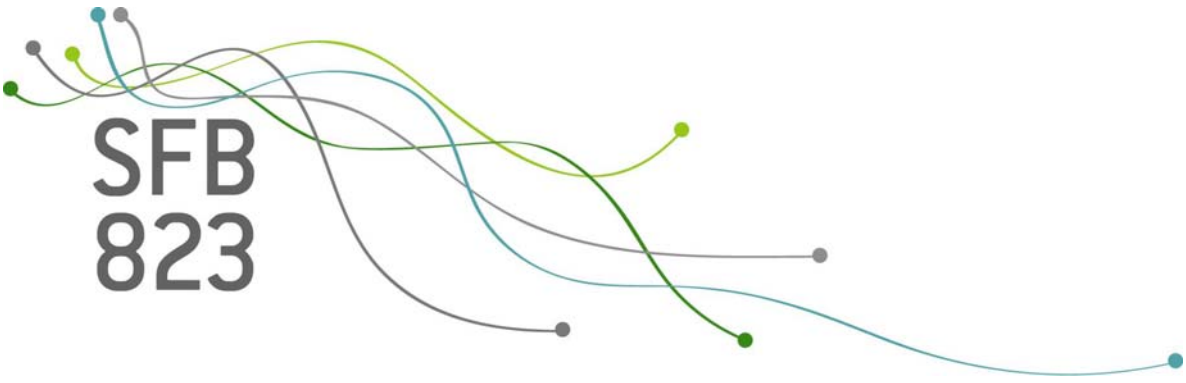
SFB 823

Fourier methods for analysing piecewise constant volatilities

Max Wornowizki, Roland Fried, Simos G. Meintanis

Nr. 43/2016

Discussion Paper



Fourier methods for analysing piecewise constant volatilities

Max Wornowizki[*], Roland Fried[*], and Simos G. Meintanis[**]

[*]Department of Statistics at TU Dortmund University, Dortmund, Germany, email: wornowizki@statistik.tu-dortmund.de
[**]Department of Economics at National and Kapodistrian University of Athens, Athens, Greece

August 9, 2016

Abstract

We develop procedures for testing the hypothesis that a parameter of a distribution is constant throughout a sequence of independent random variables. Our proposals are illustrated considering the variance and the kurtosis. Under the null hypothesis of constant variance, the modulus of a Fourier type transformation of the volatility process is identically equal to one. The approach proposed utilizes this property considering a canonical estimator for this modulus under the assumption of independent and piecewise identically distributed observations with zero mean. Using blockwise estimators we introduce several test statistics resulting from different weight functions which are all given by simple explicit formulae. The methods are compared to other tests for constant volatility in extensive Monte Carlo experiments. Our proposals offer comparatively good power particularly in the case of multiple structural breaks and allow adequate estimation of the positions of the structural breaks. An application to process control data is given, and it is shown how the methods can be adapted to test for constancy of other quantities like the kurtosis.

Keywords: Change point analysis, variance, piecewise identical distribution, independence, weight function

1 Introduction

Let us consider a real-valued sequence of independent random variables R_t corresponding to the times $t = 1, \ldots, n$ for some $n \in \mathbb{N}$. In addition, let $\theta(t)$ denote an unknown parameter of the distribution of R_t for each $t = 1, \ldots, n$. We are interested in testing whether $\theta(\cdot)$ is constant over time or not, as expressed in

the pair of hypotheses

$$\mathbb{H}_0 : \forall t, t' = 1, \ldots, n : \theta(t) = \theta(t') \text{ vs. } \mathbb{H}_1 : \exists t, t' = 1, \ldots, n : \theta(t) \neq \theta(t'). \quad (1)$$

The idea presented in the following allows us to tackle this testing problem for any parameter $\theta(\cdot)$ which can be estimated in a reasonable sense. When the parameter $\theta(\cdot)$ involves moments such procedures appear as 'tests for stability of moment condition(s)' and are popular mostly in the econometrics literature, particularly in the context of regression; see for example Juhl and Xiao (2013). For illustration purposes, in the major part of the following work we choose $\theta(t)$ as the volatility of the random variable R_t. That is, we test whether the random variables under study have the same variance as opposed to the alternative of one or several changes of the variance. After presenting and evaluating our approach for this particular testing problem, we discuss its extension to the general case and illustrate our proposal testing the constancy of the kurtosis.
Testing for changes in the volatility is a topic studied for the most part in the last forty years. First papers on this problem often rely on distributional assumptions in order to derive appropriate decision rules. For example for the Gaussian case Hsu (1977) constructs a test using cumulative sums of χ^2-type random variables, while Chen and Gupta (1997) as well as Jandhyala et al (2002) propose likelihood based procedures. To weaken the distributional assumptions later articles such as Wied et al (2012) consider asymptotic CUSUM-type tests, while Ross (2013) and others introduce nonparametric alternatives. It should be noted that most of these methods split the sample into two parts and are therefore constructed to detect at most one change at a time. This can be quite problematic in cases where several change points lead to comparable variances for any split of the data in two parts.

The method proposed in the following tries to circumvent this problem by using a test statistic calculated from multiple data blocks. It performs quite well in particular on data with several structural breaks, but is also competitive in case of only one change point. In addition, it does not make any assumptions on the distribution of the data and keeps the significance level for any sample size. In this paper we adopt a framework within which marginal distributional features are locally stationary, but globally non-constant. A blockwise constant volatility has drawn attention in Mercurio and Spokoiny (2004), Stărică and Granger (2005), Vassiliou and Demetriou (2005), Spokoiny (2009), Davies et al (2012) and Fried (2012), among others, but there are also more general frameworks such as that of blockwise local stationarity suggested in Guégan (2015).

Let us define the variance process $\sigma^2(\cdot)$ setting $\sigma^2(t) = \mathrm{Var}(R_t)$, where $\mathrm{Var}(R_t)$ is the variance of R_t for $t = 1, \ldots, n$. The assumption of a blockwise constant variance can be formulated using some specified time points $0 = t_0 < t_1 < \ldots < t_N = n$ which are understood as possible change point positions. If external knowledge allowing to choose them appropriately is not available, one can select $t_0, \ldots, t_N$ equidistantly. The possible change point positions correspond to important events which may trigger an upward or downward change in the volatility. The values of the volatility process $\sigma^2(\cdot)$ are thus allowed

to differ for some of the time blocks $B_j = \{t_{j-1}+1, t_{j-1}+2, \ldots, t_j\}$, $j = 1, ..., N$. Within each time block the volatility is assumed to be approximately constant.

In such a modelling framework, it is essential to obtain good estimates of the time points where the volatility, or more generally the parameter of interest, changes, if the null hypothesis of global constancy is rejected.

We work under the assumption of independent zero mean random variables, which are identically distributed up to scale. In other words, we assume that for some unknown but fixed distribution function F the relation $P(R_t \leq x) = F\left(x/\sqrt{\sigma^2(t)}\right)$ holds for all $t = 1, \ldots, n$ and all $x \in \mathbb{R}$. The centered Gaussian distribution is thus contained as a special case, but heavy tails as encountered in many financial applications are also included. The zero mean assumption is justifiable when dealing with returns or similar data obtained from differences of consecutive observations. In other cases it can be relaxed to blockwise constant means with known block structure, so that zero mean data results from preprocessing. For a detailed discussion of the other assumptions in the context of volatility see for example Spokoiny (2009) and the references given therein.

In the following we propose a class of tests for the testing problem (1) suiting the framework of a blockwise constant volatility process. The procedures are based on transformation of blockwise estimates to the complex plane.

