Abstract: **Detection of instabilities in various models**

Marie Hušková Charles University Prague Czech Republic

The talk focuses on two setups:

• Structural breaks in some CAPM with functional data, sequential monitoring procedures. Mostly based based on the paper:

Robust monitoring of CAPM portfolio betas II, couathors: Z.Praskova, J.Steinebach, O. Chochola, JMVA 132, 2014, 58 – 81.

• Detection of changes connected with martingale hypothesis testing Based on the paper:

FourierType Tests Involving Martingale Difference Processes,

coauthors: S. Meintanis, C. Kirch, Z. Hlavka, Econometric Reviews 36, 2017, 468 – 492.

Both parts contain introduction, theoretical results, simulations and applications.

Robust monitoring of CAPM portfolio betas II

Ondřej Chochola[a], **Marie Hušková**[b], **Zuzana Prášková**[c], **Josef G. Steinebach**[d,*]

[a,b,c] Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, CZ – 186 75 Praha 8, Czech Republic

[d] Universität zu Köln, Mathematisches Institut, Weyertal 86 – 90, D – 50 931 Köln, Germany

May 6, 2014

Abstract

In this work, we extend our study in Chochola et al. [7] and propose some robust sequential procedure for the detection of structural breaks in a *Functional* Capital Asset Pricing Model (FCAPM). The procedure is again based on M-estimates and partial weighted sums of M-residuals and "robustifies" the approach of Aue et al. [3], in which ordinary least squares (OLS) estimates have been used. Similar to [3], and in contrast to [7], high-frequency data can now also be taken into account. The main results prove some null asymptotics for the suggested test as well as its consistency under local alternatives. In addition to the theoretical results, some conclusions from a small simulation study together with an application to a real data set are presented in order to illustrate the finite sample performance of our monitoring procedure.

AMS 2000 subject classifications: 60F17, 60G10, 60J65, 62F35, 62L10, 62P05

Keywords: Robust monitoring; Functional capital asset pricing model; Portfolio beta; M-estimate; Change-point detection.

1. Introduction and statistical framework

Main aim of this work is to continue and extend our study in Chochola et al. [7] concerning the robust monitoring of CAPM portfolio betas. The Capital Asset Pricing Model (CAPM), introduced by Sharpe [18] and subsequently modified by many authors (see, e.g. Lintner [14], Merton [15] and others), is a still very popular and widely used model for evaluating the risk of a portfolio of assets with respect to the market risk. However, it is also well-known that the pricing of assets and predictions of risks may be incorrect and misleading if the model parameters $\boldsymbol{\beta}_i$ are varying over time. As in Aue et al. [3], we adopt here the arguments of Ghysels [9] and study a (piecewise) unconditional CAPM, rather than a conditional version of the latter (cf., e.g., Andersen et al. [1] for a comprehensive review), since in many cases misspecified conditional CAPMs tend to produce larger pricing errors. For a more extensive discussion of this fact, we refer to Aue et al. [3], Sections 1 and 2, and the references mentioned therein.

Indeed, contributing to avoid pricing and prediction errors was the main motivation for Aue et al. [3] in constructing a sequential monitoring procedure for the testing of the stability of portfolio betas. The corresponding stopping rules in [3] are based on comparing the (ordinary) least squares estimate (OLS) of the beta from a historical data set (training period) to that from sequentially incoming new observations, and they were able to take high-frequency data into account which is a typical situation in nowadays' market analyses (see also Chochola et al. [7] and the references mentioned therein).

Since OLS estimates may be sensitive with respect to outliers, we tried to "robustify" the Aue et al. [3] approach in [7] by making use of M-estimates instead of least squares estimates and so are also able to deal with heavier tail distributions than the OLS procedure. In a first step, however, we confined ourselves there to a study of the CAPM without high-frequency observations. Aim of our present work now is to extend the latter study to the *Functional* Capital Asset Pricing Model (FCAPM) taking also high-frequency observations into account. It will turn out that, even in this more general situation, some moment conditions may be relaxed (cf., e.g., (B.4) below compared to the corresponding assumption in [7]), but that, on the other hand and similar to Aue et al. [3], certain smoothness conditions have to be added concerning the model's intra-day behavior over time (see, e.g., (A.1)-(A.3), (B.5) and (B.7) below).

Note that, via L_p-m-approximability type conditions (cf. (B.4)-(B.5) below), our model is suitable for covering general types of weak dependencies rather than strong dependencies in the sense of long memory. Monitoring

*Corresponding author.
E-mail addresses: chochola@karlin.mff.cuni.cz (O. Chochola), huskova@karlin.mff.cuni.cz (M. Hušková) praskova@karlin.mff.cuni.cz (Z. Prášková), jost@math.uni-koeln.de (J.G. Steinebach)

procedures in the latter situation are still open for future work. On the other hand, in contrast to [3], our present approach is now applicable to data sets under heavy-tailed (leptocurtic) and contaminated distributions observed at high frequencies, which is certainly more useful in real data applications. The price to pay, however, is that more involved techniques than those used in Chochola et al. [7] are required now and the computational complexity increases as well. Nevertheless, a similar robust sequential monitoring procedure can be constructed for the FCAPM portfolio betas, now also covering a high-frequency situation as described below.

We would like to mention, however, that our focus here is on the methodological and theoretical side, trying to extend the work of Aue et al. [3] by using a robust approach and that of Chochola et al. [7] by including high-frequency situations. Moreover, for the sake of illustration and comparison, we used the same data set as in [3] for our application and a similar setting in the small simulation study of Section 3.

Our statistical framework in the sequel will be as follows. We consider the model

$$\boldsymbol{r}_i(s) = \boldsymbol{\alpha}_i + \boldsymbol{\beta}_i r_{iM}(s) + \boldsymbol{\varepsilon}_i(s), \quad i \in \mathbb{Z}, \quad s \in [0,1], \tag{1.1}$$

where $\boldsymbol{r}_i(s) = (r_{i,1}(s), \ldots, r_{i,d}(s))^T$ is a d-dimensional vector of (functional) log-returns at (say) "day" i and "intra-day time" s, $r_{iM}(s)$ is the log-return of the market portfolio at day i and time s, and $\boldsymbol{\varepsilon}_i(s) = (\varepsilon_{i,1}(s), \ldots, \varepsilon_{i,d}(s))^T$ are d-dimensional (functional) error terms. The $\boldsymbol{\alpha}_i$'s and $\boldsymbol{\beta}_i$'s are d-dimensional unknown parameters, and the $\boldsymbol{\beta}_i$'s are the parameters of interest, usually called the "portfolio betas". Note that the sequence $\{(\boldsymbol{r}_i(\cdot), r_{iM}(\cdot))\}$ is a $(d+1)$-dimensional (functional) time series satisfying certain conditions to be specified below.

We assume that a training sample of size m with no instabilities is available, i.e.,

$$\boldsymbol{\alpha}_1 = \ldots = \boldsymbol{\alpha}_m =: \boldsymbol{\alpha}_0 = (\alpha_1^0, \ldots, \alpha_d^0)^T, \quad \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m =: \boldsymbol{\beta}_0 = (\beta_1^0, \ldots, \beta_d^0)^T, \tag{1.2}$$

where $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ are unknown parameters. The problem of the instability of the portfolio betas is formulated as a testing problem, that is, we want to test the null hypothesis

$$H_0: \ \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m = \boldsymbol{\beta}_{m+1} = \ldots$$

of no "change versus" the alternative

$$H_A: \ \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_{m+k^*} \neq \boldsymbol{\beta}_{m+k^*+1} = \ldots$$

of a "structural break" at an unknown change-point $k^* = k_m^*$.

For later convenience we reformulate our model as follows:

$$r_{i,j}(s) = \alpha_j^0 + \beta_j^0 r_{iM}(s) + (\alpha_j^1 + \beta_j^1 r_{iM}(s))\delta_m I\{i > m+k^*\} + \varepsilon_{i,j}(s), \ \ j = 1, \ldots, d, \ \ i = 1, 2, \ldots, \ \ s \in [0,1], \tag{1.3}$$

where $k^* = k_m^*$ is the change-point and $\alpha_j^0, \beta_j^0, \alpha_j^1, \beta_j^1, \delta_m$ are unknown parameters.

As in [7], our test procedures will be generated by convex loss functions $\varrho_1, \ldots, \varrho_d$ with a.s. derivatives $\varrho_j' = \psi_j$ called score functions having further properties to be specified later. The estimators $\widehat{\alpha}_{jm} = \widehat{\alpha}_{jm}(\psi_j), \widehat{\beta}_{jm} = \widehat{\beta}_{jm}(\psi_j)$ of α_j^0, β_j^0 based on the training sample are defined as minimizers of

$$\sum_{i=1}^{m} \sum_{\nu=1}^{n} \varrho_j(r_{i,j}(s_\nu) - a_j - b_j r_{iM}(s_\nu)) \tag{1.4}$$

w.r.t. a_j, b_j, for $j = 1, \ldots, d$, where $s_\nu = \nu/n$, $\nu = 1, \ldots, n$, are n equidistant intra-day time-points.
The test procedure constructed below will be based on functionals of partial sums of weighted M-residuals, which are defined as follows:

$$\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)) = (\psi_1(\widehat{\varepsilon}_{i,1}(s_\nu)), \ldots, \psi_d(\widehat{\varepsilon}_{i,d}(s_\nu)))^T \tag{1.5}$$

with

$$\begin{aligned} &\widehat{\boldsymbol{\varepsilon}}_i(s_\nu) = (\widehat{\varepsilon}_{i,1}(s_\nu), \ldots, \widehat{\varepsilon}_{i,d}(s_\nu))^T, \\ &\widehat{\varepsilon}_{i,j}(s_\nu) = r_{i,j}(s_\nu) - \widehat{\alpha}_{jm} - \widehat{\beta}_{jm} r_{iM}(s_\nu). \end{aligned} \tag{1.6}$$

A suitable test statistic based on the first $m+k$ (functional) observations is

$$\widehat{Q}(k,m) = \Big(\frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+k} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu))\Big)^T \widehat{\boldsymbol{\Sigma}}_m^{-1} \Big(\frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+k} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu))\Big) \tag{1.7}$$

where $n = n(m)$ (see below) and the matrix $\widehat{\boldsymbol{\Sigma}}_m$ is an estimator of the asymptotic variance (matrix)

$$\boldsymbol{\Sigma} = \lim_{m \to \infty} var\Big\{\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \int_0^1 r_{iM}(s) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s)) ds\Big\} \tag{1.8}$$

based on the first m observations. Details will be discussed later.

