.

Tables 16, 17 and 18 show the results. As one might expect, the power increases with increasing standard deviation σ_Z $(Z = X, Y)$ and increasing sample size n. The larger σ_Z, the larger is the distance to H_0. Hence the violation of the null hypothesis is detected more probably for larger σ_Z. With larger sample size the consistency of the tests becomes more efficient. For $n = 250$ the results are not that satisfactory whereas for $n = 2500$ the rejection rate tends toward 1 very fast with growing σ_Z.

Further we explore the power of the tests for the location alternative $H_1 : F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(0.1,1)$ where the observations are contemporaneously and intertemporally independent. Here Y dominates X in sense of SD1 and SD2, hence the dominance of X has to be rejected. The results are presented in table 19. For $n = 2500$ the power is very high, for the STm test on SD2 $n = 1000$ suffices to give good results. The low power for smaller sample size is caused by the fact that the deviation from equality (and therefore weak dominance) is small.

Finally we analyze the power of the tests if the observations are contemporaneously and intertemporally independent and normally distributed and differ in both mean and variance. We consider the cases $F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(1,2)$,

		σ_X	1.1	1.2	1.3	1.4	1.5
n	b						
250	150		0.15	0.24	0.39	0.48	0.64
1000	300		0.16	0.52	0.85	0.96	1.00
2500	500		0.42	0.94	1.00	1.00	1.00

Table 16: Power of the STm test on SD2 for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The alternative considered is $H_1(\sigma_X) : F_X = \mathcal{N}(0, \sigma_X), F_Y = \mathcal{N}(0, 1)$ where the processes (X_t) and (Y_t) are contemporaneously and intertemporally independent. The number of Monte Carlo replications is $R = 100$.

		σ_X	1.1	1.2	1.3	1.4	1.5
n	b						
250	150		0.03	0.06	0.08	0.21	0.27
1000	300		0.11	0.43	0.67	0.92	0.97
2500	300		0.29	0.96	1.00	1.00	1.00

Table 17: Power of the LMWm test on SD1 for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The alternative considered is $H_1(\sigma_X) : F_X = \mathcal{N}(0, \sigma_X), F_Y = \mathcal{N}(0, 1)$ where the processes (X_t) and (Y_t) are contemporaneously and intertemporally independent. The number of Monte Carlo replications is $R = 100$.

		σ_Y	1.1	1.2	1.3	1.4	1.5
n	b						
250	150		0.00	0.01	0.08	0.10	0.19
1000	300		0.07	0.45	0.73	0.94	0.99
2500	300		0.24	0.89	1.00	1.00	1.00

Table 18: Power of the LMWm test on SD1 for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The alternative considered is $H_1(\sigma_Y) : F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(0,\sigma_Y)$ where the processes (X_t) and (Y_t) are contemporaneously and intertemporally independent. The number of Monte Carlo replications is $R = 100$.

		Test	LMWm SD1	STm SD2
n	b			
250	150		0.05	0.39
1000	300		0.29	0.69
2500	300/500		0.79	0.97

Table 19: Power of the LMWm test on SD1 and STm test on SD2 for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The alternative considered is $H_1 : F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(0.1,1)$ where the processes (X_t) and (Y_t) are contemporaneously and intertemporally independent. The number of Monte Carlo replications is $R = 100$.

$F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(-1,2)$ and vice versa. $F_X = \mathcal{N}(0,1), F_Y = \mathcal{N}(-1,2)$ is the only combination of distributions where X dominates Y in the sense of SD2, whereas SD1 does not hold in any of these cases.

The tables refpowerlocscalestm and 21 show the results. For both tests and all alternatives the power is very good for the sample size $n = 2500$; for the LMWm test $n = 1000$ already gives satisfactory results. In addition to this, both tests have good power even for $n = 250$ if $F_X = \mathcal{N}(-1,2)$, $F_Y = \mathcal{N}(0,1)$ or if $F_X = \mathcal{N}(0,1)$, $F_Y = \mathcal{N}(1,2)$ hold. In these cases, X has a smaller mean and the dominance of X in the sense of SD1 or SD2 can therefore clearly be rejected. The case $F_X = \mathcal{N}(1,2)$, $F_Y = \mathcal{N}(0,1)$ is most critical. The LMWm test has low power for $n = 250$, the STm test even for $n = 1000$. If X and Y are distributed like this, $F_X(x) \leq F_Y(x)$ holds for $x \geq -1$. The fatter left tail of X is the reason why X does not dominate Y in the sense of SD1 or SD2. However, in small samples the number of samples in the tails is very small. Hence the violation of SD is often not detected whereas for larger samples this problem becomes more and more negligible. In the case $F_X = \mathcal{N}(0,1)$, $F_Y = \mathcal{N}(-1,2)$ X dominates Y in the sense of SD2, but not of SD1. The lower right tail is the reason why X does not dominate Y in the sense of SD1. The LMWm test does not detect the violation of SD1 for $n = 250$ because there are only few observations belonging to the left tail for this small sample. For $n = 1000$ and $n = 2500$ this problem does not occur any more.

In this section we analyze the power of the tests we developed in this paper. Among the plethora of alternatives we confine ourselves to the case of independent, normally distributed observations. We see that the STm test has good power for $n = 2500$, for the LMWm test we get satisfactory results even for $n = 1000$. The larger the distance to H_0, the higher is the power. For some alternatives the power is close to one even for sample size $n = 250$.

		F_X	$\mathcal{N}(0,1)$	$\mathcal{N}(1,2)$	$\mathcal{N}(-1,2)$
		F_Y	$\mathcal{N}(1,2)$	$\mathcal{N}(0,1)$	$\mathcal{N}(0,1)$
n	b				
250	150		1.00	0.01	1.00
1000	300		1.00	0.04	1.00
2500	500		1.00	0.64	1.00

Table 20: Power of the STm test on SD2 for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The alternative considered is the combination of two normal distributions with different means and variances where the processes (X_t) and (Y_t) are contemporaneously and intertemporally independent. The number of Monte Carlo replications is $R = 100$.

		F_X	$\mathcal{N}(0,1)$	$\mathcal{N}(1,2)$	$\mathcal{N}(0,1)$	$\mathcal{N}(-1,2)$
		F_Y	$\mathcal{N}(1,2)$	$\mathcal{N}(0,1)$	$\mathcal{N}(-1,2)$	$\mathcal{N}(0,1)$
n	b					
250	150		1.00	0.08	0.11	0.99
1000	300		1.00	0.88	0.87	1.00
2500	300		1.00	1.00	1.00	1.00

Table 21: Power of the LMWm test on SD1 for the nominal value $\alpha = 0.05$ and various values of sample size n and block length b. The alternative considered is the combination of two normal distributions with different means and variances where the processes (X_t) and (Y_t) are contemporaneously and intertemporally independent. The number of Monte Carlo replications is $R = 100$.