The paper is structured as follows: in Section 2 a new type of statistics for testing changes in the volatility process is proposed. The corresponding tests as well as several other procedures for the testing problem under study are introduced in Section 3. In Section 4 all tests presented before are compared in multiple simulation scenarios. The best methods are applied to quality control data concerning face sampling in South African gold mines. In Section 5 we extend the basic construction to obtain tests for changes in kurtosis. The performance of the new kurtosis tests is investigated in a simulation study along with several competitors. Section 6 concludes and provides an outlook on possible future work.

2 Test statistics

In this section we introduce a class of test statistics for testing the constancy of the volatility process. We then derive explicit representations for some members of this class.

2.1 Motivation

To test $\mathbb{H}_0$ specified in (1) a reasonable first step is the estimation of the volatility in each time block B_j, $j = 1, \ldots, N$. Since all random variables are assumed to

have zero mean, a natural estimator for the j-th block is given by

$$\widehat{\sigma}_j^2 = \frac{1}{\tau_j} \sum_{t \in B_j} R_t^2, \ j = 1, \ldots, N. \tag{2}$$

Hereby $\tau_j = |B_j|$ denotes the number of observations with observation times in the j-th time block B_j, $j = 1, \ldots, N$.

Note that a sufficient number of observations in each block is necessary in order to ensure volatility estimations with reasonable accuracy.
Instead of considering the estimated volatilities themselves, we rather work with their logarithms. This allows us to construct scale independent test procedures as we demonstrate in the following. Under the null hypothesis specified in (1) the logarithmized volatility process $\log\left(\sigma^2(\cdot)\right)$ constantly equals $\log\left(\sigma^2\right)$ for some unknown $\sigma^2 > 0$. Thus under $\mathbb{H}_0$ the function $\varphi : \mathbb{R} \times \{1, \ldots, n\} \to \mathbb{C}$ defined by $\varphi(u,t) = e^{iu \log\left(\sigma^2(t)\right)}$, where $i = \sqrt{-1}$, does not depend on t. Hence, for any $t = 1, \ldots, n$ it can be estimated by

$$\widehat{\varphi}(u) = \sum_{j=1}^{N} \frac{\tau_j}{n} e^{iu \log\left(\widehat{\sigma}_j^2\right)}$$

in a straightforward way. Hereby, each $\tau_j/n = \tau_j/(\sum_{i=1}^{N} \tau_i)$ weights the corresponding term derived from the j-th block to the estimator according to the number of observations in the block for $j = 1, \ldots, N$. If all blocks contain the same number of observations, all weights are equal to $\frac{1}{N}$. This transformation of the blockwise estimators is closely related to the Fourier transform and characteristic functions. In this situation it has the following nice and intuitive behavior: under $\mathbb{H}_0$ the blockwise estimations are closer to each other than under the alternative. Because of that the function $f : \mathbb{R} \to \mathbb{C}, f(x) = e^{iux}$ maps them to points on the unit circle close to each other. Consequently their weighted mean $\widehat{\varphi}(u)$ lies relatively close to the unit circle for every $u \in \mathbb{R}$ and has a modulus near one. Under the alternative the logarithmized blockwise estimations differ more than under $\mathbb{H}_0$. They are thus mapped to distant points on the unit circle for most $u \in \mathbb{R}$. Hence, for most $u \in \mathbb{R}$ their weighted mean $\widehat{\varphi}(u)$ is closer to the origin than under the null hypothesis. In these cases $\widehat{\varphi}(u)$ has a small modulus. This is related to the well known property that the modulus of a given characteristic function takes on its maximum value one, identically in u, if and only if the underlying random variable is degenerate; see Loéve (1977), §14.1. In view of this fact we propose to use test statistics of the form

$$V = \int \left(1 - |\widehat{\varphi}(u)|^2\right) w(u) du$$

to test for global constant volatility, where here and henceforth integration is meant over the entire real line. Note that V is nonnegative because of $|\widehat{\varphi}(u)|^2 \leq 1$. Also, the weight function $w : \mathbb{R} \to \mathbb{R}_0^+$ should be chosen such that it ensures a finite test statistic. More details on the choice of w along with a more explicit representation of V are given in the next subsection.

We close this subsection by noting that the idea of utilizing the property of a Fourier transform taking on its maximum modulus constantly over u is, to the best of our knowledge, new in the literature. Nevertheless Fourier transforms and characteristic functions have been used before for change point detection, and to a good effect. See for instance Hušková and Meintanis (2006a,b), Hlávka et al (2012), and Hlávka et al (2015). Specifically in these papers it is shown that this type of transformation is convenient from the computational point of view, leads to theoretically sound asymptotics, and it is competitive compared to more classical approaches. The aforementioned Fourier–type tests are constructed by splitting the sample at each candidate change point and comparing the samples before and after the splitting position using two–sample techniques (Meintanis, 2005). This approach is based on the implicit assumption of only one or at least one dominant change point in the data. As opposed to this, the method suggested in our paper, while focusing on a particular, yet arbitrary, feature of the time series $\theta(\cdot)$, is designed to handle multiple structural changes by splitting the data set into several samples.

2.2 Calculation and weight functions

We now show how V can be represented explicitly. In order to handle the integral in our test statistic we set $W = \int w(u)du < \infty$ for an integrable weight function w. In this case V can be rewritten to

$$V = W - \int |\widehat{\varphi}(u)|^2 w(u)du. \tag{3}$$

Since W is independent of the data, it can be dropped. Using the definition of $\widehat{\varphi}$ the integral in (3) reduces to

$$T_{Four} = \frac{1}{n^2} \sum_{j,k=1}^{N} \tau_j \tau_k I_w \left(\log\left(\widehat{\sigma}_j^2\right) - \log\left(\widehat{\sigma}_k^2\right) \right), \tag{4}$$

where

$$I_w(x) = \int \cos(ux) w(u)du.$$

Keep in mind that since small values of V support the null hypothesis, $\mathbb{H}_0$ will be rejected for small values of T_{Four}. This test statistic depends on the data only via the terms $\log\left(\widehat{\sigma}_j^2\right) - \log\left(\widehat{\sigma}_k^2\right) = \log\left(\widehat{\sigma}_j^2/\widehat{\sigma}_k^2\right), 1 \le j < k \le N$. Therefore, thanks to taking the logarithm, any scale factor is canceled out. Hence T_{Four} is scale invariant. For that reason we propose to use the logarithmized estimated volatilities instead of the estimated volatilities themselves.
The function I_w can be expressed explicitly for several standard choices of w. These are the uniform, the Laplace and the Gaussian weighting with corresponding weight functions

$$w_U(u) = \mathbf{1}_{(-a,a)}(u), w_L(u) = e^{-a|u|} \text{ and } w_G(u) = e^{-au^2},$$

respectively. All of them depend on a parameter $a > 0$. Straightforward computations lead to

$$I_{w_U}(x) = \frac{2\sin(ax)}{x}, \quad I_{w_L}(x) = \frac{2a}{a^2 + x^2}, \quad \text{and} \quad I_{w_G}(x) = \sqrt{\frac{\pi}{a}} \exp\left(\frac{-x^2}{4a}\right)$$

for the uniform, Laplace and the Gaussian weight function, respectively (Hušková and Meintanis, 2006a). Hereby, $I_{w_U}(0)$ is defined by its limit $\lim_{x \to 0} \frac{2\sin(ax)}{x} = 2a$.

There also exist alternative choices for w. One example is the data adaptive weighting scheme proposed by Meintanis et al (2014) for goodness-of-fit testing. Another weight function is studied in Matteson and James (2014) in the context of multivariate nonparametric detection of general distributional changes. Both weight functions were considered in our simulations. The corresponding results are not included in this paper, since in accordance to our simulations in Section 4.1 the weight function does not have a large impact on the performance of our tests. In particular, the data adaptive weighting leads to slightly worse and the weighting proposed by Matteson and James to essentially the same rejection rates as the three standard weight functions introduced above. These results match earlier work utilizing such techniques showing that the performance of the procedures does not depend much on the particular functional form of the weight function chosen; see for instance Hušková and Meintanis (2006a) and Hlávka et al (2012).