For notational convenience and later use, we introduce the notations, for $i \in \mathbb{Z}$ and $n \in \mathbb{N}$,

$$\boldsymbol{z}_i = (z_{i,1}, \ldots, z_{i,d})^T = \int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))ds, \tag{1.9}$$

$$\widehat{\boldsymbol{z}}_i = \widehat{\boldsymbol{z}}_{i,n} = (\widehat{z}_{i,1}, \ldots, \widehat{z}_{i,d})^T = \frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)), \tag{1.10}$$

$$\widetilde{\boldsymbol{z}}_i = \widetilde{\boldsymbol{z}}_{i,n} = (\widetilde{z}_{i,1}, \ldots, \widetilde{z}_{i,d})^T = \frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)), \tag{1.11}$$

so that

$$\widehat{Q}(k,m) = \Big(\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\widehat{\boldsymbol{z}}_i\Big)^T \widehat{\boldsymbol{\Sigma}}_m^{-1}\Big(\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\widehat{\boldsymbol{z}}_i\Big) \qquad \text{and}$$

$$\boldsymbol{\Sigma} = \lim_{m\to\infty} var\Big\{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\boldsymbol{z}_i\Big\}.$$

Similar to [7], we reject the null hypothesis as soon as the test statistic exceeds a critical level for the first time, i.e., when

$$\widehat{Q}(k,m)/q_\gamma(k/m) \geq c$$

for an appropriately chosen $c = c_\gamma(\alpha)$, where $q_\gamma(t)$, $t \in (0,\infty)$, is a suitable boundary (weight) function. In this case we stop the procedure and confirm a structural break, otherwise we continue monitoring. The associated stopping rule is given by

$$\tau_m = \tau_m(\gamma) = \inf\{1 \leq k \leq \lfloor mT \rfloor : \widehat{Q}(k,m)/q_\gamma(k/m) \geq c\}, \tag{1.12}$$

with $\inf\emptyset := \infty$. Here T is a fixed positive number, that is, for practical reasons, we have a so-called *closed-end procedure* again. The following class of weight functions q_γ can be used, e.g.,

$$q_\gamma(t) = (1+t)^2\Big(\frac{t}{t+1}\Big)^{2\gamma}, \quad t \in (0,\infty), \tag{1.13}$$

where γ is a tuning constant taking values in $[0, 1/2)$. The critical value c will be chosen such that, under H_0, for $\alpha \in (0,1)$ (fixed),

$$\lim_{m\to\infty} P\big(\tau_m < \infty\big) = \alpha, \tag{1.14}$$

i.e., the overall asymptotic level (false alarm rate) is α and, under H_A,

$$\lim_{m\to\infty} P\big(\tau_m < \infty\big) = 1, \tag{1.15}$$

i.e., the test is consistent (has asymptotic power 1).

The rest of the paper is organized as follows. The main results including the assumptions and limit properties of the test procedures are presented and discussed in Section 2. Section 3 reports on the results of a small simulation study and an application to a real data set. The proofs of our main results are given in Section 4, whereas Section 5 contains some auxiliary lemmas to be used in the proofs.

2. Assumptions and main results

Compared to [7], the assumptions on the sequence $\{(\varepsilon_{i,1}(\cdot), \ldots, \varepsilon_{i,d}(\cdot),\, r_{iM}(\cdot))\}_{i\in\mathbb{Z}}$ and on the loss functions $\varrho_1, \ldots, \varrho_d$ (or equivalently on the score functions $\psi_1, \ldots, \psi_d$) have to be extended as follows.

We assume on ψ_j, the distributions of $\varepsilon_{0,j}(s)$ and $\lambda_j(x;s) = -E\psi_j(\varepsilon_{0,j}(s) - x)$, $j = 1, \ldots, d$, $s \in [0,1]$, $x \in \mathbb{R}$:

(A.1) ψ_j are nondecreasing functions, $\lambda_j(0,s) = 0$, $\lambda_j'(0,\cdot)$ is continuous on $[0,1]$, $\lambda_j'(x,s) := \frac{\partial}{\partial x}\lambda_j(x,s)$ exists in a neighborhood of 0 for all $s \in [0,1]$,

$$|\lambda_j'(x, s+z) - \lambda_j'(0, s+z)| \leq D_0|x|, \quad |x| \leq x_0, \quad s, s+z \in [0,1], \quad |z| \leq z_0,$$

and

$$|\lambda_j'(0, x+s) - \lambda_j'(0,s)| \leq D_0|x|, \quad |x| \leq x_0, \quad x+s, s \in [0,1],$$

for some $x_0, z_0, D_0 > 0$;

(A.2) $\int_0^1 \lambda_j'(0,s)\,ds \int_0^1 \lambda_j'(0,s)\,Er_{0M}(s)^2\,ds > \Big(\int_0^1 \lambda_j'(0,s)\,Er_{0M}(s)\,ds\Big)^2$;

[Note that, via the Cauchy-Schwarz inequality, we have at least "$\geq$" in the latter condition, so we just assume nondegeneracy.]

(A.3) $\sup_{s\in[0,1]} E|\psi_j(\varepsilon_{0,j}(s))|^3 < \infty$ and

$$E\big|\psi_j(\varepsilon_{0,j}(s)+t_2) - \psi_j(\varepsilon_{0,j}(s)+t_1)\big|^2 \leq C_1|t_2 - t_1|, \quad |t_1|,|t_2| \leq c_0,\ s\in[0,1],$$

for some $c_0, C_1 > 0$.

For later applications, let us briefly recall some of the most often considered ψ_j-functions. The classical choice $\psi_j(x) = x,\ x\in\mathbb{R}^1$, leads to the ordinary least squares (OLS) and L_2-residuals. A choice of $\psi_j(x) = \operatorname{sign} x,\ x\in\mathbb{R}^1$, leads to L_1-estimators and L_1-residuals. Huber [12] introduced $\psi_j(x) = x\,I\{|x|\leq K\} + K\operatorname{sign} x\, I\{|x|>K\},\ x\in\mathbb{R}^1$, for some $K>0$, which is one of the most often used score functions, usually known as the Huber function.

For a vector-valued random variable $\boldsymbol{X}$ define

$$\|\boldsymbol{X}\|_p = \big(E|\boldsymbol{X}|^p\big)^{1/p}, \quad p\geq 1,$$

the L_p-norm of $\boldsymbol{X}$, where $|\boldsymbol{X}|$ denotes the Euclidean norm of $\boldsymbol{X}$.

Concerning the assumptions on $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\varepsilon}_i(\cdot)\}$ we follow the setup in Aue et al. [3], but instead of fourth moment assumptions used there it typically suffices here to have second or $(2+\Delta)$-moment conditions:

(B.1) For any $i\in\mathbb{Z}$, $r_{iM}(\cdot) = h(\boldsymbol{\xi}_i(\cdot), \boldsymbol{\xi}_{i-1}(\cdot),\ldots)$, where $h(\cdot)$ is a measurable function, $\{\boldsymbol{\xi}_i(\cdot)\}$ is a sequence of i.i.d. random functions, and $\sup_{s\in[0,1]} E|r_{0M}(s)|^3<\infty$.
[Note that $\{r_{iM}(\cdot) : i\in\mathbb{Z}\}$ is a stationary and ergodic sequence.]

Remark 2.1. For the sake of simplicity, we assume a third moment condition in Assumptions (A.3) and (B.1). With some more technical effort, the latter can be replaced by a $(2+\Delta)$-moment condition with some $\Delta>0$ (cf. Lemma 5.1 (i)-(ii) below).

(B.2) For any $i\in\mathbb{Z}$, $\boldsymbol{\varepsilon}_i(\cdot) = \boldsymbol{g}(\boldsymbol{\zeta}_i(\cdot), \boldsymbol{\zeta}_{i-1}(\cdot),\ldots)$, where $\boldsymbol{g}(\cdot)$ is a measurable function, $\{\boldsymbol{\zeta}_i(\cdot)\}$ is a sequence of i.i.d. random functions having some further properties to be specified later.
[Note that $\{\boldsymbol{\varepsilon}_i(\cdot) : i\in\mathbb{Z}\}$ is also a stationary and ergodic sequence.]

(B.3) The sequences $\{\boldsymbol{\xi}_i(\cdot)\}$ and $\{\boldsymbol{\zeta}_i(\cdot)\}$ are independent.

(B.4) For all $i\in\mathbb{Z}$,

$$\sup_{s\in[0,1]} \sum_{L=1}^{\infty} \|r_{iM}(s) - r_{iM}^{(L)}(s)\|_2 < \infty,$$

where

$$r_{iM}^{(L)}(\cdot) = h(\boldsymbol{\xi}_i(\cdot), \boldsymbol{\xi}_{i-1}(\cdot),\ldots,\boldsymbol{\xi}_{i-L+1}(\cdot), \boldsymbol{\xi}_{i-L}^{(L)}(\cdot), \boldsymbol{\xi}_{i-L-1}^{(L)}(\cdot),\ldots),$$

with $\boldsymbol{\xi}_{i-L}^{(L)}(\cdot), \boldsymbol{\xi}_{i-L-1}^{(L)}(\cdot),\ldots$ being i.i.d. with the same distribution as $\boldsymbol{\xi}_0(\cdot)$ and independent of $\{\boldsymbol{\xi}_i(\cdot)\}$.
[Note that $r_{iM}^{(L)}(\cdot) \overset{\mathcal{D}}{=} r_{iM}(\cdot) \overset{\mathcal{D}}{=} r_{0M}(\cdot)$ for all $i\in\mathbb{Z}$ and $L\geq 1$.]

(B.5) With $\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot)) = (\psi_1(\varepsilon_{i,1}(\cdot)),\ldots,\psi_d(\varepsilon_{i,d}(\cdot)))^T$, for all $i\in\mathbb{Z}$, it holds that

$$\sup_{s\in[0,1]} \sup_{|\boldsymbol{a}|\leq a_0} \sum_{L=1}^{\infty} \|\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s)-\boldsymbol{a}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i^{(L)}(s)-\boldsymbol{a})\|_2 < \infty$$

for some $a_0>0$, where

$$\boldsymbol{\varepsilon}_i^{(L)}(\cdot) = \boldsymbol{g}(\boldsymbol{\zeta}_i(\cdot), \boldsymbol{\zeta}_{i-1}(\cdot),\ldots,\boldsymbol{\zeta}_{i-L+1}(\cdot), \boldsymbol{\zeta}_{i-L}^{(L)}(\cdot), \boldsymbol{\zeta}_{i-L-1}^{(L)}(\cdot),\ldots),$$

with $\boldsymbol{\zeta}_{i-L}^{(L)}(\cdot), \boldsymbol{\zeta}_{i-L-1}^{(L)}(\cdot),\ldots$ being i.i.d. with the same distribution as $\boldsymbol{\zeta}_0(\cdot)$ and independent of $\{\boldsymbol{\zeta}_i(\cdot)\}$.

Remark 2.2. Assumption (B.5) could be weakened as follows, but then the proofs would require somewhat more technicalities:

(B.5') $\limsup_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} \sup_{|\boldsymbol{a}|\leq a_0} \sum_{L=1}^{\infty} \|\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)-\boldsymbol{a}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i^{(L)}(s_\nu)-\boldsymbol{a})\|_2 < \infty$

for some $a_0>0$, with $s_\nu = \nu/n$, $\nu=1,\ldots,n$, and $\{\boldsymbol{\varepsilon}_i^{(L)}(\cdot)\}$ as in (B.5).