7 Testing for Stochastic Dominance in German Stock Returns

In the following we apply the modified ST and LMW tests to the daily returns of the 30 stocks contained in the German stock index DAX. After a descriptive comparison we test on stochastic dominance between each pair of stocks. From the test results we get the efficient sets. A random variable is called efficient if it is not dominated.

The daily spot stock prices p_t are taken from Datastream. The returns at day t are defined by $r_t = \ln(\frac{p_t}{p_{t-1}})$. In our investigation we consider the 10-year period between 16 September 1994 and 15 September 2004 and the 1-year period between 16 September 2003 and 15 September 2004. There are 2522 observations for the 10-year period and 255 observations for the 1-year period. For the 10-year period we only consider the stocks which were inclosed in the DAX for the entire period.

The tables 22 and 31 show the annualized means and standard deviations of the stock returns. In these tables the abbreviations of the firms, which are also used in the other tables, are specified. 8 of 30 stocks have negative mean returns in the 1-year period, this number diminishes to 4 of 22 for the 10-year period. The stocks of the travel agency TUI had the strongest decline which might have been caused by the tourism crisis after September 11, 2001.

As mentioned in the previous sections, financial data feature contemporaneous correlation and conditional heteroskedasticity. The estimated correlations are summarized in the tables 23 and 24 for the 1-year period and in the tables 32 and 33 for the 10-year period. The stocks are all positively correlated with each other, the correlations range from 0.11 to 0.76 for the 1-year period and from 0.15 to 0.74 for the 10-year period. The estimations for the parameters α_1 and β_1 in the GARCH model $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ are presented in the descriptive statistic results in tables 22 and 31 for the 1-year and 10-year period, respectively. For many stocks the sum $\alpha_1 + \beta_1$ is close to unity. This phenomenon is more broad for the 10-year period.

We start the investigation of stochastic dominance with a descriptive comparison. X descriptively dominates Y in the sense of SDi if for the empirical distribution functions $\hat{F}^{(i)}_{X,n}(x) \leq \hat{F}^{(i)}_{Y,n}(x)$ holds for all $x \in \mathbb{R}$. Descriptively, no dominance relationship in the sense of SD1 can be established between any pair of return series, neither for the 1-year nor for the 10-year period. Every pair of empirical distribution functions crosses at least once. Hence all stocks are SD1 efficient descriptively.

Concerning SD2 the findings are different. The results are reported in the tables 27 and 28 for the 1-year period and in the table 35 for the 10-year period. In the descriptive sense, SD2 can be established in 187 of 870 comparisons for the 1-year period and in 91 of 462 comparisons for the 10-year period. For the 1-year period there are only 4 of 30 stocks which are efficient: Adidas-Salomon, BASF, Continental, RWE. 8 of 22 stocks are efficient for the 10-year period: Altana, BASF, Continental, Eon, Henkel, RWE, SAP, Schering. For a larger sample size, a descriptive dominance relationship is harder to establish. Hence the fraction of the SD2 efficient stocks becomes larger with increasing length of the period.

The problem of too large efficient sets was already mentioned by Nelson/ Pope (1991) and Stein/ Pfaffenberger/ Kumar (1983). For getting established results concerning stochastic dominance we use the LMWm test on SD1 and the STm test on SD2. For the 1-year period, we choose the block length $b = 150$ for all tests. Due to the fact that $b = 300$ is a good choice for the LMWm test and both $n = 1000$ and $n = 4000$, we choose $b = 300$ for the application of the LMWm test to the 10-year period. For this period we choose $b = 500$ for the STm test because the simulations show that this is an appropriate block length for $n = 2500$. In addition to this, this is the recommended block length if we apply the rule $b(n) = 10\sqrt{n}$ for the appropriate block length.

The tables 25, 26, 29 and 30 show the test results for the 1-year period, the tables 34 and 36 for the 10-year period. We find that in most cases dominance cannot be rejected to the size $\alpha = 0.05$. For many comparisons this holds in both directions. For instance, the application of the LMWm test on SD1 for the 1-year period (see table 25) yields no rejection for dominance of Adidas-Salomon against Allianz and vice versa. This finding suggests that the empirical distributions are very close to each other.

From the obtained dominance results we determine the efficient sets in the following way. If $X \succ_i Y$ is rejected whereas $Y \succ_i X$ is not and Y has larger mean than X, SDi of Y against X is established. A stock is called efficient if it is not dominated in this sense. The condition concerning the means is required in order to prevent the paradox result that a stock dominates another one with larger mean. A necessary condition for stochastic dominance of any order is that the mean of the dominant random variable has to be at least as large as the mean of the dominated one.

Table 37 summarizes the results. For the 1-year period Adidas-Salomon, Continental, Eon and RWE are SD1 and SD2 efficient whereas Altana, Deutsche Bank, Deutsche Post, Metro and Siemens are only SD1 efficient. The SD1 efficient stocks for the 10-year period are Altana, BASF, Continental, Henkel, SAP, Schering and Siemens. These stocks are also SD2 efficient, furthermore Eon, Linde and RWE. It seems to be paradox that for the 10-year period some stocks are found to be SD2 efficient, but not SD1 efficient. This might be due to sampling errors.

Note that we have to take the efficiency results with a pinch of salt. The considered tests do not assert significantly that one random variable dominates another one, instead, the tests do or do not reject the hypothesis of dominance. In many cases the empirical distributions are very close to each other, therefore often the test cannot reject stochastic dominance.

Further we have to be aware of the fact that even a dominated stock can be a useful member of a portfolio. Diversification diminishes risk, and this effect can be stronger than the one caused by stochastic dominance. Hence, in many cases, a dominated stock should not be eliminated from a portfolio.