In general there also might be some influence from the value of the parameter a. In this connection and for the weight functions considered herein, a large value of a lets w decay more sharply while the opposite is true for smaller values of the same parameter. In the former case (resp. latter case) more emphasis is placed on $\widehat{\varphi}(u)$ for u near the origin (resp. away from the origin). In turn recall that the behaviour of the characteristic function near the origin reflects the tail behaviour of the underlying distribution. Therefore caution must be exercised in choosing a. If a is chosen too large then we might overemphasize the tail of the underlying law, whereas in the opposite case of a being too small, these tails will be obscured; see for instance Epps (1993). Other than that the whole issue of making an educated guess for a is typically highly technical and it has been theoretically investigated only under very stringent parametric assumptions both about the hypothesis being tested as well as about the direction of possible deviation from this hypothesis; see Epps (1999), Tenreiro (2009). Fortunately, the simulations in Section 4.1 show that our testing procedures seem not that sensitive to the choice of a so that these problems are alleviated.

3 Testing for constant volatility

In this section we first show how the hypothesis of global constant volatility can be tested using the statistics defined in Section 2. Hereafter, a natural estimator of the structural break position in case of rejection is defined. The procedure allows for the location of multiple presumable structural break positions. The

section closes by briefly introducing four alternative methods for the testing problem taken from or inspired by the literature.

3.1 Testing procedure

The distribution of our test statistics strongly depends on the distribution of the random variables $R(1), \ldots, R(n)$. Getting critical values without imposing distributional assumptions is thus impossible at least for small sample sizes. In such situations resampling strategies are often of great help. Since under the null hypothesis in (1) the observations stem from identically distributed random variables, we propose tests for global constant volatility using the permutation principle introduced by Fisher (1935). Research in various fields shows that it can lead to quite powerful tests; see Good (2005) for a monograph treatment with applications to various fields and Hušková and Meintanis (2006b) in the context of characteristic functions. In our application the method works as follows: given the complete original sample first generate p new samples by randomly permuting the observations p times. Then determine the test statistic T_{Four} for each of the $p+1$ samples assuming that the data was observed in the given order. Thereby choose the same parameters N, w, a and block lengths $\tau_1, \ldots, \tau_N$ for all computations. Under the null hypothesis the $p+1$ test statistics stem from identically distributed random variables. Thus the permutation test rejects $\mathbb{H}_0$ at the predefined significance level α, if the test statistic determined on the original sample falls below the empirical α-quantile of all $p+1$ test statistics.

3.2 Localization of structural breaks

If one of our tests rejects $\mathbb{H}_0$, we are interested in localizing the first presumable change point. A rough approximation for that is given by t_{j^*}, where $j^* = \operatorname{argmax} |\log\left(\widehat{\sigma}_j^2\right) - \log\left(\widehat{\sigma}_{j+1}^2\right)|$ and the maximization is performed over the blocks $j = 1, \ldots, N-1$. However, the resolution of this estimator is limited by the block lengths. This is particularly problematic if potential change point positions $t_1, \ldots, t_N$ are not determined by a priori knowledge. In order to alleviate this problem the presumable change point position can be fine tuned as follows: since we expect a structural break near the rough estimate t_{j^*}, we consider the union of the two blocks around t_{j^*}, $B = \{t_{j^*-1}+1, t_{j^*-1}+1, \ldots, t_{j^*}, \ldots, t_{j^*+1}\}$. For the moment we focus solely on the observations with indices in B and exclude the ones from the remaining blocks from this part of the analysis. Our goal is to find the index $t^* \in B$ such that the empirical variance before t^*, $\widehat{\sigma}_1^2(t^*)$, differs the most from the empirical variance after t^*, $\widehat{\sigma}_2^2(t^*)$. Hereby both $\widehat{\sigma}_1^2(t^*)$ and $\widehat{\sigma}_2^2(t^*)$ are computed analogously to (2) using observations from B only. The position of the presumable structural break is thus estimated by

$$t^* = \underset{t \in B}{\operatorname{argmax}} |\log\left(\widehat{\sigma}_1^2(t)\right) - \log\left(\widehat{\sigma}_2^2(t)\right)|.$$

In order to ensure meaningful estimations $\sigma_1^2(t)$ and $\sigma_2^2(t)$ we do not maximize over all $t \in B$ but have to leave out values of t near the bounds of B.
Multiple structural break positions are located in a recursive manner in the spirit of Vostrikova (1981). After identifying the first presumable change point as described above, the sample is split into two parts at that point. The test procedure is then repeated on each of the subsamples large enough to ensure reasonable estimations. In case of new rejections the corresponding presumable change points are determined and the splitting continues. As soon as no splitting is performed anymore, the current data blocks seem homogeneous and the method stops. This testing procedure attains a predefined significance level α under the null hypothesis, since under $\mathbb{H}_0$ the permutation test conducted on the full sample rejects in only α percent of the cases.

3.3 Alternative methods

The literature offers several approaches for checking global constant volatility of a series of observations. The CUSUM procedure is a standard tool in the detection of structural breaks. Since a lot of work is available on it, we choose the method proposed by Wied et al (2012) as a representative for this class of tests. It is based on the CUSUM statistic

$$T_{CUS} = \max_{1 \leq t \leq n} \left| \widehat{D} \frac{t}{\sqrt{n}} (\widehat{\sigma}_{1:t}^2 - \widehat{\sigma}_{1:n}^2) \right|,$$

where $\widehat{\sigma}_{1:l}^2$ denotes the empirical variance of the first l observations for $l = 1, \ldots, n$. The normalising scalar $\widehat{D}$ is necessary to attain the asymptotic distribution. The CUSUM approach compares the discrepancies between the estimated variance on the whole sample to all estimated variances on proper subsamples. It then determines the maximal deviation signaling a possible structural break. The test is designed to detect at most one change in volatility and critical values are derived from asymptotics.
As opposed to the CUSUM strategy, Peña (2005) also compares variances estimated on subsamples to a measure of volatility estimated on the complete sample. He proposes a test statistic built in a blockwise manner. This statistic is

$$T_{Log} = n \log \left(\sum_{t=1}^{n} R_t^2 \right) - \sum_{j=1}^{N} \tau_j \log \left(\widehat{\sigma}_j^2 \right). \tag{5}$$

Its distribution under the null hypothesis also heavily depends on the data, as for our test statistics. In order to obtain distribution-free tests, we again apply the permutation principle for the construction of a test.
Another approach for testing the hypothesis of global constant volatility is given by Ross (2013). It is motivated by the classical distribution-free procedure proposed by Mood (1954). Instead of using $R(1), \ldots, R(n)$, their ranks in the complete sample denoted by $r(1), \ldots, r(n)$ are determined. Hereafter, the sample is divided into two subsamples for each possible split position $t = 1, \ldots, n$.