As in Aue et al. [3] and Chochola et al. [7], the above assumptions are motivated by the work of Hörmann and Kokoszka [10] on the concept of L_p-m-approximability, but could be relaxed here to a certain extent.

The following conditions, assuming that the processes under consideration are smooth functions of the intra-day parameter $s \in [0,1]$, are weakened versions of the corresponding conditions in Aue et al. [3].

First we also make the following "high-frequency" assumption:

(B.6) We let $n = n(m) \to \infty$ as $m \to \infty$.

Secondly, we assume smoothness of the $r_{iM}(\cdot)$'s and $\psi_j(\varepsilon_{i,j}(\cdot))$'s:

(B.7) For all $i \in \mathbb{Z}$, $j = 1, \ldots, d$, with $s_\nu = 1/n$ as above and $n = n(m) \to \infty$,

a) $$\lim_{m\to\infty} (\log m) \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|r_{iM}(s_\nu) - r_{iM}(s_\nu - h)\|_2 = 0$$

and

b) $$\lim_{m\to\infty} (\log m) \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|\psi_j(\varepsilon_{i,j}(s_\nu)) - \psi_j(\varepsilon_{i,j}(s_\nu - h))\|_2 = 0.$$

Remark 2.3. It will be obvious from the proofs below that, if the L_2-approximability conditions in Assumptions (B.4) and (B.5) are replaced by corresponding $L_{2+\Delta}$-approximability (with some $\Delta > 0$), then the convergence rate condition in (B.7) can be avoided, i.e., (B.7) can be replaced by

(B.7') For all $i \in \mathbb{Z}$, $j = 1, \ldots, d$, with $s_\nu = 1/n$ as above and $n = n(m) \to \infty$,

a) $$\lim_{m\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|r_{iM}(s_\nu) - r_{iM}(s_\nu - h)\|_{2+\Delta} = 0$$

and

b) $$\lim_{m\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|\psi_j(\varepsilon_{i,j}(s_\nu)) - \psi_j(\varepsilon_{i,j}(s_\nu - h))\|_{2+\Delta} = 0.$$

Remark 2.4. The theoretical results below as well as the applications to the real data set work with equidistant grid points being the same for all components. Nevertheless, going through the proofs this assumption can be relaxed, e.g., working with more general $s_{\nu,j}$'s, $j = 1, \ldots, d$, under accordingly modified assumptions. Moreover, having a closer look at the test statistic defined through (1.7), (1.10) and (2.4), we realize that the test procedures depend on the observations through

$$\widehat{\boldsymbol{z}}_i = \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)),$$

that are averages over time grids s_ν, i.e., averages over the intra-day behavior, which also work for asynchronous data.

Next we present our results on the limit behavior of the test procedures, both under the null hypothesis H_0 as well as under the alternative H_A.

2.1. Asymptotic results

Theorem 2.1. *Let Assumptions* (A.1)-(A.2), (B.1)-(B.7) *and* (1.13) *with* $\gamma \in [0, 1/2)$ *be satisfied and*

$$\widehat{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma} = o_P(1) \quad (m \to \infty), \tag{2.1}$$

where, with the $\boldsymbol{z}_i$*'s from* (1.9),

$$\boldsymbol{\Sigma} = \lim_{m\to\infty} var\left\{\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \boldsymbol{z}_i\right\} = E[\boldsymbol{z}_0 \boldsymbol{z}_0^T] + \sum_{i=1}^{\infty} E[\boldsymbol{z}_0 \boldsymbol{z}_i^T + \boldsymbol{z}_i \boldsymbol{z}_0^T], \tag{2.2}$$

and $\boldsymbol{\Sigma}$ *is a positive definite matrix. Then, under the null hypothesis* H_0 *,*

$$\max_{1\le k\le \lfloor mT\rfloor} \left(\frac{\widehat{Q}(k,m)}{q_\gamma(k/m)}\right) \xrightarrow{\mathcal{D}} \sup_{0<t<T/(T+1)} \left(\frac{\sum_{j=1}^{d} W_j^2(t)}{t^{2\gamma}}\right) \quad (m \to \infty),$$

where $\{W_j(t),\ t \in [0,1]\}$*,* $j = 1, \ldots, d$*, are independent (standard) Brownian motions (Wiener processes).*

The proof of Theorem 2.1 is postponed to Section 4.

It follows from Assumptions (A.1)-(A.2) and (B.1)-(B.5) that $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot))\}$ are independent sequences. Then Lemma 2.1 and Theorem 4.2 in Hörmann and Kokoszka [10] imply that the series in (2.2) converges (component-wise) absolutely.

Now we turn to the model under local alternatives, i.e.

$$r_{i,j}(s) = \alpha_j^0 + \beta_j^0 r_{iM}(s) + (\alpha_j^1 + \beta_j^1 r_{iM}(s))\delta_m I\{i > m+k^*\} + \varepsilon_{i,j}(s), \quad j = 1, \dots, d, \quad i = 1, 2, \dots, \quad s \in [0,1], \tag{2.3}$$

with $\delta_m \to 0$ and $k^* < \lfloor mT \rfloor$.

Theorem 2.2. *Let Assumptions* (A.1)-(A.2), (B.1)-(B.7) *and* (1.13) *with* $\gamma \in [0, 1/2)$ *be satisfied and*

$$\widehat{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma} = o_P(1) \quad (m \to \infty),$$

where $\boldsymbol{\Sigma}$ *is as in Theorem 2.1. Then, under* (2.3), *with* $\delta_m \to 0$, $|\delta_m| m^{1/2} \to \infty$, $\liminf_{m\to\infty}(\lfloor mT \rfloor - k^*)/m > 0$, *and* $\beta_j^1 \neq 0$ *for at least one* j,

$$\max_{1 \le k \le \lfloor mT \rfloor} \left(\frac{\widehat{Q}(k,m)}{q_\gamma(k/m)} \right) \xrightarrow{P} \infty \quad (m \to \infty).$$

The proof of Theorem 2.2 is also postponed to Section 4.

Remark 2.5. a) By Theorem 2.1, the assertion (1.14) holds true if $c_\gamma(\alpha)$ satisfies

$$P\left(\sup_{0<t<T/(T+1)} \left(\frac{\sum_{j=1}^d W_j^2(t)}{t^{2\gamma}} \right) \ge c_\gamma(\alpha) \right) = \alpha,$$

where $c_\gamma(\alpha)$ can either be obtained by simulation of the limit distribution or by an application of a suitable form of bootstrap based on the training sample.

b) Theorem 2.2 implies the consistency of the test, i.e., the validity of (1.15) (asymptotic power 1).

2.2. Estimation of the variance matrix

In this section we deal with an estimator of the asymptotic variance (matrix) $\boldsymbol{\Sigma}$ as given in (2.2). Notice that $\boldsymbol{\Sigma} = \sum_{k=-\infty}^{\infty} \boldsymbol{\Gamma}_k$, where $\boldsymbol{\Gamma}_k = E[\boldsymbol{z}_0 \boldsymbol{z}_k^T]$ for $k \ge 0$ and $\boldsymbol{\Gamma}_{-k} = \boldsymbol{\Gamma}_k^T$.

We consider an estimator of $\boldsymbol{\Sigma}$ based on the first m (functional) observations defined as

$$\widehat{\boldsymbol{\Sigma}}_m = \sum_{|k|<q} \omega_q(k) \widehat{\boldsymbol{\Gamma}}_k \tag{2.4}$$

where $q = q(m)$, $\omega_q(k) = \omega(k/q)$ and ω is a kernel specified below, and $\widehat{\boldsymbol{\Gamma}}_k$ is the k-th lag sample covariance corresponding to $\boldsymbol{\Gamma}_k$, i.e.,

$$\widehat{\boldsymbol{\Gamma}}_k = \begin{cases} \dfrac{1}{m} \displaystyle\sum_{i=1}^{m-k} \widehat{\boldsymbol{z}}_i \widehat{\boldsymbol{z}}_{i+k}^T, & k \ge 0, \\ \widehat{\boldsymbol{\Gamma}}_{-k}^T, & k < 0, \end{cases} \tag{2.5}$$

with the $\widehat{\boldsymbol{z}}_i$'s as defined in (1.10), based on the $r_{iM}(\cdot)$'s from (1.1) and $\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i)$'s according to the M-residuals as given in (1.5) and (1.6).

Theorem 2.3. *Let Assumptions* (A.1), (A.2), *and* (B.1)-(B.7) *be satisfied. Let* $\widehat{\boldsymbol{\Sigma}}_m$ *be the estimator of* $\boldsymbol{\Sigma}$ *given in* (2.4) *with a kernel* $\omega_q(k) = \omega(k/q)$ *satisfying the following conditions:*

(i) $\omega(0) = 1$*;*

(ii) ω *is a symmetric and Lipschitz-continuous function;*

(iii) ω *has bounded support;*

(iv) *the Fourier transform of* ω *is also Lipschitz-continuous and integrable;*

(v) $q(m) = O(\log m)$ $(m \to \infty)$.

Then

$$\widehat{\boldsymbol{\Sigma}}_m = \boldsymbol{\Sigma} + o_P(1) \quad (m \to \infty).$$

We can work, e.g., either with the Bartlett kernel

$$\omega(x) = (1 - |x|) I\{|x| \le 1\} \tag{2.6}$$

or with the flat-top kernel

$$\omega(x) = \begin{cases} 1, & |x| \le \frac{1}{2}, \\ 2(1 - |x|), & \frac{1}{2} < |x| < 1, \\ 0, & |x| \ge 1. \end{cases} \tag{2.7}$$

3. Applications and simulations

In this section we present some results from a small simulation study as well as an application to a real data set in order to illustrate the finite sample performance of our monitoring procedure based on the test statistic (1.7) with boundary function (1.13).

First we discuss some aspects which are common to both the simulation study and the application. Since the asymptotic distribution of the test statistic given in Theorem 2.1 coincides with the one derived in Chochola et al. [7] (cf. also Remark 2.5), we can use the critical values given in Table 1 of [7].

The question of the choice of the tuning constant γ has also been discussed in [7] and the recommendation given there remains valid, i.e., if a change is to be expected "early" after the training period, then γ near to 0.5 is advisable, whereas for "late change scenarios", small γ's are recommended. A choice of $\gamma = 0.25$ provides a reasonably good balance between these two scenarios and is thus used here.

We consider the L_2, Huber and L_1 ψ-functions and always apply the same function to all coordinates.