Name of Stock		Mean ×100	Std.dev. ×100	GARCH parameters α_1	β_1
Adidas-Salomon	ADS	37.102	19.653	0.0000	0.0051
Allianz	ALL	1.428	27.287	0.1131	0.5529
Altana	ALT	−13.082	22.955	0.2275	0.4364
BASF	BAS	10.990	20.491	0.0802	0.8882
Bayer	BAY	9.588	26.676	0.0827	0.7675
Bay. Hypo-Vereinsbank	BHV	9.869	37.097	0.0861	0.8268
BMW	BMW	−2.270	23.468	0.0732	0.8407
Commerzbank	CBK	9.384	29.970	0.1038	0.8351
Continental	CON	63.495	26.080	0.0847	0.8184
Daimler-Chrysler	DAC	4.386	24.242	0.0870	0.8453
Deutsche Bank	DBK	4.208	24.454	0.0207	0.7580
Deutsche Boerse	DBO	−6.987	23.102	0.0503	0.0000
Deutsche Lufthansa	DLH	−15.810	28.634	0.0536	0.7385
Deutsche Post	DPO	10.685	26.732	0.0287	0.9681
Deutsche Telekom	DTL	8.772	20.466	0.0375	0.8183
Eon	EON	27.285	19.744	0.0185	0.9781
Fresenius	FRE	17.162	21.091	0.0154	0.9816
Henkel	HEN	4.080	20.561	0.0776	0.7292
Infineon	INF	−41.718	35.534	0.0495	0.8289
Linde	LIN	17.162	22.209	0.0621	0.8004
MAN	MAN	32.334	30.329	0.0977	0.7602
Metro	MET	17.416	25.109	0.0912	0.8217
Muenchner Rueckvers.	MRV	−11.271	26.067	0.1203	0.7791
RWE	RWE	46.308	23.150	0.1148	0.3936
SAP	SAP	11.297	31.008	0.1187	0.7656
Schering	SCH	22.797	21.511	0.0642	0.8161
Siemens	SIE	9.894	25.184	0.0000	0.9991
Thyssen-Krupp	TYK	17.416	28.330	0.0293	0.9670
TUI	TUI	−5.075	32.008	0.1079	0.4465
Volkswagen	VW	−28.229	24.105	0.0337	0.9288

Table 22: Descriptive statistics and estimated GARCH parameters of the *annualized* daily returns of DAX stocks for the 1-year period.

	ADS	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DAC	DBK	DBO	DLH	DPO	DTL
ADS		36	18	37	34	30	27	33	29	32	34	19	25	25	39
ALL	36		27	64	63	61	62	64	52	67	66	33	63	52	66
ALT	18	27		17	17	22	25	22	15	18	23	13	20	12	19
BAS	37	64	17		71	46	59	52	50	60	62	24	50	50	63
BAY	34	63	17	71		46	56	52	52	54	56	20	50	42	61
BHV	30	61	22	46	46		46	67	43	50	52	31	48	43	49
BMW	27	62	25	59	56	46		49	56	69	53	22	52	40	51
CBK	33	64	22	52	52	67	49		46	52	60	33	49	49	56
CON	29	52	15	50	52	43	56	46		55	46	14	42	43	48
DAC	32	67	18	60	54	50	69	52	55		54	26	53	39	53
DBK	34	66	23	62	56	52	53	60	46	54		32	49	53	57
DBO	19	33	13	24	20	31	22	33	14	26	32		32	31	32
DLH	25	63	20	50	50	48	52	49	42	53	49	32		49	52
DPO	25	52	12	50	42	43	40	49	43	39	53	31	49		49
DTL	39	66	19	63	61	49	51	56	48	53	57	32	52	49	
EON	27	49	27	61	50	41	47	44	43	46	46	21	42	39	47
FRE	24	32	26	40	32	27	36	33	31	37	29	15	27	28	29
HEN	27	44	27	35	38	34	37	39	34	35	39	19	35	36	30
INF	31	55	25	49	47	50	44	49	32	45	52	29	55	45	52
LIN	33	49	19	52	48	40	47	48	39	47	50	31	47	46	44
MAN	26	58	16	51	53	41	60	46	46	51	47	28	58	42	49
MET	27	58	25	58	46	47	46	48	44	50	48	22	43	44	46
MRV	36	76	21	63	61	58	54	63	49	60	58	25	56	46	60
RWE	34	48	19	52	44	37	45	38	37	42	45	22	37	42	46
SAP	28	50	12	50	51	46	46	47	37	43	50	30	47	44	52
SCH	21	31	36	30	35	33	33	39	34	30	28	14	21	21	29
SIE	39	74	24	70	66	55	61	59	53	63	66	31	61	54	65
TYK	37	66	19	58	55	56	60	52	58	60	52	17	58	49	58
TUI	30	56	11	49	46	38	46	41	39	48	46	18	52	44	54
VW	33	65	20	59	57	55	75	50	61	69	59	27	57	49	51

Table 23: Correlation coefficients of the daily returns of DAX stocks $\times$ 100 for the 1-year period.

	EON	FRE	HEN	INF	LIN	MAN	MET	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ADS	27	24	27	31	33	26	27	36	34	28	21	39	37	30	33
ALL	49	32	44	55	49	58	58	76	48	50	31	74	66	56	65
ALT	27	26	27	25	19	16	25	21	19	12	36	24	19	11	20
BAS	61	40	35	49	52	51	58	63	52	50	30	70	58	49	59
BAY	50	32	38	47	48	53	46	61	44	51	35	66	55	46	57
BHV	41	27	34	50	40	41	47	58	37	46	33	55	56	38	55
BMW	47	36	37	44	47	60	46	54	45	46	33	61	60	46	75
CBK	44	33	39	49	48	46	48	63	38	47	39	59	52	41	50
CON	43	31	34	32	39	46	44	49	37	37	34	53	58	39	61
DAC	46	37	35	45	47	51	50	60	42	43	30	63	60	48	69
DBK	46	29	39	52	50	47	48	58	45	50	28	66	52	46	59
DBO	21	15	19	29	31	28	22	25	22	30	14	31	17	18	27
DLH	42	27	35	55	47	58	43	56	37	47	21	61	58	52	57
DPO	39	28	36	45	46	42	44	46	42	44	21	54	49	44	49
DTL	47	29	30	52	44	49	46	60	46	52	29	65	58	54	51
EON		42	41	33	46	36	45	53	69	34	36	46	52	44	49
FRE	42		23	26	25	30	30	29	32	30	31	34	37	29	39
HEN	41	23		30	45	41	38	45	32	29	33	41	37	26	41
INF	33	26	30		39	45	42	53	29	66	25	67	50	43	49
LIN	46	25	45	39		48	44	47	43	42	36	51	49	41	50
MAN	36	30	41	45	48		39	49	37	47	28	61	59	44	59
MET	45	30	38	42	44	39		52	37	36	29	54	48	36	50
MRV	53	29	45	53	47	49	52		47	47	29	66	58	47	59
RWE	69	32	32	29	43	37	37	47		33	32	41	45	39	45
SAP	34	30	29	66	42	47	36	47	33		26	63	47	42	49
SCH	36	31	33	25	36	28	29	29	32	26		34	25	25	33
SIE	46	34	41	67	51	61	54	66	41	63	34		67	52	66
TYK	52	37	37	50	49	59	48	58	45	47	25	67		52	61
TUI	44	29	26	43	41	44	36	47	39	42	25	52	52		48
VW	49	39	41	49	50	59	50	59	45	49	33	66	61	48	

Table 24: Correlation coefficients of the daily returns of DAX stocks × 100 for the 1-year period.