For each of these splittings the standardized test statistic of the Mood test is calculated. The expected value $\mu_t = t(n^2-1)/12$ and the standard deviation $\sigma_t = \sqrt{t(n-t)(n+1)(n^2-4)/180}$ used hereby hold under the null hypothesis. Taking the maximum over the possible split positions $t = 1, \ldots, n$ results in

$$T_{Mood} = \max_{t=1,\ldots,n} \frac{\left| \sum_{h=1}^{t} \left(r(h) - \frac{n+1}{2} \right)^2 - \mu_t \right|}{\sigma_t}.$$

Since only the ranks of the observations contribute to the test statistic, the procedure is distribution-free. Appropriate critical values depend only on the sample size n and can be derived by simulations. For several critical values and more details we refer to Ross (2013).
In addition we consider the monitoring procedure based on characteristic functions introduced by Steland and Rafajłowicz (2014). The authors develop a test statistic comparable to ours. It has the advantage that changes in the location process do not affect the monitoring of the volatility and vice versa. According to the authors

$$S_j = \int \left[\left(\widehat{U}_j(u) \right)^2 + \left(\widehat{V}_j(u) \right)^2 \right] w(u) du \text{ for } j = 1, \ldots, N$$

is an estimator in the context of characteristic functions, which reflects the volatility in the j-th block. Hereby, w is a weight function as before. $\widehat{U}_j$ and $\widehat{V}_j$ denote the empirical estimators of the real and imaginary part of the characteristic function for the random variables in the j-th block:

$$\widehat{U}_j(u) = \frac{1}{\tau_j} \sum_{t \in B_j} \cos\left(u \cdot R_t\right) \text{ and } \widehat{V}_j(u) = \frac{1}{\tau_j} \sum_{t \in B_j} \sin\left(u \cdot R_t\right), \; j = 1, \ldots, N.$$

We adopt the monitoring procedure to the retrospective case in the following way: since the null hypothesis should be rejected if the volatilities in two blocks are substantially different, we propose the quantity

$$T_{cf} = \max_{1 \le j < k \le N} |S_j - S_k|$$

as a test statistic for the testing problem under study. The testing is carried out via the permutation principle. Note that for any $j = 1, \ldots, N$ one can rewrite S_j as

$$S_j = \frac{1}{\tau_j^2} \sum_{t,t' \in B_j} I_w\left(R_t - R_t'\right),$$

so that by (4) T_{Four} can be interpreted as a weighted version of S_j computed on the pseudo observations $\log\left(\widehat{\sigma}_1^2\right), \ldots, \log\left(\widehat{\sigma}_N^2\right)$.
All four methods presented in this section reject the hypothesis of a global constant volatility for large values of the corresponding test statistic.

4 Evaluation of the methods

In this section we compare the performance of the permutation tests based on the statistics presented in Section 2 to the competitors listed in Section 3.3. This is achieved by determining the empirical powers of the tests in different data scenarios. We thereby address the choice of parameters and weighting functions. Hereafter, the best methods are applied to process control data.

4.1 Choice of settings

As a first step of the analysis we assess the influence of the weight function w, its parameter a and the number of the blocks on the two tests using weight functions. These are the ones based on the Fourier type statistics T_{Four} and the statistics T_{cf} derived from characteristic functions. Since both methods are constructed using the permutation principle, they attain a predefined significance level α under the null hypothesis of global constant volatility. Therefore, their empirical powers under alternatives are adequate performance measures. The tests are evaluated on datasets consisting of 200 observations. The first half of each sample is generated from the standard Gaussian distribution. The second 100 observations are sampled from the Gaussian distribution with increased standard deviation 1.5 and mean 0. In this manner we proceed generating 1000 datasets and apply both tests to them. Following the discussion in Section 2.2, appropriate values for the weight parameter $a > 0$ are chosen from the literature on empirical characteristic functions, which are comparable to our quantity $\widehat{\varphi}_N$; see for instance Jiménez-Gamero et al (2009), Potgieter and Genton (2013) and Pardo-Fernández et al (2015). In accordance with this prior experience we consider the values $a = 0.5,\ 1,\ 1.5$ for the uniform (w_U), Laplacian (w_L) and Gaussian (w_G) weight functions. The number of equidistant blocks is set to $N = 5$ and $N = 10$ and 2000 permutations are conducted for both tests. The corresponding rejection rates are given in Table 1.

Table 1: Rejection rates in case of one volatility change in the middle of the sample for the permutation tests based on T_{Four} and T_{cf} for different weight functions w, the parameter values $a = 0.5, 1, 1.5$ and two numbers of initial equidistant blocks N given in rounded percent

		w_U			w_L			w_G		
		0.5	1	1.5	0.5	1	1.5	0.5	1	1.5
N=5	T_{Four}	80	80	81	81	80	80	80	80	81
	T_{cf}	68	69	65	60	65	68	66	68	69
N=10	T_{Four}	72	72	72	71	72	71	72	72	71
	T_{cf}	47	50	42	34	42	46	43	48	49

According to the rejection rates the choice of the weight function w and its

parameter a do not have a large influence on the empirical powers of the test using T_{Four}. The second method seems more affected by them and in particular does not show the same behavior in a for each choice of w. Unsurprisingly, both tests heavily depend on the initial number of blocks, because a few large blocks in general allow better estimations of the blockwise volatilities. For that reason the methods lead to lower rejection rates for $N = 10$ in comparison to $N = 5$. This is the case despite of the fact that for $N = 5$ the true structural break lies in the middle of one block, which certainly has a negative effect. We also observe that the test based on T_{Four} considerably outperforms the one using T_{cf} for all parameter constellations.
In the following we restrict ourselves to the Gaussian weighting and $a = 1.5$ for both tests due to the almost identical results for T_{Four} and good results for T_{cf} in the previous analysis. Repeating the simulation for $N = 2, \ldots, 10$ yields the rejection rates in Table 2:

Table 2: Rejection rates in case of one volatility change in the middle of the sample for the permutation tests based on T_{Four} and T_{cf} for different numbers of initial equidistant blocks N given in rounded percent

N	2	3	4	5	6	7	8	9	10
T_{Four}	96	82	93	79	85	71	79	66	74
T_{cf}	98	77	86	67	68	56	62	48	50

As expected choosing $N = 2$ blocks leads to the highest rejection rates, since this data splitting perfectly fits the true volatility clusters. For the same reason even numbers of blocks N attain higher results than the corresponding odd neighbors $N - 1$ and $N + 1$, for which the structural change is positioned withing a block. Also, quite unsurprisingly, the rejection rates decrease with N in general. This is the case, because a higher number of blocks leads to worse estimations of the individual volatilities in each block, since less data is available. In addition for large N some blocks, say blocks j and k, estimate the same volatility. In this case the difference $\log(\hat{\sigma}_k^2) - \log(\hat{\sigma}_j^2)$ figuring in the test statistic T_{Four} is not equal to 0 solely due to sampling error. Thus more terms than necessary contribute to the mean T_{Four}. Therefore, the impact of the relevant terms which reflect an actual change of the volatility is downweighted. Lastly, we observe that, up to a certain degree, the correct position of the blocks is more relevant than their actual number. For example even though $N = 2$ is optimal in this case, $N = 6$ is preferable to $N = 3$ and $N = 5$ for the test based on T_{Four}.
To study the latter point in more detail, we repeat the simulation with the structural change now taking place after 20% of the observations rather than in the middle of the sample:

According to Table 3 choosing few oversized blocks is more problematic than choosing too many small blocks. In this particular case $N = 5$ blocks free of

Table 3: Rejection rates in case of one volatility change after 20% of the observations for the permutation tests based on T_{Four} and T_{cf} for different numbers of initial equidistant blocks N given in rounded percent

N	2	3	4	5	6	7	8	9	10
T_{Four}	25	34	48	69	53	44	42	49	50
T_{cf}	26	35	47	70	48	37	36	32	36

structural breaks attain the best results, although actually just two volatility regions exist. For $N = 2$, the correct number of blocks, the data is split in the wrong position leading to worse rejection rates. As before we observe a general decrease of power after the optimal choice $N = 5$ and relatively good results for its multiple $N = 10$. All results mentioned consistently apply for both tests. Hereby T_{Four} always leads to higher rejection rates than T_{cf}.