It remains to choose the kernel function and especially its bandwidth q in the estimator of the variance matrix suggested in (2.4). In this aspect, we use the results of Chochola [6] which show that it can be difficult to set a proper q a priori for the Bartlett or the flat-top kernel, because it depends on the degree of dependency of the data. Thus better results can be obtained using a data-driven adaptive choice of the bandwidth based on the work of Andrews [2] and implemented in the statistical software R as described in Zeileis [19]. Differences between possible kernel choices are not too big, so that we always use the Bartlett kernel here.

As an illustration of a possible application of our robust monitoring, we investigate the data set used in Aue et al. [3] in more detail. Recalling this data set, it consists of five stocks from different sectors of S&P 100, namely Boeing (BA), Bank of America (BAC), Microsoft (MSFT), AT&T (T), and Exxon Mobile (XOM). As the market portfolio, the S&P 100 index itself is used.

The intra-day behavior of the process $\{\boldsymbol{r}_i(s) : s \in [0;1]; i \in \mathbb{Z}\}$, which is defined at time s as the difference between the log-prices of the stocks at time s and $s+15$ min, is thus sampled every 15 minutes during any trading day i. The process $r_{iM}(\cdot)$ is defined analogously.

The historical training period starts on January 29, 2001 and consists of 120 trading days for which the values of the portfolio betas under consideration appear reasonably stable. The choice of the beginning of the period is motivated by the fact that, prior to January 29, 2001, the tick size (i.e. the smallest value the price can change) was different. The monitoring horizon for the closed-end procedure was selected as 360 days, corresponding to $T = 3$ for our stopping rule in (1.12). This covers the 9/11 event, the influence of which we want to study.

The stability of the historical portfolio betas was checked via moving windows estimates presented in Figure 1. The figure shows Huber estimates of portfolio betas based on moving windows of 10 trading days for each company throughout the historical and monitoring periods, but the figures look similar for L_2 estimates. The solid black vertical line marks the end of the historical period (120 days), whereas the dashed black line marks the last day, when the estimate is not influenced by the observations from the monitoring period. The grey lines refer in the same way to the 9/11 event. Since "no change" during the historical period is assumed, we tested for a change in this period via L_2 and Huber retrospective procedures and this assumption could be confirmed.

The BAC and T estimates seem to be stable throughout the whole period, whereas there is a small temporary influence of the 9/11 event on MSFT and a very big one on BA. Finally there seems to be a shift in the portfolio beta of XOM right after the end of the training period. We come back to these observations later on.

Next we discuss the robust monitoring itself. Figure 2 shows values of the normalized test statistic, i.e., $\widehat{Q}(k,m)/(c_{0.25}(0.05)\,q_\gamma(k/m))$, for the L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedure and for various combinations of stocks, which are given in the heading of each chart. On the x-axis the number of trading days is shown starting from the beginning of the monitoring. A vertical grey dashed line marks the September 11, 2001, terrorist attack, the horizontal one (at value 1) indicates the critical line, due to the normalization of the statistic.

When all companies are considered together, we get the same results as in Aue et al. [3] for the L_2 procedure. The critical value is extremely exceeded. For the Huber and L_1 procedures the crossing still occurs, but in a much more moderate way.

It is possible to get further insight by looking at the stocks individually. In view of the conclusions from Figure 1, we examined Boeing (BA) and Exxon (XOM). Portfolio betas of the three remaining companies (BAC, MSFT and T) do not show any sign of a change as can be seen from the last chart. For Boeing (BA) we can see the extreme influence of the 9/11 event on the L_2 monitoring procedure. In case of robust procedures this has a much smaller impact, the critical value, however, is still crossed right after the event. For Exxon (XOM) and robust monitoring, the critical line is crossed already before the 9/11 attack - there is a steady increase in the test statistic from the beginning of the monitoring on, which is in line with the conclusions from Figure 1. By applying the retrospective procedure to the XOM data for the first 120 and 240 trading days, respectively, it turned out that no change could be detected based on the period of length 120, but using 240 days a change was indicated close to trading day 110. This explains why the critical line is already crossed before the 9/11 event.

It is of further interest, whether the change in Boeing's (BA) portfolio betas after the 9/11 was only temporary or persistent. In order to find out, we use the same monitoring procedure, but exclude 5 or 10 trading days after the 9/11 from the monitoring. This can be seen in Figure 3. We can see that, if 5 days are excluded, then the

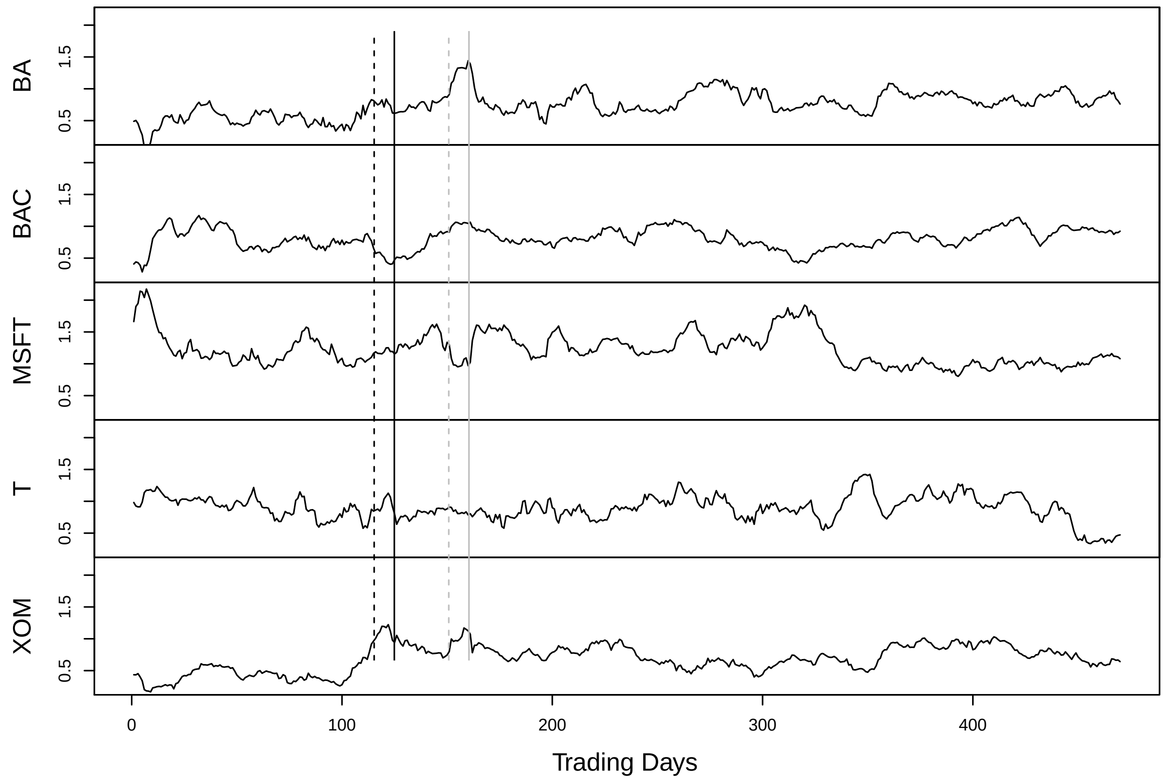

Figure 1: Huber estimates of portfolio beta based on moving windows of 10 trading days. Black solid vertical line marks end of training period, grey one marks the 9/11 event. Dashed lines indicate the beginning of moving windows that are already influenced by these events.

crossings are much smaller and, if 10 days are excluded, then the terrorist attack has no impact at all and the change is not indicated until mid of March 2002.

In order to further quantify the finite sample properties of the monitoring procedure, a small simulation study has been conducted. We simulated data according to the model (1.1), with $d = 2$, $\boldsymbol{\alpha}_0 = (1/2, 1/2)^T$, $\boldsymbol{\beta}_0 = (1, 1)^T$ for simplicity. Various settings have been used for the market portfolio log-returns $r_{iM}(\cdot)$ and the error terms $\boldsymbol{\varepsilon}_i(\cdot)$. The $r_{iM}(\cdot)$'s were either independent standard Brownian motions (denoted B_i) or, similarly as in Aue et al. [3], chosen as a functional AR(1) process, i.e.,

$$r_{iM}(s) = \rho \int_0^1 K(s,t) r_{i-1,M}(t)dt + \eta_i(s), \quad s \in [0, 1],$$

where $\{\eta_i(\cdot) : i \in \mathbb{Z}\}$ denotes a sequence of independent standard Brownian motions and $K(s;t) = c\exp(-|t - s|)$, with c such that the norm of K equals one. We chose $\rho = 0.1$ and $\rho = 0.4$ as the dependency coefficient and denote the models as AR(1;0.1) or AR(1;0.4). The random errors, in both coordinates, are either standard Brownian motions or, to illustrate the robustness of the monitoring procedures, we use a 5% contamination with Brownian motion having larger variance, i.e. $10B_i$ (denoted Mix).

Finally $m = 100$ or $m = 200$ and $T = 5$ were chosen, with a tuning constant $\gamma = 0.25$ in the boundary function, nominal level $\alpha = 5\%$, and the Bartlett kernel is used with an adaptive choice of the bandwidth q, as discussed at the beginning of this section. All results are based on 2000 repetitions.

First we have a look at the empirical levels presented in Table 1. We can see that the levels are approximately kept for the Huber and L_1 procedures in all scenarios considered. This, however, is no longer true for the L_2 procedure, especially in the case of the contaminated model. So, in order to compare the different procedures one would have to adjust them to possess the same empirical size.

In order to illustrate the properties of the test under the alternative hypothesis, we chose $k^* = 10$ and a unit change in both parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and in both coordinates. Figure 4 shows the densities of the detection delays $\tau_m - k^*$ for various choices of distributions of $r_{i,M}$ and $\boldsymbol{\varepsilon}_i$. As long as both are standard Brownian motions (B_i, B_i), the L_2 procedure performs better than Huber and L_1, while in case of (B_i, Mix) the L_2 procedure is outperformed by the robust ones, in particular by the Huber procedure. The latter effect is even more visible if all procedures are adjusted to the same empirical size (see also Table 1).