	ADS	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DAC	DBK	DBO	DLH	DPO	DTL
ADS		0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALL	0		0	0	0	0	0	0	0	0	0	0	0	0	0
ALT	0	0		0	0	0	0	0	0	0	0	0	0	0	0
BAS	0	1	0		0	0	0	0	0	0	0	0	0	0	0
BAY	1	0	0	0		0	0	0	1	0	0	0	0	0	0
BHV	1	1	0	1	0		1	0	0	0	0	0	0	0	1
BMW	1	0	0	0	0	0		0	0	0	1	0	1	0	0
CBK	1	0	0	1	0	0	0		0	0	0	0	0	0	0
CON	1	0	0	0	0	0	0	0		0	0	0	0	0	0
DAC	0	0	0	0	0	0	0	0	0		0	0	0	0	0
DBK	0	0	0	0	0	0	0	0	0	0		0	0	0	0
DBO	0	0	0	0	0	0	0	0	1	0	1		0	0	0
DLH	0	0	0	0	0	0	0	0	0	0	0	0		0	1
DPO	0	0	0	0	0	0	0	0	0	0	0	0	0		0
DTL	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
EON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRE	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
HEN	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
INF	1	0	0	1	0	0	1	0	0	0	1	1	0	0	0
LIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAN	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
MET	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MRV	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
RWE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAP	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
SCH	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
SIE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TYK	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TUI	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
VW	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1

Table 25: LMWm test results on stochastic dominance (1st order) of the daily returns of DAX stocks for the 1-year period (1 $\sim$ rejection of dominance).

	EON	FRE	HEN	INF	LIN	MAN	MET	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ADS	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
ALL	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ALT	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
BAS	0	0	0	1	0	1	0	0	0	0	0	1	0	1	0
BAY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BHV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BMW	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
CBK	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0
CON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DAC	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0
DBK	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
DBO	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
DLH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DPO	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
DTL	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
EON		0	0	1	0	0	0	0	0	0	0	0	0	1	0
FRE	0		0	1	0	0	0	0	0	0	0	0	0	0	0
HEN	0	0		1	0	0	1	0	0	0	0	0	0	0	0
INF	1	0	1		1	0	0	1	1	1	1	0	0	0	0
LIN	0	0	0	1		0	0	0	0	0	0	0	1	0	0
MAN	1	0	0	0	1		0	0	0	0	0	0	0	0	0
MET	0	0	0	0	0	0		0	0	0	0	0	0	0	0
MRV	1	0	1	1	0	0	0		0	0	0	0	0	0	0
RWE	0	0	0	1	0	0	0	0		0	0	0	0	0	0
SAP	1	0	0	0	0	0	0	0	0		0	0	0	0	0
SCH	0	0	0	0	0	0	0	0	0	0		0	0	0	0
SIE	0	0	0	1	0	0	0	0	0	0	0		0	1	0
TYK	1	0	0	0	0	0	0	0	0	0	0	0		0	0
TUI	0	0	1	0	0	0	0	0	0	0	0	0	0		0
VW	1	0	0	1	1	0	0	0	0	0	1	0	0	1	

Table 26: LMWm results on stochastic dominance (1st order) between the daily returns of DAX stocks for the 1-year period (1 ∼ rejection of dominance).

	ADS	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DAC	DBK	DBO	DLH	DPO	DTL
ADS		0	0	1	0	0	0	0	1	0	0	0	0	0	0
ALL	1		1	1	1	1	1	1	1	1	1	1	1	1	1
ALT	1	1		1	1	1	1	1	1	1	1	1	0	1	1
BAS	1	0	0		0	0	0	0	1	0	0	0	0	1	1
BAY	1	1	1	1		1	1	1	1	1	1	1	1	1	1
BHV	1	1	1	1	1		1	1	1	1	1	1	1	1	1
BMW	1	1	1	1	1	1		1	1	1	1	1	1	1	1
CBK	1	1	1	1	1	1	1		1	1	1	1	1	1	1
CON	1	1	1	1	1	0	1	0		1	1	1	1	0	1
DAC	1	0	1	1	1	1	1	1	1		1	1	0	1	1
DBK	1	1	1	1	1	1	1	1	1	1		1	0	1	1
DBO	1	1	1	1	1	1	1	1	1	1	1		1	1	1
DLH	1	1	1	1	1	1	1	1	1	1	1	1		1	1
DPO	1	1	1	1	1	0	1	1	1	1	1	1	1		1
DTL	1	0	0	1	1	1	0	1	1	1	1	0	0	1	
EON	1	0	0	1	0	0	0	0	1	1	1	0	0	0	1
FRE	1	0	1	1	0	0	0	0	1	1	1	0	0	0	1
HEN	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
INF	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LIN	1	0	1	1	0	0	0	0	1	1	1	0	0	0	1
MAN	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
MET	1	0	1	1	0	0	0	0	1	1	1	1	1	0	1
MRV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RWE	1	0	1	1	0	0	0	0	1	1	1	0	0	0	1
SAP	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
SCH	1	0	1	1	0	0	0	0	1	1	1	0	0	0	1
SIE	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
TYK	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
TUI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
VW	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 27: Descriptive results for stochastic dominance (2nd order) of the daily returns of DAX stocks for the 1-year period ($1 \sim$ rejection of dominance).

	EON	FRE	HEN	INF	LIN	MAN	MET	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ADS	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
ALL	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1
ALT	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
BAS	1	1	1	0	1	1	1	0	1	1	1	0	1	0	0
BAY	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1
BHV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BMW	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1
CBK	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1
CON	1	1	1	0	1	0	1	0	1	1	1	0	1	0	1
DAC	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0
DBK	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0
DBO	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DLH	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
DPO	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
DTL	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0
EON		1	0	0	0	1	0	0	1	0	0	0	0	0	0
FRE	1		1	0	1	1	1	0	1	0	1	0	1	0	0
HEN	1	1		0	1	1	1	0	1	1	1	1	1	0	1
INF	1	1	1		1	1	1	1	1	1	1	1	1	1	1
LIN	1	1	1	0		1	1	0	1	0	1	0	1	0	0
MAN	1	1	1	1	1		1	1	1	1	1	1	1	1	1
MET	1	1	1	0	1	1		0	1	1	1	1	1	0	1
MRV	1	1	1	1	1	1	1		1	1	1	1	1	1	1
RWE	1	1	1	0	1	0	0	0		0	1	0	0	0	0
SAP	1	1	1	0	1	1	1	1	1		1	1	1	0	1
SCH	1	1	0	0	1	1	0	0	1	0		0	0	0	0
SIE	1	1	1	0	1	1	1	1	1	1	1		1	1	1
TYK	1	1	1	0	1	1	1	1	1	1	1	1		1	1
TUI	1	1	1	0	1	1	1	1	1	1	1	1	1		1
VW	1	1	1	0	1	1	1	1	1	1	1	1	1	1	

Table 28: Descriptive results for stochastic dominance (2nd order) of the daily returns of DAX stocks for the 1-year period (1 $\sim$ rejection of dominance).