4.2 Method comparison

We now apply all tests introduced in Section 3 in five data scenarios listed in the following. To present the settings in a clear and compact way, let $|n, \sigma = a|\, m, \sigma = b|$ denote n observations with standard deviation a followed by m observations with standard deviation b. The data cases under study are:

1) $|200, \sigma = 1|$ ($\mathbb{H}_0$)

2) $|100, \sigma = 1|\, 100, \sigma = 1.5|$ (one structural break)

3) $|100, \sigma = 1|\, 100, \sigma = 1.5\, |100, \sigma = 1|$ (two structural breaks)

4) $|100, \sigma = 1|\, 100, \sigma = 1.4\, |100, \sigma = 1|\, 100, \sigma = 1.4\, |100, \sigma = 1|$ (four structural breaks)

5) $|100, \sigma = 1|\, 50, \sigma = 1.6\, |100, \sigma = 1|\, 150, \sigma = 1.2\, |100, \sigma = 1|$ (four non-equidistant structural breaks)

For each of these five data cases three different distributions are considered. These are a zero mean Gaussian distribution (G), a t-distribution with 5 degrees of freedom (t5) and an exponential distribution with parameter $\lambda = 1$ (exp) shifted to have zero mean. We make use of scaling in order to obtain the desired standard deviations. This leads us to 15 data cases. For each data case 10 000 replications are conducted and all methods introduced in Section 3 are applied at a significance level of 5%. All permutation tests are executed with 2000 permutations. For every blockwise procedure we always divide the data into $N = 10$ equidistant blocks and proceed in the same way on subsamples in case of splitting. In order to reduce the computational burden the two tests which rely on weight functions are only carried out for $a = 1.5$ and the Gaussian

weighting based on the analysis in the previous subsection. The corresponding rejection rates are given in Table 4.

Table 4: Rejection rates for all tests introduced in Section 3 in the five different volatility scenarios on page 12 for Gaussian distributions (G), t-distributions with 5 degrees of freedom (t5) and exponential distributions (exp) given in percent

		T_{Four}	T_{CUS}	T_{Mood}	T_{cf}	T_{Log}
$\mathbb{H}_0$	G	5.2	4.9	5.3	5.0	5.4
	t5	5.0	3.5	5.1	5.3	5.1
	exp	5.0	3.0	4.8	5.2	5.1
1 break	G	71.0	89.3	78.4	44.1	74.6
	t5	47.2	51.1	65.8	35.1	47.4
	exp	49.1	31.8	99.0	26.3	41.4
2 breaks	G	78.5	2.6	38.1	55.2	82.7
	t5	50.6	1.9	29.0	42.9	52.3
	exp	44.3	1.4	83.6	31.0	41.1
4 breaks	G	77.7	0.7	14.6	55.6	80.1
	t5	70.1	0.7	20.4	64.8	69.7
	exp	56.1	0.9	70.4	43.7	53.7
4 noneq. breaks	G	88.9	0.3	14.6	80.8	92.0
	t5	56.2	0.7	12.6	65.5	60.8
	exp	44.7	0.8	42.6	43.3	46.4

In general we observe that the rejection rates for data generated from the t-distribution with five degrees of freedom are always lower than the corresponding ones for Gaussian data. Apparently, heavier tails make the detection of the breaks more difficult. Under the null hypothesis all methods keep the significance level of 5%. For the data with $n = 200$ observations stemming from the t- and the exponential distribution the asymptotics of the CUSUM test does not provide a good approximation. Its rejection rates are quite low under the null. As expected it leads to the best results for Gaussian data with one volatility change, but loses a considerable amount of power in presence of multiple structural breaks due to masking effects. The Mood type test clearly outperforms its competitors in the case of exponential data in all but the fifth data scenario. The procedure therefore might have problems in non-equidistant settings. For data from the Gaussian and the t-distribution, we observe a similar loss of efficiency as for the CUSUM test. The problem is quite similar, because

the procedure proposed by Ross is based on a two-sample test. Much like the CUSUM approach, it therefore implicitly anticipates one structural break at a time and thus always divides the sample into two parts. The tests based on T_{Four} and T_{Log} lead to competitive results overall. In particular they suffer considerably less from multiple structural breaks in comparison to the CUSUM and the Mood type test. As a consequence, they clearly outperform the CUSUM test for all distributions under study and the test proposed by Ross for the symmetric distributions in case of more than one volatility change. Among the two, T_{Log} leads to slightly higher rejection rates under alternatives. The procedure using T_{cf} performs similarly to them, but is inferior to both in all considered scenarios. This is in accordance with the results in Section 4.1.
Next, we want to check which impact a change of the structural break position has on the results. Therefore, we repeat the simulation for Gaussian data and one volatility change increasing from $\sigma = 1$ to $\sigma = 1.5$. However, this time the structural break occurs after 10%, 20%, ..., 80%, or 90% of the 200 observations. We apply the same tests as before but leave out the permutation test based on T_{cf} due to its poor results in the previous simulations. The rejection rates based on 1000 replications are presented in Table 5.

Table 5: Rejection rates of different volatility tests in case of one volatility change after different proportions of the data, denoted by p, given in rounded percent

p	10%	20%	30%	40%	50%	60%	70%	80%	90%
T_{Four}	28.6	50.3	65.3	67.4	75.1	70.1	58.7	45.3	20.3
T_{CUS}	9.3	25.3	57.1	80.1	88.5	88.5	79.6	55.7	11.8
T_{Mood}	16.2	47.7	66.9	71.2	78.5	75.0	68.9	59.7	38.2
T_{Log}	25.3	47.3	63.1	68.2	76.8	75.4	65.8	57.6	34.2

For all methods the best results are attained when the structural break is in the middle of a sample. The rejection rates decrease the more the volatility clusters differ in size. The best procedure for a break in the middle of the sample, the CUSUM test, is most severely affected by this. Its rejection rates drop the most when the data is almost exclusively generated from one model. Hereby in general for all tests better results are achieved if the majority of the data has a standard deviation of 1 rather than 1.5. In those cases the overall variance of the data is smaller and thus the decision process more precise.
We also want to assess the importance of assumption we made dealing with zero mean data. We thus repeat the simulations for Gaussian samples in case of one and two structural breaks, respectively. The permutation test based on T_{Four} is applied to 1000 replications of each scenario. Hereafter, each of the 2000 data sets is centered by the corresponding arithmetic mean and again passed to the permutation test. In case of one volatility change we get rejection rates

of 70.3% for the uncentered data and 70.2% for the centered equivalents. For two strucural breaks the results are 76% and 75.2%. We thus conclude that demeaning has a negligable effect on the results. Therefore, the zero mean assumption might be weakened to an unknown but constant mean.
Finally, we compare the tests on artificial data with linearly increasing variance. For this purpose we simulate 500 datasets containing 200 observations each. Each time the first 100 observations are generated from the standard Gaussian distributions. The remaining ones are drawn from Gaussian distributions with mean zero and a variance increasing by 0.025 each time, that is, $\sigma^2(101) = 1.025, \sigma^2(102) = 1.05, \ldots, \sigma^2(200) = 3.5$. The rejection rates are 83% for T_{Four}, 94% for T_{CUS}, 88% for T_{Log} and 89 % for T_{Mood}. We repeat the simulation now increasing the variance by 0.1 each time. This time all test reject the null hypotheses of constant variance in all cases. The histograms of the estimated change point positions for each test are given in Figure 1.