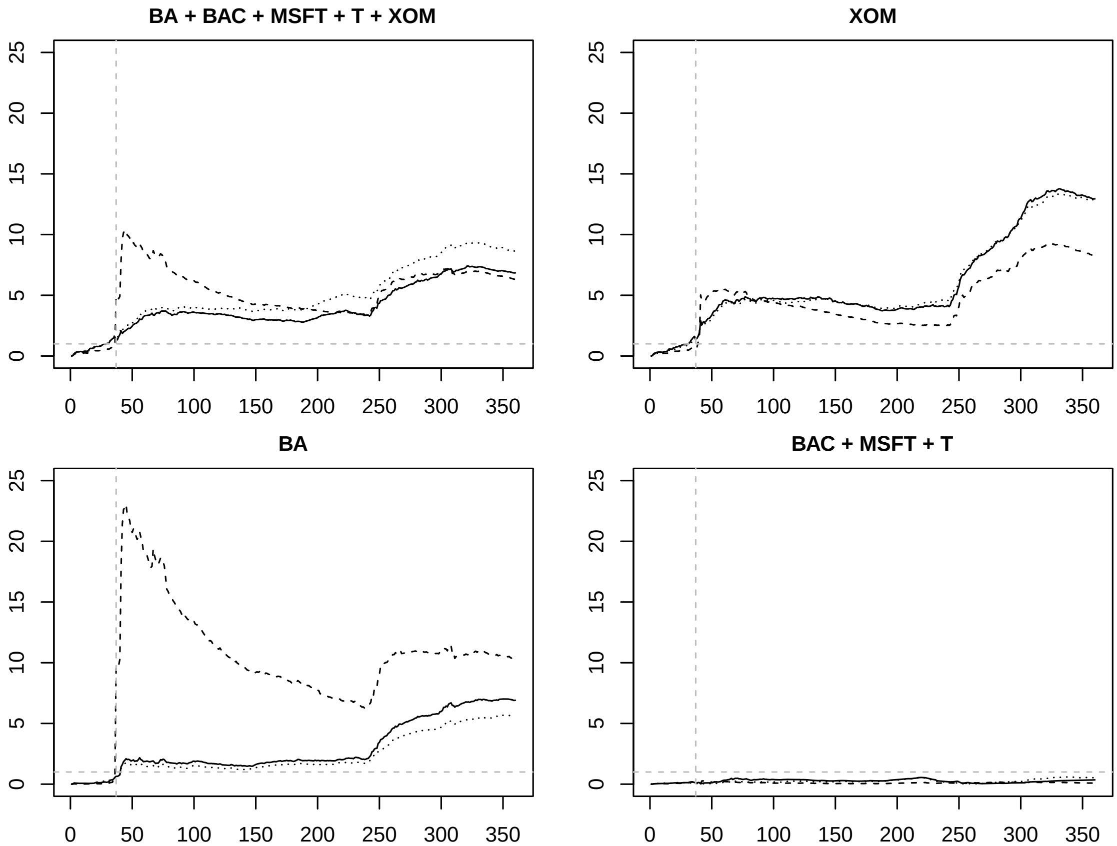

Figure 2: Normalized test statistics for the L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedures, various combinations of stocks - given in the heading of each chart. x-axis shows number of trading days from the beginning of the monitoring on.

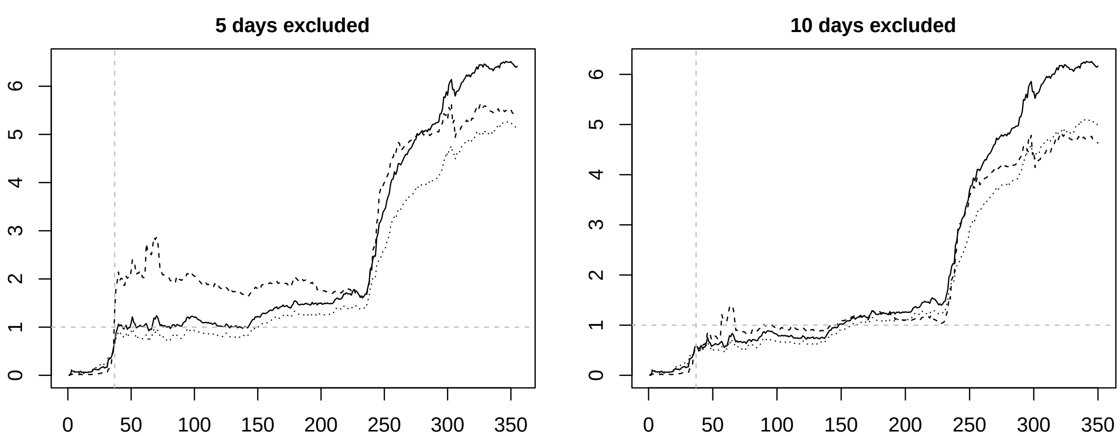

Figure 3: Boeing stock, normalized test statistics for L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedures. 5 or 10 days excluded from the monitoring after the 9/11.

In conclusion, in certain situations the robust monitoring procedures suggested in this work show definite advantages over the much more sensitive L_2 approach. They usually avoid overrejection of the test and are able to keep the approximate size. A choice of Huber's ψ-function seems to provide a good balance between robust and sensitive monitoring. If no prior knowledge is available on where to expect a possible change, a choice of the tuning constant $\gamma = 0.25$ in (1.13) appears to be appropriate.

r_{iM}	ε_i	m	L_2	Huber	L_1
B_i	B_i	100	8.4	7.0	5.9
		200	5.3	4.8	4.1
B_i	Mix	100	25.7	7.4	5.8
		200	14.5	4.6	3.9
AR(1; 0.1)	B_i	100	8.4	7.0	5.0
		200	6.4	5.4	4.4
AR(1; 0.4)	B_i	100	9.3	7.6	6.2
		200	6.7	6.1	5.6

Table 1: Empirical sizes at nominal level $\alpha = 5\%$ under H_0.

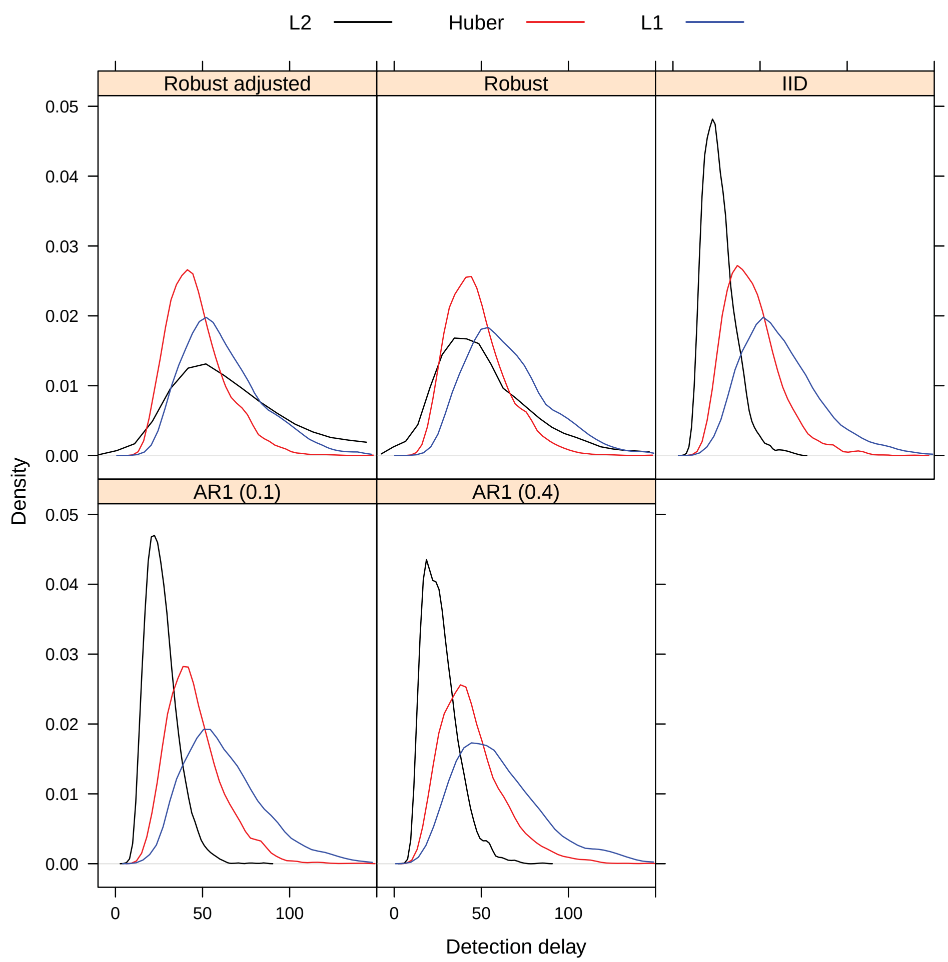

Figure 4: Densities of the detection delays for L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedures.

4. Proofs

Proof of Theorem 2.1. Similar to Chochola et al. [7], the proof can be given in three steps. Let us recall that we work with the model

$$r_{i,j}(s) = \alpha_j^0 + \beta_j^0 r_{iM}(s) + (\alpha_j^1 + \beta_j^1 r_{iM}(s))\delta_m I\{i > m + k^*\} + \varepsilon_{i,j}(s), \quad j = 1, \dots, d, \quad i = 1, 2, \dots, \quad s \in [0, 1], \tag{4.1}$$

as defined in (1.3).

1. In a first step we make use of asymptotic representations of the estimators $\widehat{\alpha}_{jm}$, $\widehat{\beta}_{jm}$ of α_j^0, β_j^0, $j = 1, \dots, d$, from (1.4). These estimators are based on the training sample only, so that we are in a non-sequential setup and can proceed in the same way as in treating the behavior of multivariate M-estimators. However, we need to take care of the dependency structure of the random error functions.

In the following it is convenient to introduce auxiliary estimators $\widehat{\alpha}^*_{jm}$ and $\widehat{\beta}^*_{jm}$ as minimizers of

$$\sum_{i=1}^{m}\sum_{\nu=1}^{n}\varrho_j(\varepsilon_{i,j}(s_\nu)-a_j^*/\sqrt{m}-b_j^* r_{iM}(s_\nu)/\sqrt{m}) \tag{4.2}$$

w.r.t. a_j^* and b_j^*, for $j=1,\dots,d$, where $s_\nu=\nu/n$, $\nu=1,\dots,n$. Clearly,

$$\widehat{\alpha}^*_{jm}=\sqrt{m}(\widehat{\alpha}_{jm}-\alpha_j^0),\quad \widehat{\beta}^*_{jm}=\sqrt{m}(\widehat{\beta}_{jm}-\beta_j^0). \tag{4.3}$$

Usually, the estimators $\widehat{\alpha}^*_{jm}$ and $\widehat{\beta}^*_{jm}$ can be obtained as solutions of the equations

$$\sum_{i=1}^{m}\sum_{\nu=1}^{n}\psi_j(\varepsilon_{i,j}(s_\nu)-(a_j^*+b_j^* r_{iM}(s_\nu))/\sqrt{m})=0, \tag{4.4}$$

$$\sum_{i=1}^{m}\sum_{\nu=1}^{n}\psi_j(\varepsilon_{i,j}-(a_j^*+b_j^* r_{iM}(s_\nu))/\sqrt{m})\widetilde{r}_{iM}=0, \tag{4.5}$$

w.r.t. a_j^*, b_j^*, for $j=1,\dots,d$.