	ADS	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DAC	DBK	DBO	DLH	DPO	DTL
ADS		0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALL	1		0	0	0	0	0	0	1	0	0	0	0	0	1
ALT	1	0		0	0	0	0	0	1	0	0	0	0	0	0
BAS	1	0	0		0	0	0	0	1	0	0	0	0	0	0
BAY	1	0	0	1		0	0	0	1	0	0	0	0	0	1
BHV	1	1	0	1	0		0	0	1	0	0	0	0	0	1
BMW	1	1	0	1	0	0		0	1	0	0	0	0	0	1
CBK	1	0	0	1	0	0	0		1	0	0	0	0	0	1
CON	1	0	0	0	0	0	0	0		0	0	0	0	0	0
DAC	1	0	0	0	0	0	0	0	1		0	0	0	0	0
DBK	1	0	0	0	0	0	0	0	1	0		0	0	0	0
DBO	1	0	0	0	0	0	0	0	1	0	0		0	0	0
DLH	1	0	0	1	0	0	0	0	1	0	1	1		0	1
DPO	1	0	0	1	0	0	0	0	1	0	0	0	0		1
DTL	1	0	0	0	0	0	0	0	1	0	0	0	0	0	
EON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRE	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
HEN	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
INF	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
LIN	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
MAN	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MET	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0
MRV	1	0	0	1	0	0	0	0	1	0	0	0	0	1	1
RWE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAP	1	0	0	1	0	0	0	0	1	0	0	1	0	0	1
SCH	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
SIE	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1
TYK	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1
TUI	1	0	0	1	1	0	1	0	1	1	1	1	0	1	1
VW	1	0	0	1	1	0	1	1	1	0	0	0	0	1	1

Table 29: STm results on stochastic dominance (2nd order) between the daily returns of DAX stocks for the 1-year period (1 $\sim$ rejection of dominance).

	EON	FRE	HEN	INF	LIN	MAN	MET	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ADS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALL	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0
ALT	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0
BAS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BAY	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0
BHV	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0
BMW	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0
CBK	1	1	1	0	1	0	0	0	0	0	1	0	0	0	0
CON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DAC	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0
DBK	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0
DBO	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0
DLH	1	0	1	0	1	1	0	0	1	0	0	1	0	0	0
DPO	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0
DTL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
EON		0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRE	0		0	0	0	0	0	0	0	0	0	0	0	0	0
HEN	0	0		0	0	1	0	0	1	0	0	0	0	0	0
INF	1	1	1		1	1	1	1	1	1	1	1	1	0	1
LIN	0	0	0	0		0	0	0	0	0	0	0	0	0	0
MAN	1	0	1	0	1		0	0	1	0	0	0	0	0	0
MET	1	1	0	0	0	0		0	0	0	0	0	0	0	0
MRV	1	1	1	0	1	1	1		1	0	1	0	0	0	0
RWE	0	0	0	0	0	0	0	0		0	0	0	0	0	0
SAP	1	1	1	0	1	0	1	0	1		1	0	0	0	0
SCH	0	0	0	0	0	0	0	0	0	0		0	0	0	0
SIE	0	0	0	0	0	0	0	0	0	0	0		0	0	0
TYK	0	0	0	0	1	0	0	0	0	0	0	0		0	0
TUI	1	1	1	0	1	1	1	0	1	0	1	1	0		0
VW	1	1	1	0	1	1	1	0	1	1	1	1	1	0	

Table 30: STm results on stochastic dominance (2nd order) between the daily returns of DAX stocks for the 1-year period (1 $\sim$ rejection of dominance).

Name of Stock		Mean ×100	Std.dev. ×100	GARCH parameters α_1	β_1
Allianz	ALL	−2.093	36.778	0.1037	0.8921
Altana	ALT	16.065	36.972	0.0561	0.9439
BASF	BAS	10.441	28.955	0.0991	0.8741
Bayer	BAY	1.639	34.344	0.0713	0.9193
Bay. Hypo-Vereinsbank	BHV	−1.412	41.449	0.1290	0.8710
BMW	BMW	10.113	35.729	0.0927	0.9073
Commerzbank	CBK	−0.530	35.098	0.1469	0.8530
Continental	CON	13.114	32.427	0.0828	0.8806
Deutsche Bank	DBK	5.321	33.996	0.0909	0.9070
Deutsche Lufthansa	DLH	0.378	36.791	0.0655	0.9180
Eon	EON	7.566	29.132	0.0793	0.9118
Henkel	HEN	7.188	29.902	0.0681	0.9188
Linde	LIN	1.009	29.750	0.0563	0.9390
MAN	MAN	2.926	34.809	0.0538	0.9381
Muenchner Rueckvers.	MRV	2.623	38.763	0.1013	0.8944
RWE	RWE	5.120	29.702	0.0694	0.9249
SAP	SAP	21.992	51.177	0.1603	0.7917
Schering	SCH	11.324	29.480	0.0700	0.9092
Siemens	SIE	9.685	36.505	0.0554	0.9435
Thyssen-Krupp	TYK	0.303	34.034	0.0937	0.9001
TUI	TUI	−4.111	35.894	0.0935	0.9044
Volkswagen	VW	3.354	34.490	0.0731	0.9180

Table 31: Descriptive statistics and estimated GARCH parameters of the *annualized* daily returns of DAX stocks for the 10-year period.