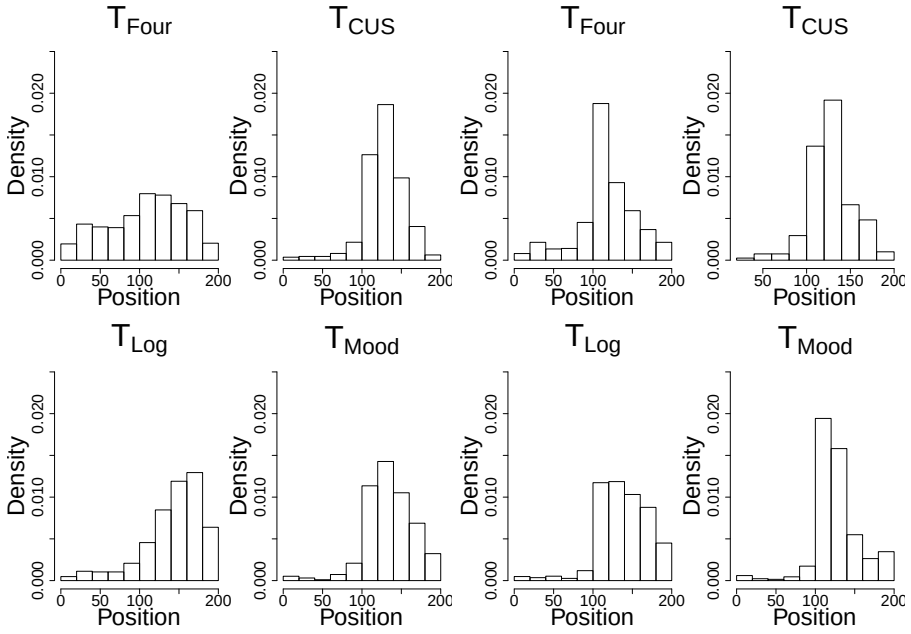


Figure 1: Histograms of estimated positions of potential structural changes for different tests in case of a linear increase of the variance by 0.025 (left) and 0.1 (right) after 100 observations.

The results suggest that the test based on T_{Four} is not able to allocate the structural break position well in case of a comparatively small linear increase of the variance. In contrast, its competitors correctly detect potential change points almost exclusively after the 100th observation. For the scenario with increased growth of the variance shown in the right part of the figure the test based on T_{Four} performs much better. Apparently it is more sensitive to the

first change in the variance process than to the linear increase after the 100th observation.

4.3 Estimated number and position of estimated structural breaks

A good test for structural breaks should have high rejection rates under various alternatives. However, it also must determine the causes of the heterogeneity adequately after a rejection. Otherwise, it connects the correct rejection with an irrelevant event leading to false conclusions. We therefore take a closer look at the estimated number and location of the structural breaks of the methods. Thereby, we focus on the results for the blockwise procedures based on the statistics T_{Four} and T_{Log}. They performed the best in terms of power for the t-distributions with five degrees of freedom with multiple structural breaks. Both tests are conducted in the recursive manner as explained in Section 3.2. In Table 6 the mean number of presumable structural breaks is listed for both methods and all five scenarios introduced on page 12 for data stemming from the t-distributions studied in the previous section.

Table 6: Mean number of detected structural breaks estimated for the permutation tests using the blockwise statistics T_{Four} and T_{Log} on data generated by the t-distribution with five degrees of freedom for each of the five data cases introduced in Section 4.2 given in percent. In brackets the mean number of detected structural breaks among the samples with rejection is given

	$\mathbb{H}_0$	1 break	2 breaks	4 breaks	4 noneq. breaks
T_{Four}	0.07 (1.4)	0.70 (1.48)	1.05 (2.08)	1.66 (2.37)	2.16 (3.84)
T_{Log}	0.08 (1.56)	0.76 (1.60)	1.15 (2.20)	1.74 (2.48)	2.28 (3.75)

The results reveal that on average both methods do not detect all structural breaks, particularly if several volatility changes are present. It is particularly problematic to detect all four volatility changes placed equidistantly in setting four. Note that each data case is chosen such that all rejection rates are below 1 in order to make the tests comparable. Therefore, it is obvious that not all structural breaks are detected. They are simply not that clear to all procedures under study by construction. The test based on T_{Log} rejects more often and thus unsurprisingly finds more structural breaks on average. The differences are quite small though. Next, we consider the cases where the tests reject. As indicated by the results given in brackets, both methods determine a reasonable number of presumable structural breaks if they reject at all. Hereby, the test based on T_{Four} estimates the number of structural breaks more adequately under the null hypothesis and in presence of four structural breaks, but the results are again similar.

The presumable change point positions for the case of two and four equidistant structural breaks are presented in Figure 4.3 for both methods. Keep in mind that in case of rejection both tests locate presumable change points in the same way, see Section 3.2. The different results are thus mainly a consequence of additional rejections of the method based on T_{Log} on subsamples. These additional presumable change point positions do not coincide with the true ones in most of the times. Hence, the test based on T_{Four} locates the true change point positions much more precisely. This suggests that tackling the detection of structural breaks via Fourier type transformations is advantageous in comparison to a simpler blockwise approach. The price paid by a somewhat smaller rejection rate is outweighed by a more exact location of the change position.

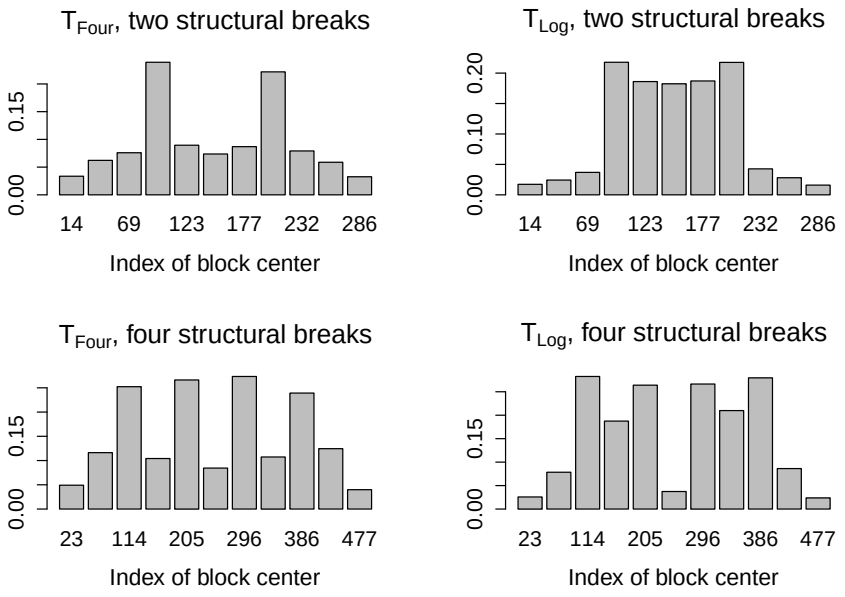


Figure 2: Change point positions estimated by the permutation procedure for two (top) and four (bottom) present structural changes for the test based on T_{Four} (left) and T_{Log} (right)

4.4 Application

For further illustration of the methods we consider a data set related to gold mining in South Africa introduced in Rowland and Sichel (1960). The extraction and processing of gold is quite expensive. Therefore, ore samples are collected in mines and checked for their gold content by chemical examinations to discover promising cultivating regions. However, taking representative samples is

quite complicated due to the highly irregular gold concentration. This leads to unreliable results in particular for new samplers. Therefore, experienced supervisors are used to resample part of the work. The data considered in the following and presented in Figure 3 is taken from Jandhyala et al (2002). It consists of 157 logarithmised ratios of gold contents in samples collected by a junior sampler to samples mined by a supervisor at the same locations. The data is arranged in chronological order. Values with a large modulus therefore indicate a high disagreement between the two corresponding samples indicating a non representative sampling by the junior sampler. We follow the arguments of Jandhyala et al (2002) stating that unreliable sampling does not induce a bias in the data, but rather leads to more unstable results. In accordance to these authors we thus assume zero mean observations. In order to check the reliability of the samples collected by the junior sampler we test the constancy of the variance of the measurements with the tests performing best on artificial data.