Lemmas 5.2 and 5.3 below ensure that $\widehat{\alpha}^*_{jm}=O_P(1)$ and $\widehat{\beta}^*_{jm}=O_P(1)$ and, moreover, we get the asymptotic representations, as $m\to\infty$,

$$\widehat{\alpha}^*_m=\frac{1}{\int_0^1\lambda'(0,z)dz}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1\psi(\varepsilon_i(s))ds-\widehat{\beta}^*_m\frac{\int_0^1\lambda'(0,z)Er_{0M}^2(z)dz}{\int_0^1\lambda'(0,z)dz}+O_P(m^{-\eta}), \tag{4.6}$$

$$\widehat{\beta}^*_m=\frac{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1\psi(\varepsilon_i(s))\Big(r_{iM}(s)-\frac{\int_0^1\lambda'(0,z)Er_{0M}(z)dz}{\int_0^1\lambda'(0,z)dz}\Big)ds}{\int_0^1\lambda'(0,z)Er_{0M}^2(z)dz-\frac{\big(\int_0^1\lambda'(0,z)Er_{0M}(z)dz\big)^2}{\int_0^1\lambda'(0,z)dz}}+O_P(m^{-\eta}), \tag{4.7}$$

with some $\eta>0$ (cf. Remark 5.1).

2. Next, as a consequence of Lemmas 5.2-5.4 in combination with Remarks 5.1-5.2, we observe that the limit behavior of the weighted partial sums

$$\widehat{\boldsymbol{H}}(m,k)=(\widehat{H}_1(m,k),\dots,\widehat{H}_d(m,k))^T=\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)),\quad k=1,\dots,\lfloor mT\rfloor,$$

is the same as that of

$$\widetilde{\boldsymbol{H}}(m,k)=\frac{1}{\sqrt{m}}\Big(\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu))-\frac{k}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu))\Big),\quad k=1,\dots,\lfloor mT\rfloor.$$

In view of Lemma 5.5 (ii) together with Assumption (2.1), this further implies that the limit behavior of

$$\max_{1\le k\le\lfloor mT\rfloor}\widehat{Q}(k,m)/q_\gamma(k/m)$$

is the same as that of

$$\max_{1\le k\le\lfloor mT\rfloor}Q(k,m)/q_\gamma(k/m),$$

where

$$Q(k,m)=\boldsymbol{H}(m,k)^T\boldsymbol{\Sigma}^{-1}\boldsymbol{H}(m,k), \tag{4.8}$$

with

$$\boldsymbol{H}(m,k)=\frac{1}{\sqrt{m}}\Big(\sum_{i=m+1}^{m+k}\int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))\,ds-\frac{k}{m}\sum_{i=1}^{m}\int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))\,ds\Big),\quad k=1,\dots,\lfloor mT\rfloor.$$

3. In order to obtain the limit behavior of

$$\max_{1\le k\le\lfloor mT\rfloor}Q(k,m)/q_\gamma(k/m),$$

with $Q(k,m)$ from (4.8), we follow the lines of proof of Theorem 2.1 in Chochola et al. [7]. We just have to replace the random sequences and processes $\{\boldsymbol{Z}_i\}$, $\{\boldsymbol{Z}_i^{(L)}\}$ and $\{\boldsymbol{Z}_m(t)\}$ introduced there by

$$\boldsymbol{Z}_i=(Z_{i,1},\dots,Z_{i,d})^T=\int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))\,ds,\quad i=1,2,\dots,$$

$$\boldsymbol{Z}_i^{(L)}=(Z_{i,1}^{(L)},\dots,Z_{i,d}^{(L)})^T=\int_0^1 r_{iM}^{(L)}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i^{(L)}(s)),\quad i=1,2,\dots,\quad\text{and}$$

$$\boldsymbol{Z}_m(t)=\frac{1}{\sqrt{m}}\sum_{i=1}^{\lfloor mt\rfloor}\boldsymbol{Z}_i,\quad 0\le t\le T+1,$$

where $\int_0^1$ is to be taken componentwise.

The main step, that is, the weak convergence in the Skorokhod space $\mathcal{D}^d[0,T+1]$

$$\boldsymbol{Z}_m(\cdot) \overset{\mathcal{D}^d[0,T+1]}{\longrightarrow} \boldsymbol{W}_{\boldsymbol{\Sigma}}(\cdot),$$

where $\{\boldsymbol{W}_{\boldsymbol{\Sigma}}(t) : t \in [0,T+1]\}$ is a centered Gaussian process with covariance function $E[\boldsymbol{W}_{\boldsymbol{\Sigma}}(t)\boldsymbol{W}_{\boldsymbol{\Sigma}}^T(s)] = \min(t,s)\boldsymbol{\Sigma}$, is again a consequence of Billingsley [5], Theorem 21.1. An application of the continuous mapping theorem then completes the proof. For details we refer to Chochola et al. [7], pp. 383-385. □

Proof of Theorem 2.2. It suffices to show that

$$\frac{\widehat{Q}(\tilde{k},m)}{q_\gamma(\tilde{k}/m)} \overset{P}{\longrightarrow} \infty$$

for suitably chosen $\tilde{k}$. We take $\tilde{k} = k^* + (mT - k^*)/2$. In view of our assumptions on $\widehat{\boldsymbol{\Sigma}}_m$ and the choice of $\tilde{k}$ it suffices to treat

$$\begin{aligned} &\frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j(\widehat{\varepsilon}_{i,j}(s_\nu)) \\ &= \frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu) - (\widehat{\alpha}^*_{mj} + \widehat{\beta}^*_{mj} r_{iM}(s_\nu))/\sqrt{m} + (\alpha_j^1 + \beta_j^1 r_{iM}(s_\nu))\delta_m\big), \end{aligned}$$

where $\widehat{\alpha}^*_{mj} = O_P(1)$ and $\widehat{\beta}^*_{mj} = O_P(1)$. Therefore it is enough to study

$$\frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu) - (a + b r_{iM}(s_\nu))/\sqrt{m} + (\alpha_j^1 + \beta_j^1 r_{iM}(s_\nu))\delta_m\big)$$

for $|a| + |b| \le C$, $C > 0$. Proceeding analogously to the proof of Lemma 5.4 and recalling that $\delta_m \to 0$, but $|\delta_m|\sqrt{m} \to \infty$, we get

$$\Big|E^* \frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu) - (a + b r_{iM}(s_\nu))/\sqrt{m} + (\alpha_j^1 + \beta_j^1 r_{iM}(s_\nu)\delta_m\big)\Big| \overset{P}{\longrightarrow} \infty$$

and

$$var^*\Big\{\frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu) - (a + b r_{iM}(s_\nu))/\sqrt{m} + (\alpha_j^1 + \beta_j^1 r_{iM}(s_\nu))\delta_m\big)\Big\} = O_P(1),$$

uniformly in $|a| + |b| \le C$, $C > 0$.

From here, after some standard steps, we receive the desired assertion. □

Proof of Theorem 2.3.

Let $\boldsymbol{z}_i, \widehat{\boldsymbol{z}}_i$ and $\widetilde{\boldsymbol{z}}_i$ be as given in (1.9), (1.10) and (1.11), respectively. Recall $\widehat{\boldsymbol{\Gamma}}_k$ from (2.5) and further define, for $k \ge 0$,

$$\begin{aligned} \boldsymbol{\Gamma}_k &= \frac{1}{m} \sum_{i=1}^{m-k} \boldsymbol{z}_i \boldsymbol{z}_{i+k}^T, \\ \widetilde{\boldsymbol{\Gamma}}_k &= \frac{1}{m} \sum_{i=1}^{m-k} \widetilde{\boldsymbol{z}}_i \widetilde{\boldsymbol{z}}_{i+k}^T, \end{aligned}$$

and, for $k < 0$, put $\boldsymbol{\Gamma}_k = \boldsymbol{\Gamma}_{-k}^T$ and $\widetilde{\boldsymbol{\Gamma}}_k = \widetilde{\boldsymbol{\Gamma}}_{-k}^T$, respectively.

Let $\widehat{\boldsymbol{\Sigma}}_m$ be as given in (2.4) and put

$$\boldsymbol{\Sigma}_m = \sum_{|k|<q} \omega_q(k)\boldsymbol{\Gamma}_k$$

and

$$\widetilde{\boldsymbol{\Sigma}}_m = \sum_{|k|<q} \omega_q(k)\widetilde{\boldsymbol{\Gamma}}_k.$$

Then we have

$$\widehat{\boldsymbol{\Sigma}}_m = \boldsymbol{\Sigma}_m + (\widehat{\boldsymbol{\Sigma}}_m - \widetilde{\boldsymbol{\Sigma}}_m) + (\widetilde{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma}_m).$$

First, let us consider $\mathbf{\Sigma}_m$. Note, that $\{\boldsymbol{z}_i : i \in \mathbb{Z}\}$ is a stationary, L_2-approximable, centered sequence with $E\|\boldsymbol{z}_0\|^2 < \infty$, which follows from Assumptions (B.1)-(B.5) together with Lemma 2.1 in Hörmann and Kokoszka [10]. With a kernel ω_q satisfying conditions (i)-(v), all assumptions of Theorem 16.6 in Horváth and Kokoszka [11] are fulfilled. According to the latter theorem, we get

$$\mathbf{\Sigma}_m \xrightarrow{P} \mathbf{\Sigma} \quad \text{as} \quad m \to \infty. \tag{4.9}$$

In the next step we will show that

$$\widehat{\mathbf{\Sigma}}_m - \widetilde{\mathbf{\Sigma}}_m = O_p(q(m)m^{-1/4}). \tag{4.10}$$

Here we can proceed quite analogously to the corresponding part of the proof of Theorem 2.3 in Chochola et al. [7]. Obviously,

$$\widehat{\mathbf{\Sigma}}_m - \widetilde{\mathbf{\Sigma}}_m = \sum_{|k|<q} \omega_q(k)(\widehat{\mathbf{\Gamma}}_k - \widetilde{\mathbf{\Gamma}}_k)$$

and, since

$$\widehat{\boldsymbol{z}}_i \widehat{\boldsymbol{z}}_{i+k}^T - \widetilde{\boldsymbol{z}}_i \widetilde{\boldsymbol{z}}_{i+k}^T = (\widehat{\boldsymbol{z}}_i - \widetilde{\boldsymbol{z}}_i)(\widehat{\boldsymbol{z}}_{i+k} - \widetilde{\boldsymbol{z}}_{i+k})^T + (\widehat{\boldsymbol{z}}_i - \widetilde{\boldsymbol{z}}_i)\widetilde{\boldsymbol{z}}_{i+k}^T + \widetilde{\boldsymbol{z}}_i(\widehat{\boldsymbol{z}}_{i+k} - \widetilde{\boldsymbol{z}}_{i+k})^T,$$

we have

$$\sum_{0\le k<q} \omega_q(k)(\widehat{\mathbf{\Gamma}}_k - \widetilde{\mathbf{\Gamma}}_k) = \boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3,$$

where

$$\begin{aligned}
\boldsymbol{S}_1 &= \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu)[\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\mu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))][\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_{i+k}(s_\nu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T,\\
\boldsymbol{S}_2 &= \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu)[\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\mu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))]\boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))^T,\\
\boldsymbol{S}_3 &= \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))[\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_{i+k}(s_\nu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T.
\end{aligned}$$