	A L L	A L T	B A S	B A Y	B H V	B M W	C B K	C O N	D B K	D L H	E O N
ALL		23	49	48	58	45	58	37	61	46	39
ALT	23		23	21	19	19	21	17	23	17	21
BAS	49	23		68	44	50	46	37	51	48	41
BAY	48	21	68		41	45	44	34	48	41	37
BHV	58	19	44	41		40	62	32	59	40	32
BMW	45	19	50	45	40		45	40	48	44	36
CBK	58	21	46	44	62	45		36	64	45	34
CON	37	17	37	34	32	40	36		37	36	29
DBK	61	23	51	48	59	48	64	37		47	39
DLH	46	17	48	41	40	44	45	36	47		33
EON	39	21	41	37	32	36	34	29	39	33	
HEN	31	18	39	37	29	37	30	31	32	32	31
LIN	39	19	45	43	35	41	39	32	40	37	31
MAN	43	17	44	42	40	41	43	35	44	41	30
MRV	74	25	47	45	53	44	53	35	55	41	39
RWE	41	24	40	40	35	37	34	26	39	31	58
SAP	37	19	34	32	32	31	38	25	42	34	23
SCH	35	24	38	38	31	35	33	25	35	29	33
SIE	52	23	47	44	44	44	48	36	56	43	34
TYK	42	17	46	45	37	44	42	33	43	39	32
TUI	45	15	43	38	38	41	43	32	45	44	32
VW	48	20	54	48	44	61	48	44	52	48	40

Table 32: Correlation coefficients of the daily returns of DAX stocks $\times$ 100 for the 10-year period.

	HEN	LIN	MAN	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ALL	31	39	43	74	41	37	35	52	42	45	48
ALT	18	19	17	25	24	19	24	23	17	15	20
BAS	39	45	44	47	40	34	38	47	46	43	54
BAY	37	43	42	45	40	32	38	44	45	38	48
BHV	29	35	40	53	35	32	31	44	37	38	44
BMW	37	41	41	44	37	31	35	44	44	41	61
CBK	30	39	43	53	34	38	33	48	42	43	48
CON	31	32	35	35	26	25	25	36	33	32	44
DBK	32	40	44	55	39	42	35	56	43	45	52
DLH	32	37	41	41	31	34	29	43	39	44	48
EON	31	31	30	39	58	23	33	34	32	32	40
HEN		35	33	32	32	19	28	28	30	27	36
LIN	35		44	38	34	28	30	36	43	39	43
MAN	33	44		40	32	32	28	44	45	40	45
MRV	32	38	40		42	32	36	46	39	41	47
RWE	32	34	32	42		23	34	35	32	33	37
SAP	19	28	32	32	23		25	50	29	30	38
SCH	28	30	28	36	34	25		31	28	31	33
SIE	28	36	44	46	35	50	31		41	42	49
TYK	30	43	45	39	32	29	28	41		43	45
TUI	27	39	40	41	33	30	31	42	43		44
VW	36	43	45	47	37	38	33	49	45	44	

Table 33: Correlation coefficients of the daily returns of DAX stocks $\times$ 100 for the 10-year period.

	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DBK	DLH	EON	HEN	LIN	MAN	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ALL		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
ALT	0		0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
BAS	0	0		0	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0
BAY	0	0	0		0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0
BHV	0	0	0	0		0	1	0	1	0	1	0	0	0	0	0	1	1	0	0	0	0
BMW	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
CBK	1	0	0	0	1	1		0	0	1	0	0	0	1	1	0	1	0	0	0	0	0
CON	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1	0	0	0	0	0
DBK	0	0	0	0	0	0	0	0		0	1	0	0	0	0	0	1	0	0	0	0	0
DLH	1	1	1	1	0	0	1	1	1		1	1	1	0	0	1	1	1	0	0	0	0
EON	0	0	0	0	1	1	0	0	0	1		0	0	0	1	0	1	0	1	0	0	1
HEN	0	0	0	0	0	0	0	0	0	1	0		0	0	0	0	1	0	0	0	0	0
LIN	1	0	0	0	1	1	0	0	1	1	0	0		1	1	0	1	0	0	1	0	1
MAN	0	0	1	0	0	0	0	0	0	0	1	0	0		0	0	1	1	0	0	0	0
MRV	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0	0	0
RWE	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1		1	0	0	0	0	1
SAP	0	1	1	1	0	0	1	1	1	0	1	1	1	1	0	1		1	0	1	0	1
SCH	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1		0	0	0	1
SIE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0
TYK	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0		0	0
TUI	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	1	1	1	1	0		0
VW	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	

Table 34: LMWm results on stochastic dominance (1st order) between the daily returns of DAX stocks for the 10-year period (1 $\sim$ rejection of dominance).

	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DBK	DLH	EON	HEN	LIN	MAN	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ALL		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ALT	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BAS	0	1		0	0	0	0	1	0	0	1	1	0	0	0	1	1	1	0	0	0	0
BAY	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BHV	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BMW	1	1	1	1	0		1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
CBK	1	1	1	1	0	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CON	0	1	1	0	0	0	0		1	0	1	1	1	1	0	1	1	1	1	0	0	1
DBK	0	1	1	1	0	1	1	1		0	1	1	1	1	0	1	1	1	1	1	0	1
DLH	1	1	1	1	0	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1
EON	0	1	1	0	0	1	0	1	0	0		1	0	0	0	1	1	1	1	0	0	0
HEN	0	1	1	0	0	1	1	1	1	0	1		1	1	0	1	1	1	1	0	0	0
LIN	0	1	1	1	0	1	0	1	1	0	1	1		1	1	1	1	1	1	0	0	1
MAN	0	1	1	1	0	1	1	1	1	0	1	1	1		0	1	1	1	1	1	0	1
MRV	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1
RWE	0	1	1	0	0	1	0	1	1	0	1	1	1	0	0		1	1	1	0	0	0
SAP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1
SCH	0	1	1	0	0	0	0	1	0	0	1	1	1	1	0	1	1		1	0	0	0
SIE	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1		1	1	1
TYK	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1		0	1
TUI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1
VW	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	

Table 35: Descriptive results for stochastic dominance (2nd order) of the daily returns of DAX stocks for the 10-year period (1 $\sim$ rejection of dominance).

	ALL	ALT	BAS	BAY	BHV	BMW	CBK	CON	DBK	DLH	EON	HEN	LIN	MAN	MRV	RWE	SAP	SCH	SIE	TYK	TUI	VW
ALL		0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
ALT	0		1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	0	0
BAS	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BAY	0	0	1		0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
BHV	1	0	1	1		0	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0
BMW	0	0	1	0	0		0	0	0	0	1	1	1	0	0	1	0	1	0	0	0	0
CBK	0	0	1	0	0	0		0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
CON	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0
DBK	0	0	0	0	0	0	0	0		0	1	0	0	0	0	1	0	0	0	0	0	0
DLH	0	0	1	1	0	0	1	1	1		1	1	1	0	0	1	0	1	0	1	0	1
EON	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0
HEN	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
LIN	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
MAN	0	0	1	0	0	0	0	0	0	0	1	0	1		0	1	0	1	0	0	0	0
MRV	0	0	1	1	0	0	0	0	1	0	1	0	1	0		1	0	1	0	0	0	0
RWE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0
SAP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1
SCH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0
SIE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0
TYK	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	0		0	0
TUI	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0		0
VW	0	0	1	0	0	0	0	1	0	0	1	0	1	0	0	1	0	1	0	0	0	

Table 36: STm results on stochastic dominance (2nd order) of the daily returns of DAX stocks for the 10-year period (1 $\sim$ rejection of dominance).