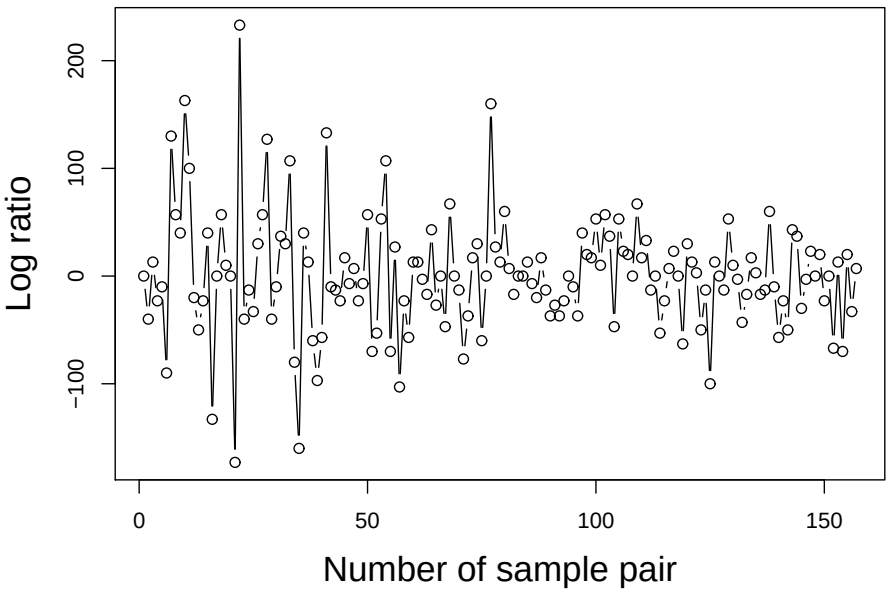


Figure 3: Logarithmised ratios of gold contents in samples collected by a junior sampler to samples mined by a supervisor at the same 157 locations given in Jandhyala et al (2002).

Using five almost equidistant blocks, the Gaussian weight function and $a = 1$ the test based on T_{Four} detects a variance change at observation 81. This is consistent with the results in Jandhyala et al (2002) and seems plausible in view of the data. Note that the test applied in Jandhyala et al (2002) relies on normally distributed data, while the detection by our procedure is valid

regardless of this assumption. The test based on T_{Log} detects a change point at observation 42 neglecting both the region of comparatively high variance around observation 55 as well as the high peak at 77. Therefore, the procedure based on T_{Four} appears to be more favorable in this case.

5 Extension to testing structural breaks in kurtosis

The concept introduced in Section 2.1 as well as the testing procedure proposed in 3.1 are not restricted to testing for constant volatility. They can easily be adapted to test the constancy of any relevant parameter of the data for which reasonable estimators are available. This can be achieved in a straightforward way: one simply substitutes the estimator (2) by another measure reflecting the quantity of choice.
To illustrate the procedure and give a first impression on its performance, we consider structural changes in the kurtosis. Kurtosis has recently gained additional attention in financial applications and is increasingly regarded as an alternative risk measure, see for example Martín et al (2012) and the references given therein. In analogy to Section 4.2, we consider 1000 repetitions of each of the four data cases

1) $|6000, N(0,1)|$ $(\mathbb{H}_0)$
2) $|3000, N(0,1)|\, 3000, t_{10}|$ (one structural break)
3) $|3000, N(0,1)|\, 3000, t_{10}\, |3000, N(0,1)|$ (two structural breaks)
4) $|3000, N(0,1)|\, 3000, t_{10}\, |3000, N(0,1)|\, 3000, t_{10}\, |3000, N(0,1)|$ (four structural breaks)

Hereby $N(0,1)$ denotes standard Gaussian data, while t_{10} stands for observations drawn from a t-distribution with 10 degrees of freedom. In order to eliminate the effect of different volatilities, the data from the t-distribution is standardized by the corresponding theoretical standard deviation. Note that we work with considerably larger sample sizes than for the volatility, because the kurtosis is much harder to estimate.
We apply four tests to the datasets. The following three of them rely on the permutation principle. The first one is the proposed adaption to the procedure motivated in Section 2.1. The test statistic is

$$\widetilde{T}_{Four} = \frac{1}{n^2} \sum_{j,k=1}^{N} \tau_j \tau_k I_w \left(\log\left(\widehat{\kappa}_j\right) - \log\left(\widehat{\kappa}_k\right) \right),$$

where

$$\widehat{\kappa}_j = \frac{1}{\tau_j} \sum_{t \in B_j} R_t^4 \quad \text{for} \quad j = 1, \ldots, N$$

is the natural estimator of kurtosis in the j-th block for zero mean observations with unit variance. $\widetilde{T}_{Four}$ resembles the statistic T_{Four} in (4), but the volatility estimators $\widehat{\sigma}_j^2$ are replaced by the kurtosis estimators $\widehat{\kappa}_j$ for all $j = 1, \ldots, N$. The weighting is conducted using the Gaussian weight function with parameter $a = 1.5$. Additional simulations not reported here show that both the weighting scheme as well as the parameter a do not have a huge effect on the results, as it is the case for the volatility. In the same way we adopt the statistics T_{Log} introduced in (5) resulting in

$$\widetilde{T}_{Log} = n \log \left(\sum_{t=1}^{n} R_t^4 \right) - \sum_{j=1}^{N} \tau_j \log \left(\widehat{\kappa}_j \right).$$

The third statistic under study is

$$\widetilde{T}_{Max} = \max_{1 \leq j,k \leq N} |\widehat{\kappa}_j - \widehat{\kappa}_k|.$$

It is an intuitive measure to capture changes in the kurtosis process. In addition to the three permutation tests, an asymptotic CUSUM test for the kurtosis is considered. Following Bertram (unpublished manuscript available at `http://www.wiwi.uni-hannover.de/fileadmin/statistik/papers/CUSQ.pdf`) we define

$$\tilde{T}_{CUS} = \max_{1 \leq h \leq n} \sqrt{n} \left| \frac{\widehat{\kappa}_{1:h}}{\widehat{\kappa}_{1:n}} - \frac{h}{n} \right|$$

and derive corresponding critical values from its asymptotics. Hereby $\widehat{\kappa}_{1:h} = \frac{1}{h} \sum_{t=1}^{h} R_t^4$ denotes the kurtosis estimator on the first j observations for $1 \leq h \leq n$.

The results in Table 7 lead to similar conclusions as for the volatility. The CUSUM procedure attains the highest rejection rates as long as the data contain only one or two structural breaks. With increasing number of changes in kurtosis, the method's performance worsens in comparison to the other tests, although this happens not as fast as in the volatility case. The tests using $\tilde{T}_{Four}$ and $\tilde{T}_{Log}$ on the contrary reject more often with increasing number of structural breaks. As before, their results are quite close. Both tests outperform the method using $\tilde{T}_{Max}$ in all settings considered. In analogy to Section 4.3 we examined the positions of the structural breaks determined by these two tests. As for the volatility, the test based on $\tilde{T}_{Four}$ determines the location of the true structural breaks more accurately, cf. Section 4.3.