For $s \in [0,1]$, set $\boldsymbol{d}_i(s) = \boldsymbol{a} + \boldsymbol{b} r_{iM}(s)$, where $\boldsymbol{a} = (a_1, \ldots, a_d)^T, \boldsymbol{b} = (b_1, \ldots, b_d)^T, \boldsymbol{d}_i(s) = (d_{i,1}(s), \ldots, d_{i,d}(s))^T$, and introduce

$$\begin{aligned}
\boldsymbol{S}_1^0 &= \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu)[\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu) - \boldsymbol{d}_i(s_\mu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))]\\
&\qquad \times [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu) - \boldsymbol{d}_{i+k}(s_\nu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T,\\
\boldsymbol{S}_2^0 &= \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu)[\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu) - \boldsymbol{d}_i(s_\mu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))]\\
&\qquad \times \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))^T,\\
\boldsymbol{S}_3^0 &= \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))\\
&\qquad \times [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu) - \boldsymbol{d}_{i+k}(s_\nu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T.
\end{aligned}$$

Now, for any $1 \le j, \ell \le d$,

$$\begin{aligned}
&E|r_{iM}(s_\mu) r_{i+k,M}(s_\nu)[\psi_j(\varepsilon_{i,j}(s_\mu) - d_{i,j}(s_\mu)/\sqrt{m}) - \psi_j(\varepsilon_{i,j}(s_\mu))]\\
&\qquad \times [\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu) - d_{i+k,\ell}(s_\nu)/\sqrt{m}) - \psi_\ell(\varepsilon_{i+k,\ell}(s_\nu))]|\\
&\le E|r_{iM}(s_\mu) r_{i+k,M}(s_\nu)|(E^*|\psi_j(\varepsilon_{i,j}(s_\mu) - d_{i,j}(s_\mu)/\sqrt{m}) - \psi_j(\varepsilon_{i,j}(s_\mu))|^2)^{1/2}\\
&\qquad \times (E^*|\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu) - d_{i+k,\ell}(s_\nu)/\sqrt{m}) - \psi_\ell(\varepsilon_{i+k,\ell}(s_\nu))|^2)^{1/2} \le C m^{-1/2},
\end{aligned}$$

uniformly in $\boldsymbol{a}, \boldsymbol{b}$ such that $\max_{1\le j\le d}(|a_j| + |b_j|) < C$ for some constant $C > 0$, where we have used the rule of iterated expectations (with E^* being the conditional expectation given $r_{iM}, i = 1, \ldots, m$), independence of $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot))\}$ (cf. Assumption (B.3)), the Cauchy-Schwarz inequality, Assumptions (B.1) and (A.3) and the boundedness of ω_q. From here we can conclude that, as $m \to \infty$, each (j,ℓ)-th element of $\boldsymbol{S}_1^0$ is $O_p(q(m)m^{-1/2})$, and so is $\boldsymbol{S}_1^0$, uniformly in $\boldsymbol{a}, \boldsymbol{b}$ such that $\max_{1\le j\le d}(|a_j| + |b_j|) < C$, with some $C > 0$.

Proceeding in the same way, we obtain $\boldsymbol{S}_2^0 = O_p(q(m)m^{-1/4})$ and $\boldsymbol{S}_3^0 = O_p(q(m)m^{-1/4})$, as $m \to \infty$, uniformly in $\boldsymbol{a}, \boldsymbol{b}$ such that $\max_{1\le j\le d}(|a_j| + |b_j|) < C$ for some constant $C > 0$.

Since $\widehat{\varepsilon}_{i,j}(s) = \varepsilon_{i,j}(s) - \widehat{\alpha}^*_{jm}/\sqrt{m} - \widehat{\beta}^*_{jm} r_{iM}(s)/\sqrt{m}$ and $\widehat{\alpha}^*_{jm} = O_P(1), \widehat{\beta}^*_{jm} = O_P(1)$, for all $j = 1, \dots, d$ (see (4.6) and (4.7), respectively), we obtain, due to the monotonicity of the ψ_j's, that $\boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3 = O_p(q(m)m^{-1/4})$. Combining this with the corresponding estimates for $-q < k < 0$, we get

$$\widehat{\boldsymbol{\Sigma}}_m - \widetilde{\boldsymbol{\Sigma}}_m = O_P\big(q(m)m^{-1/4}\big) \quad (m \to \infty),$$

i.e. (4.10).

It remains to estimate $\widetilde{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma}_m$.

First, notice that, for $k \ge 0$,

$$\widetilde{\boldsymbol{\Gamma}}_k - \boldsymbol{\Gamma}_k = \frac{1}{m}\sum_{i=1}^{m-k}[(\widetilde{\boldsymbol{z}}_i - \boldsymbol{z}_i)(\widetilde{\boldsymbol{z}}_{i+k} - \boldsymbol{z}_{i+k})^T + (\widetilde{\boldsymbol{z}}_i - \boldsymbol{z}_i)\boldsymbol{z}_{i+k}^T + \boldsymbol{z}_i(\widetilde{\boldsymbol{z}}_{i+k} - \boldsymbol{z}_{i+k})^T].$$

Further, for $i \in \mathbb{Z}, n \in \mathbb{N}$,

$$\begin{aligned}
\widetilde{\boldsymbol{z}}_i - \boldsymbol{z}_i &= \frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) - \sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu} r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))ds \\
&= \sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}[r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) - r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))]ds \\
&= \sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}[r_{iM}(s_\nu) - r_{iM}(s)]\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu))ds + \sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu} r_{iM}(s)[\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) - r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))]ds \\
&= \boldsymbol{u}_i + \boldsymbol{v}_i.
\end{aligned}$$

Thus,

$$\sum_{0\le k<q}\omega_q(k)(\widetilde{\boldsymbol{\Gamma}}_k - \boldsymbol{\Gamma}_k) = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3,$$

where

$$\boldsymbol{A}_1 = \sum_{0\le k<q}\omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}[\boldsymbol{u}_i\boldsymbol{u}_{i+k}^T + \boldsymbol{u}_i\boldsymbol{v}_{i+k}^T + \boldsymbol{v}_i\boldsymbol{u}_{i+k}^T + \boldsymbol{v}_i\boldsymbol{v}_{i+k}^T], \tag{4.11}$$

$$\boldsymbol{A}_2 = \sum_{0\le k<q}\omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}(\boldsymbol{u}_i + \boldsymbol{v}_i)\widetilde{\boldsymbol{z}}_{i+k}^T, \tag{4.12}$$

$$\boldsymbol{A}_3 = \sum_{0\le k<q}\omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}\boldsymbol{z}_i(\boldsymbol{u}_{i+k} + \boldsymbol{v}_{i+k})^T. \tag{4.13}$$

Since all the matrices appearing on the right-hand side of (4.11) are of the same type, we shall only treat one of them. Consider, for example, the (j,ℓ)-th element of the matrix $\boldsymbol{u}_i\boldsymbol{v}_{i+k}^T$. We have

$$u_{i,j}v_{i+k,\ell} = \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\int_{s_{\mu-1}}^{s_\mu}\int_{s_{\nu-1}}^{s_\nu}\Big([r_{iM}(s_\mu) - r_{iM}(s)]\psi_j(\varepsilon_{i,j}(s_\mu))r_{i+k,M}(t)[\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu)) - \psi_\ell(\varepsilon_{i+k,\ell}(t))]\Big)dsdt,$$

and from here, using the independence of $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot))\}$, the Cauchy-Schwarz inequality and stationarity,

$$\begin{aligned}
E\,|u_{i,j}v_{i+k,\ell}| \le\ & \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\int_{s_{\mu-1}}^{s_\mu}\int_{s_{\nu-1}}^{s_\nu}\Big(\|r_{iM}(s_\mu) - r_{iM}(s)\|_2\cdot\|r_{i+k,M}(t)\|_2 \\
&\times\|\psi_j(\varepsilon_{i,j}(s_\mu))\|_2\cdot\|\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu)) - \psi_\ell(\varepsilon_{i+k,\ell}(t))\|_2\Big)dsdt \\
\le\ & \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\int_{s_{\mu-1}}^{s_\mu}\int_{s_{\nu-1}}^{s_\nu}\Big(\sup_{s\in[0,1]}\|\psi_j(\varepsilon_{i,j}(s)\|_2\cdot\sup_{h\in[0,1/n]}\|r_{iM}(s_\mu - r_{iM}(s_\mu - h)\|_2 \\
&\times\sup_{t\in[0,1]}\|r_{i+k,M}(t)\|_2\cdot\sup_{h\in[0,1/n]}\|\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu) - \psi_\ell(\varepsilon_{i+k,\ell}(s_\nu - h))\|_2\Big)dsdt \\
=\ & \sup_{s\in[0,1]}\|\psi_j(\varepsilon_{0,j}(s))\|_2\cdot\sup_{t\in[0,1]}\|r_{0M}(t)\|_2\cdot\frac{1}{n}\sum_{\mu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\mu) - r_{0M}(s_\mu - h)\|_2 \\
&\times\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|\psi_\ell(\varepsilon_{0,\ell}(s_\nu) - \psi_\ell(\varepsilon_{0,\ell}(s_\nu - h))\|_2.
\end{aligned}$$

Now, using Assumptions (A.3), (B.1) and (B.7a)-(B.7b), together with the fact that ω_q is bounded and $q(m) = O(\log m)$, we can easily deduce that

$$\sum_{0\le k<q} \omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k} u_{i,j}v_{i+k,\ell} = o_p(1) \quad (m\to\infty).$$

The same result holds for all elements of the matrix $\boldsymbol{A}_1$. Concerning the matrices $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$, we can proceed in the same way. It suffices to write $\boldsymbol{z}_i = \sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu} r_{iM}(s)\boldsymbol{\psi}(\varepsilon_i(s))ds$ and make use of Assumptions (A.3) and (B.1) again together with either (B.7a) or (B.7b). Combining all the asymptotics above with the corresponding estimates for $-q<k<0$, we get

$$\widetilde{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma}_m = o_p(1) \quad (m\to\infty), \tag{4.14}$$

which together with (4.9) and (4.10) concludes the proof. □

5. Some auxiliary results

In the sequel, C and D denote generic positive constants, which may vary from case to case.

For the sake of brevity, we let $\{x_i(\cdot)\}$ denote any of the sequences $\{r_{iM}(\cdot) - Er_{iM}(\cdot)\}$, $\{\psi_j(\varepsilon_{i,j}(\cdot))\}$ or $\{r_{iM}(\cdot)\psi_j(\varepsilon_{i,j}(\cdot))\}$ and write $\{x_i^{(L)}(\cdot)\}$ for the corresponding counterparts of $\{r_{iM}^{(L)}(\cdot) - Er_{iM}^{(L)}(\cdot)\}$, $\{\psi_j(\varepsilon_{i,j}^{(L)}(\cdot))\}$ or $\{r_{iM}^{(L)}(\cdot)\psi_j(\varepsilon_{i,j}^{(L)}(\cdot))\}$, respectively.