Name of stock	1-year period LMWm SD1	1-year period STm SD2	10-year period LMWm SD1	10-year period STm SD2
Adidas-Salomon	•	•	–	–
Allianz				
Altana	•		•	•
BASF			•	•
Bayer				
Bay. Hypo-Vereinsbank				
BMW				
Commerzbank				
Continental	•	•	•	•
Daimler-Chrysler			–	–
Deutsche Bank	•			
Deutsche Boerse			–	–
Deutsche Lufthansa				
Deutsche Post	•		–	–
Deutsche Telekom			–	–
Eon	•	•		•
Fresenius			–	–
Henkel			•	•
Infineon			–	–
Linde				•
MAN				
Metro	•		–	–
Muenchner Rueckvers.				
RWE	•	•		•
SAP			•	•
Schering			•	•
Siemens	•		•	•
Thyssen-Krupp				
TUI				
Volkswagen				

Table 37: Efficiency results concerning the application of the considered tests on stochastic dominance to the daily returns of DAX stocks for the 1-year and the 10-year period. The efficient stocks are denoted with a bullet. The stocks not considered for the longer period are denoted with a hyphen.

8 Conclusion

We investigate the tests on stochastic dominance of Schmid/ Trede, Xu/ Fisher/ Willson and Linton/ Maasoumi/ Whang using simulation study. They determine the critical values of the tests based on subsampling methods. Schmid/ Trede use the permutation principle, Xu/ Fisher/ Willson the moving block bootstrap and Linton/ Maasoumi/ Whang subsampling estimation. Simulations show that all these tests perform rather poorly for finite samples if the data are generated by GARCH(1,1) processes, which is an appropriate choice for financial data.

We develop several modifications to heal these shortcomings. Our modification consists in using circular block methods: the modified ST test uses circular block permutation, the modified XFW test circular block bootstrap and the modified LMW test circular subsampling.

We show analytically that the modifications of the XFW and the LMW tests are asymptotically equivalent to the original tests. We argue that the circular block methods are more appropriate for resampling the dependence structure. By Monte Carlo simulation we show that for finite samples the modifications make the ST and the LMW tests robust to conditional heteroskedasticity. In contrast, the modifications do not improve the XFW test.

For the modified ST and LMW tests the appropriate block length has to be chosen. We find that for both tests the optimal block length grows approximately with rate $\sqrt{n}$ where n is the sample size. For the LMWm test the range of suitable block lengths is broader for larger sample sizes whereas this is is not true for the STm test.

Then we investigate the power of the developed tests. For various combinations of normally distributed observations we see that the power increases with increasing sample size and with increasing distance to the null hypothesis.

Finally we apply the tests to the daily returns of German stocks which feature typical time series properties. From the test results we determine the SD efficient sets of stocks.

In the power investigation we consider only a small part of the complex alternative of non-dominance. In further research one could explore the power if the data are generated by conditionally heteroskedastic time series. Another point of interest is how the modified tests perform when we assume some more sophisticated models, e.g. stochastic volatility or non-Gaussian copulas.

Further, one could search for a test with non-dominance in the null hypothesis and dominance in the alternative. In researching this topic similar complexity problems as in the power investigation arise.

Another interesting topic of further research could be investigation of dominance between various portfolios. For this goal, the tests developed in this paper can be applied as for single stocks.

References

Akgiray, V. (1989): Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts. Journal of Business 62, 55–80.

Davis, R.A., Mikosch, T., Basrak, B. (1999): Sample ACF of Multivariate Stochastic Recurrence Equations With Application to GARCH. Technical Report, University of Groningen.

Efron, B. (1979): Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics 7(1), 1–16.

Engle, R., Kroner, K. (1995): Multivariate Simultaneous Generalized ARCH. Econometric Theory 11, 122–150.

Härdle, W., Horowitz, J., Kreiss, J.-P. (2003): Bootstrap Methods for Time Series. International Statistical Review 71(2), 435–459.

Hall, P., Heyde, C.C. (1980): Martingale Limit Theory and its Applications. Academic Press, New York.

Künsch, H.R. (1989): The Jackknife and the Bootstrap for General Stationary Observations. Annals of Statistics 17(3), 1217–1241.

Lahiri, S.N. (1999): Theoretical Comparisons of Block Bootstrap Methods. The Annals of Statistics 27(1), 386–404.

Levy, H. (1992): Stochastic Dominance and Expected Utility: Survey and Analysis. Management Science 38, 555–593.

Lifshits, M.A. (1982): On the Absolute Continuity of Distributions of Functionals of Random Processes. Theory of Probability and Its Applications 27, 600–607.

Linton, O., Maasoumi, E., Whang, Y.-J. (2003): Consistent Testing for Stochastic Dominance: A Subsampling Approach. Discussion Paper, London / Dallas / Seoul.

Liu, R.Y., Singh, K. (1992): Moving Blocks Jackknife and Bootstrap Capture Weak Dependence. In: LePage, R., Billard, L.: Exploring the Limits of Bootstrap. Wiley & Sons, New York, 225–248.

Nelson, R.D., Pope, R.D. (1991): Bootstrapped Insights into Empirical Applications of Stochastic Dominance. Management Science 37(9), 1182–1194.

Politis, D.N., Romano, J.P. (1992): A Circular Block-Resampling Procedure for Stationary Data. In: LePage, R., Billard, L.: Exploring the Limits of Bootstrap. Wiley & Sons, New York, 263–270.

Politis, D.N., Romano, J.P. (1994): Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions. The Annals of Statistics 22(4), 2031–2050.

Politis, D.N., Romano, J.P., Wolf, M. (1999): Subsampling. Springer, New York.

Schmid, F., Trede, M. (1997): Nonparametric Inference for Second Order Stochastic Dominance from Paired Observations: Theory and Empirical Application. Wirtschafts- und Sozialstatistik heute, Theorie und Praxis; Festschrift für Walter Krug.

Stein, W., Pfaffenberger, R., Kumar, P.C. (1983): On the Estimation Risk in First-Order Stochastic Dominance: A Note. Journal of Financial and Quantitative Analysis 18(4), 471–476.