6 Conclusion and outlook

In this work we consider the problem of testing whether a sequence of random variables has constant volatility over time. In accordance with Spokoiny (2009), Davies et al (2012) and other authors we assume that the volatility is approximately piecewise constant. We thus propose a new test based on blockwise volatility estimates and a Fourier type transformation. Simulations suggest

Table 7: Rejection rates for tests introduced in Section 5 in the four different kurtosis scenarios

	$\tilde{T}_{Four}$	$\tilde{T}_{CUS}$	$\tilde{T}_{Log}$	$\tilde{T}_{Max}$
$\mathbb{H}_0$	4.7	5.1	4.6	4.9
1 break	22.2	81.1	22.7	16.0
2 breaks	28.1	56.6	29.5	22.0
4 breaks	65.1	56.5	65.7	43.6

that it attains competitive rejection rates in comparison to other methods testing for global constant volatility and symmetric distributions. In particular, it is recommendable if several structural changes are possible. In case of rejection it locates the structural break positions adequately. An illustrative application to process control data also leads to meaningful results. Since our concept is not only applicable when testing for constant volatility, we extend the method to testing for a global constant kurtosis and obtain comparable results. Analogous measures can be constructed in order to check the constancy of skewness, any quantiles or more complicated quantities. Also the concept allows to apply any other notion of volatility, tail-behavior, etc. suitable in a given context. For example, robust estimators of scale may be preferable if outliers are an issue. Such extensions, asymptotics for the test statistics introduced, as well as adequate algorithms for the choice of blocks are subjects of current research. Another issue to be considered is the extension of the procedures to dependent observations and in an even more general framework than just volatility; see for instance Brooks et al (2005); Feunou and Tédongap (2012); Harvey and Siddique (1999). In this connection, appropriate resampling techniques in the dependent case are also goals of future work. In closing we note that the suggested method has been investigated basically by means of Monte Carlo results. A theoretical investigation is also possible and first results show that the test statistic attains a simple asymptotic limit, in probability, and that under certain assumptions and with suitable scaling the asymptotic null distribution distribution of the test is normal. These results however are the subject of ongoing work.

Acknowledgements: The work was supported in part by the Collaborative Research Center 823, Project C3 (Analysis of Structural Change in Dynamic Processes), of the German Research Foundation and by grant Nr.11699 of the Special Account for Research Grants (ELKE) of the National and Kapodistrian University of Athens.

References

Brooks C, Burke SP, Heravi S, Persand G (2005) Autoregressive Conditional Kurtosis. Journal of Financial Econometrics 3(3):399–421

Chen J, Gupta AK (1997) Testing and Locating Variance Changepoints with Application to Stock Prices. Journal of the American Statistical Association 92(438):739–747

Davies L, Höhenrieder C, Krämer W (2012) Recursive Computation of Piecewise Constant Volatilities. Computational Statistics & Data Analysis 56(11):3623–3631

Epps TW (1993) Characteristic Functions and Their Empirical Counterparts: Geometrical Interpretations and Applications to Statistical Inference. The American Statistician 47(1):33–38

Epps TW (1999) Limiting behavior of the ICF test for normality under Gram-Charlier alternatives. Statistics & Probability Letters 42(2):175–184

Feunou B, Tédongap R (2012) A Stochastic Volatility Model with Conditional Skewness. Journal of Business and Economic Statistics 30(4):576–591

Fisher RA (1935) The Design of Experiments. Oliver & Boyd, Oxford

Fried R (2012) On the Online Estimation of Local Constant Volatilities. Computational Statistics & Data Analysis 56(11):3080–3090

Good P (2005) Multiple tests. Permutation, Parametric and Bootstrap Tests of Hypotheses pp 79–84

Guégan D (2015) Non–stationary samples and meta–distribution. Basu,A, Samanta, T and Sen Gupta, A (eds): Statistical Paradigms Recent Advances and Reconciliations pp 121–141

Harvey CR, Siddique A (1999) Autoregressive Conditional Skewness. Journal of Financial & Quantitative Analysis 34(4):465–487

Hlávka Z, Hušková M, Kirch C, Meintanis SG (2012) Monitoring Changes in the Error Distribution of Autoregressive Models Based on Fourier Methods. Test 21(4):605–634

Hlávka Z, Hušková M, Kirch C, Meintanis SG (2015) Fourier–type tests involving martingale difference processes. Econometric Reviews DOI 10.1080/07474938.2014.977074

Hsu DA (1977) Tests for Variance Shift at an Unknown Time Point. Journal of the Royal Statistical Society 26(3):279–284

Hušková M, Meintanis SG (2006a) Change Point Analysis Based on Empirical Characteristic Functions. Metrika 63(2):145–168

Hušková M, Meintanis SG (2006b) Change-point Analysis Based on Empirical Characteristic Functions of Ranks. Sequential Analysis 25(4):421–436

Jandhyala VK, Fotopoulos SB, Hawkins DM (2002) Detection and estimation of abrupt changes in the variability of a process. Computationl Statistics & Data Analysis 40(1):1–19

Jiménez-Gamero MD, Alba-Fernández V, Muñoz-García J, Chalco-Cano Y (2009) Goodness-of-Fit Tests Based on Empirical Characteristic Functions. Computational Statistics & Data Analysis 53(12):3957–3971

Juhl T, Xiao Z (2013) Nonparametric Tests of Moment Condition Stability. Econometric Theory 29(1):90–114

Loéve M (1977) Probability Theory I. Springer–Verlag, New York

Martín MdML, García CG, Pérez JG (2012) Treatment of kurtosis in financial markets. Physica A: Statistical Mechanics and its Applications 391(5):2032–2045

Matteson DS, James NA (2014) A Nonparametric Approach for Multiple Change Point Analysis of Multivariate Data. Journal of the American Statistical Association 109(505):334–345

Meintanis SG (2005) Permutation Tests for Homogeneity Based on the Empirical Characteristic Function. Nonparametric Statistics 17(5):583–592

Meintanis SG, Swanepoel J, Allison J (2014) The Probability Weighted Characteristic Function and Goodness-of-Fit Testing. Journal of Statistical Planning and Inference 146:122–132

Mercurio D, Spokoiny V (2004) Statistical Inference for Time-inhomogeneous Volatility Models. The Annals of Statistics 32(2):577–602

Mood AM (1954) On the Asymptotic Efficiency of Certain Nonparametric Two-Sample Tests. The Annals of Mathematical Statistics 25(3):514–522

Pardo-Fernández JC, Jiménez-Gamero MD, El Ghouch A (2015) A Nonparametric ANOVA–type Test for Regression Curves Based on Characteristic Functions. Scandinavian Journal of Statistics 42:197–213

Peña D (2005) Analisis de Series Temporales. Alianza Editorial, 319, Madrid

Potgieter CJ, Genton MG (2013) Characteristic Function-based Semiparametric Inference for Skew-symmetric Models. Scandinavian Journal of Statistics 40(3):471–490

Ross GJ (2013) Modelling Financial Volatility in the Presence of Abrupt Changes. Physica A: Statistical Mechanics and its Applications 392(2):350–360

Rowland RSJ, Sichel H (1960) Statistical quality control of routine underground sampling. Journal of the South African Institute of Mining and Metall 60:251–284

Spokoiny V (2009) Multiscale Local Change Point Detection with Applications to Value-at-Risk. The Annals of Statistics 37(3):1405–1436

Stărică C, Granger C (2005) Nonstationarities in Stock Returns. Review of Economics and Statistics 87(3):503–522

Steland A, Rafajłowicz E (2014) Decoupling Change-point Detection Based on Characteristic Functions: Methodology, Asymptotics, Subsampling and Application. Journal of Statistical Planning and Inference 145:49–73

Tenreiro C (2009) On the choice of the smoothing parameter for the BHEP goodness-of-fit test. Computational Statistics and Data Analysis 53:1038–1053

Vassiliou E, Demetriou IC (2005) An Adaptive Algorithm for Least Squares Piecewise Monotonic Data Fitting. Computational Statistics & Data Analysis 49(2):591–609

Vostrikova LJ (1981) Detecting Disorder in Multidimensional Random Process. Soviet Mathematics Doklady 24:55–59

Wied D, Arnold M, Bissantz N, Ziggel D (2012) A New Fluctuation Test for Constant Variances with Applications to Finance. Metrika 75(8):1111–1127