Whitmore, G., Findlay, M.C. (1978): Stochastic Dominance: An Approach to Decision-Making Under Risk. Heath, Lexington.

Xu, K., Fisher, G., Willson, D. (1997): Tests for Stochastic Dominance Based on α-Mixing with an Application in Finance. Working Paper, Dalhousie University.

Appendix

Proof of theorem 1. We begin with proving part 1. Let $\hat{G}_{n,b}$ be the empirical distribution function of $\{\sqrt{b}d_{n,b,k} : k = 1, \ldots, n\}$ and G the distribution function of the asymptotic distribution of $T_{n,1}$ under H_0^1. As Linton/ Maasoumi/ Whang state, G is absolutely continuous according to Lifshits (1982), Theorem 1. Therefore, to prove part 1, it suffices to show

$$\hat{G}_{n,b}(w) \xrightarrow{p} G(w) \ \forall w \in \mathbb{R}.$$

By definition of G

$$G_b(w) := P(\sqrt{b}d_{n,b,1} \leq w) \underset{b\to\infty}{\longrightarrow} G(w)$$

holds for all $w \in \mathbb{R}$. Hence we have to show $\hat{G}_{n,b}(w) \underset{n\to\infty}{\longrightarrow} G_b(w)$ for all $w \in \mathbb{R}$; note that $b \to \infty$ as $n \to \infty$. Let $I_k = 1(\sqrt{b}d_{n,b,k} \leq w)$ for $k = 1, \ldots, n$. $E(I_k) = P(\sqrt{b}d_{n,b,k} \leq w) = P(\sqrt{b}d_{n,b,1} \leq w) = G_b(w)$ holds for $k = 1, \ldots, n-b+1$. This yields

$$E(\hat{G}_{n,b}(w)) = \frac{1}{n}\sum_{k=1}^{n} E(I_k) = \frac{n-b+1}{n}G_b(w) + \frac{1}{n}\sum_{k=n-b+2}^{n} E(I_k)$$

and therefore

$$|E(\hat{G}_{n,b}(w)) - G_b(w)| = \Big|\frac{1}{n}\sum_{k=n-b+2}^{n} E(I_k) - \frac{b-1}{n}G_b(w)\Big| \leq \frac{b-1}{n} \underset{n\to\infty}{\longrightarrow} 0. \qquad (1)$$

Further,

$$\begin{aligned} Var(\hat{G}_{n,b}(w)) &= \frac{1}{n^2}\Big(\sum_{k=1}^{n} Var(I_k) + 2\sum_{1\leq k<l\leq n} Cov(I_k, I_l)\Big) \\ &= \frac{1}{n^2}\Big(\sum_{k=1}^{n} Var(I_k) + 2\sum_{m=1}^{n-1}\sum_{k=1}^{n-m} Cov(I_k, I_{k+m})\Big) \\ &= S_{n,0} + 2\sum_{m=1}^{n-1} S_{n,m} \end{aligned}$$

where

$$S_{n,m} = \frac{1}{n^2}\sum_{k=1}^{n-m} Cov(I_k, I_{k+m}).$$

$|I_k| \leq 1$ for all k yields $|Cov(I_k, I_{k+m})| \leq 1$ and so $|S_{n,m}| \leq \frac{1}{n}$ for all m. Therefore

$$\Big|S_{n,0} + 2\sum_{m=1}^{b-1} S_{n,m} + 2\sum_{m=n-b+1}^{n-1} S_{n,m}\Big| \leq O\Big(\frac{b}{n}\Big) = o(1).$$

Further, we have

$$\begin{aligned} \Big|2\sum_{m=b}^{n-b} S_{n,m}\Big| &= \frac{2}{n^2}\Big|\sum_{m=b}^{\lfloor\frac{n}{2}\rfloor}\sum_{k=1}^{n-m} Cov(I_k, I_{k+m}) + \sum_{m=\lfloor\frac{n}{2}\rfloor+1}^{n-b}\sum_{k=1}^{n-m} Cov(I_k, I_{k+m})\Big| \\ &\leq \frac{8}{n^2}\Big|\sum_{m=b}^{\lfloor\frac{n}{2}\rfloor}\sum_{k=1}^{n-m} \alpha(m-b+1) + \sum_{m=\lfloor\frac{n}{2}\rfloor+1}^{n-b}\sum_{k=1}^{n-m} \alpha(n-m-b+1)\Big| \qquad (2) \\ &\leq \frac{8}{n}\Big|\sum_{m=1}^{\lfloor\frac{n}{2}\rfloor-b+1} \alpha(m) + \sum_{m=1}^{\lceil\frac{n}{2}\rceil-b} \alpha(m)\Big| \\ &\leq O(n^{-1}) = o(1) \qquad (3) \end{aligned}$$

where (2) holds by Hall/Heyde (1980), Theorem A.5, and (3) holds by the assumption that $\alpha(m) = O(m^{-3})$. Hence we have shown

$$\lim_{n\to\infty} Var(\hat{G}_{n,b}(w)) = 0. \tag{4}$$

(1) and (4) yield $\hat{G}_{n,b}(w) \xrightarrow{p} G_b(w)$ and therefore $\hat{G}_{n,b}(w) \xrightarrow{p} G(w)$. This establishes part 1 of the theorem.

For the proof of part 2 first note that under H_1^1 we have

$$d := \sup_{x\in\mathbb{R}}(F_X(x) - F_Y(x)) > 0.$$

Analogously to Linton/ Maasoumi/ Whang $d_n(W_1, \ldots, W_n) \xrightarrow{p} d$ holds. Let $\hat{G}^0_{n,b}$ and $g^0_{n,b}$ be the empirical distribution and quantile function of $\{d_{n,b,k} : k = 1, \ldots, n\}$ and G^0_b the distribution function of $d_{n,b,1}$. Due to the mixing condition the convergence $\hat{G}^0_{n,b}(w) \xrightarrow{p} G^0_b(w)$ holds; it can be shown analogously to part 1. With $d_b(W_1, \ldots, W_b) \xrightarrow{p} d$ this yields $\hat{G}^0_{n,b}(w) \xrightarrow{d} \delta_d$ where δ_d denotes the Dirac distribution in d. Therefore we have $g^0_{n,b}(1-\alpha) \xrightarrow{p} d$. Because of $g_{n,b}(1-\alpha) = \sqrt{b}g^0_{n,b}(1-\alpha)$ this yields

$$P(T_{n,1} > g_{n,b}(1-\alpha)) \underset{n\to\infty}{\longrightarrow} 1$$

as in Linton/ Maasoumi/ Whang. □