Lemma 5.1. *Under the assumptions of Theorem 2.1, possibly extended to an $L_{2+\Delta}$-approximability condition in* (B.4) *and* (B.5) *(cf. Remark 2.5),*

(i) *there is a constant $C>0$ such that, for every $\ell\in\mathbb{Z}$, $K\in\mathbb{N}$, and $s\in[0,1]$,*

$$E\Big|\sum_{i=\ell+1}^{\ell+K} x_i(s)\Big|^p \le C\sup_{s\in[0,1]}\|x_0(s)\|_p^p\, K^{p/2}, \quad 2\le p\le 2+\Delta,$$

and, for $b_1\ge b_2\ge\ldots\ge b_K>0$,

$$E\max_{1\le k\le K}\Big|b_k\sum_{i=\ell+1}^{\ell+k} x_i(s)\Big|^2 \le C\sup_{s\in[0,1]}\|x_0(s)\|_2^2\,(\log K)^2\sum_{k=1}^{K} b_k^2, \tag{5.1}$$

$$E\max_{1\le k\le K}\Big|b_k\sum_{i=\ell+1}^{\ell+k} x_i(s)\Big|^p \le C\sup_{s\in[0,1]}\|x_0(s)\|_p^p\,\sum_{k=1}^{K} b_k^p\, k^{p/2-1}, \quad 2<p\le 2+\Delta; \tag{5.2}$$

(ii) *for some $D>0$ and all $m\in\mathbb{N}$, $s\in[0,1]$,*

$$E\Big(\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k} x_i(s)\Big|\Big)^2 \le D\sup_{s\in[0,1]}\|x_0(s)\|_2^2\,(\log m)^2, \tag{5.3}$$

$$E\Big(\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k} x_i(s)\Big|\Big)^p \le D\sup_{s\in[0,1]}\|x_0(s)\|_p^p, \quad 2<p\le 2+\Delta; \tag{5.4}$$

For the proof of (5.2) *and* (5.4)*, however, it is necessary to replace the L_2-approximability conditions in Assumptions* (B.4) *and* (B.5) *by a corresponding $L_{2+\Delta}$-approximability assumption, with some $\Delta>0$.*

(iii) *uniformly in $s\in[0,1]$ and for any $q_m\to\infty$,*

$$\max_{1\le i\le\lfloor m(T+1)\rfloor}|r_{iM}(s)| = O_P(m^{1/3}), \tag{5.5}$$

$$\sup_{s\in[0,1]} P\big(\max_{1\le i\le\lfloor m(T+1)\rfloor}|r_{iM}(s)|\ge q_m m^{1/3}\big)\to 0. \tag{5.6}$$

Proof. (i) Making use of the $L_{2+\Delta}$-approximability from Assumption (B.4) (with $\Delta\ge 0$, cf. Remarks 2.1 and 2.3), the first bound has been obtained in Berkes et al. [4], Proposition 4. Observe that, in our case,

$$\|x_i(s)\|_p = \|x_i^{(L)}(s)\|_p \le \sup_{s\in[0,1]}\|x_0(s)\|_p, \quad \text{for} \quad 2\le p\le 2+\Delta, \quad s\in[0,1].$$

Similarly, for the two other bounds confer, e.g., Kirch [13], Theorems B.1 and B.3, which are based on earlier results of Moricz [16] and Móricz et al. [17] in combination with Fazekas and Klesov [8].

Note that the sequence $\{r_{iM}(s)\psi_j(\varepsilon_{i,j}(s))\}$ also satisfies the $L_{2+\Delta}$-approximability condition, uniformly in $s \in [0,1]$, since

$$\begin{aligned}
&\|r_{iM}(s)\psi_j(\varepsilon_{i,j}(s)) - r_{iM}^{(L)}(s)\psi_j(\varepsilon_{i,j}^{(L)}(s))\|_{2+\Delta} \\
&\leq \|(r_{iM}(s) - r_{iM}^{(L)}(s))\psi_j(\varepsilon_{i,j}(s))\|_{2+\Delta} + \|r_{iM}^{(L)}(s)(\psi_j(\varepsilon_{i,j}(s)) - \psi_j(\varepsilon_{i,j}^{(L)}(s)))\|_{2+\Delta} \\
&\leq \sup_{s\in[0,1]} \|r_{iM}(s) - r_{iM}^{(L)}(s)\|_{2+\Delta} \sup_{s\in[0,1]} \|\psi_j(\varepsilon_{i,j}(s))\|_{2+\Delta} \\
&\quad + \sup_{s\in[0,1]} \|r_{iM}^{(L)}(s)\|_{2+\Delta} \sup_{s\in[0,1]} \|\psi_j(\varepsilon_{i,j}(s)) - \psi_j(\varepsilon_{i,j}^{(L)}(s))\|_{2+\Delta},
\end{aligned}$$

where, for the second inequality, we have used the independence of the sequences $\{r_{iM}(s)\}$ and $\{\psi_j(\varepsilon_{i,j}(s))\}$.
(ii) It follows immediately from the fact that the sequence $\{x_i(s)\}$, $s \in [0,1]$ fixed, satisfies Assumptions (B.1) and (B.4) together with the estimates in (5.1) and (5.2).
(iii) By (i),

$$E\Big|\sum_{i=\ell+1}^{\ell+K}(r_{iM}(s) - Er_{iM}(s))\Big|^3 \leq C \sup_{s\in[0,1]} ||r_{0M}(s) - Er_{0M}(s)||_3^3 \; K^{3/2}, \quad s \in [0,1].$$

We also have, for $s \in [0,1]$,

$$\max_{1\leq i\leq \lfloor m(T+1)\rfloor} |r_{iM}(s)| \leq \max_{1\leq i\leq \lfloor m(T+1)\rfloor} |r_{iM}(s) - Er_{iM}(s)| + \max_{1\leq i\leq \lfloor m(T+1)\rfloor} |Er_{iM}(s)|, \qquad \text{and}$$

$$\max_{1\leq i\leq \lfloor m(T+1)\rfloor} |r_{iM}(s) - Er_{iM}(s)| \leq D\Big(\frac{1}{\lfloor m(T+1)\rfloor}\sum_{i=1}^{\lfloor m(T+1)\rfloor} |r_{iM}(s) - Er_{iM}(s)|^3\Big)^{1/3} \lfloor m(T+1)\rfloor^{1/3}.$$

Since, by our assumptions, for fixed $s \in [0,1]$, $\{r_{iM}(s) - Er_{iM}(s)\}$ is a stationary and ergodic sequence and $\sup_{s\in[0,1]} E|r_{iM}(s) - Er_{iM}(s)|^3 < \infty$, the ergodic theorem implies, as $m \to \infty$,

$$\frac{1}{\lfloor m(T+1)\rfloor}\sum_{i=1}^{\lfloor m(T+1)\rfloor} |(r_{iM}(s) - Er_{iM}(s))|^3 \to E|(r_{0M}(s) - Er_{0M}(s))|^3 \leq \sup_{s\in[0,1]} E|r_{0M}(s) - Er_{0M}(s)|^3 < \infty.$$

Combining all this we get (5.5), which immediately implies (5.6). □

In the following E^* and var^* denote the conditional expectation and conditional variance given $r_{iM}(\cdot)$, $i = 1,\ldots,m; m+1,\ldots,\lfloor mT\rfloor$. We omit the index j, i.e., we write $\varepsilon_i(s), \psi, \ldots$ instead of $\varepsilon_{i,j}(s), \psi_j(s), \ldots$.

Lemma 5.2. *Let the assumptions of Theorem 2.1 be satisfied. Then, as $m \to \infty$,*

$$\sup_{|a|+|b|\leq C} |Z_m(a,b) - E^*Z_m(a,b)| = O_P\big(m^{-\eta}\big),$$

$$E^*Z_m(a,b) = \frac{1}{n}\sum_{\nu=1}^{n}\frac{\lambda'(0,s_\nu)}{2}\frac{1}{m}\sum_{i=1}^{m}\big(a + br_{iM}(s_\nu)\big)^2 + O_P\big(m^{-\eta}(|a|^3 + |b|^3)\big),$$

and

$$\sup_{|a|+|b|\leq C}\Big|Z_m(a,b) - \frac{1}{n}\sum_{\nu=1}^{n}\frac{\lambda'(0,s_\nu)}{2}\frac{1}{m}\sum_{i=1}^{m}\big(a + br_{iM}(s_\nu)\big)^2\Big| = O_P(m^{-\eta}),$$

for some $\eta > 0$, where

$$Z_m(a,b) = \frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m}\big(\rho(\varepsilon_i(s_\nu) - a/\sqrt{m} - br_{iM}(s_\nu)/\sqrt{m}) - \rho(\varepsilon_i(s_\nu)) + (a/\sqrt{m} + br_{iM}(s_\nu)/\sqrt{m})\psi(\varepsilon_i(s_\nu))\big).$$

Proof. The lines of the proof are quite standard. We just need to derive a proper approximation for the conditional expectation and variance of $Z_m(a,b)$.

Whenever convenient we use the short-hand notations

$$\begin{aligned}
&d_i(s_\nu) = a + br_{iM}(s_\nu) \quad \text{and} \\
&g(\varepsilon_i(s_\nu), x, d_i(s_\nu)) = \operatorname{sign} d_i(s_\nu)\big(-\psi(\varepsilon_i(s_\nu) - x \operatorname{sign} d_i(s_\nu)) + \psi(\varepsilon_i(s_\nu))\big), \quad i \in \mathbb{Z}.
\end{aligned}$$

Note that, for any d,

$$\rho(\varepsilon_i - d) - \rho(\varepsilon_i) + d\psi(\varepsilon_i) = \operatorname{sign} d\int_0^{|d|}\big(-\psi(\varepsilon_i - x\operatorname{sign} d) + \psi(\varepsilon_i)\big)dx \geq 0, \quad i \in \mathbb{Z}.$$

Direct calculations in combination with Lemma 5.1 result in

$$
\begin{aligned}
E^* Z_m(a,b) = &\ \frac{1}{n}\sum_{\nu=1}^{n} E^* \sum_{i=1}^{m} \Big(\int_0^{|d_i(s_\nu)|/\sqrt{m}} g(\varepsilon_i(s_\nu), x, d_i(s_\nu)) dx \\
= &\ \frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m} \lambda'(0,s_\nu) d_i^2(s_\nu)\frac{1}{2m} + O_P\Big(\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m} |d_i(s_\nu)|^3 \frac{1}{m^{3/2}}\Big) \\
= &\ \sum_{\nu=1}^{n} \frac{1}{2}\lambda'(0,s_\nu)\big(a + 2ab\frac{1}{m}\sum_{i=1}^{m} r_{iM}(s_\nu) + b^2\frac{1}{m}\sum_{i=1}^{m} r_{iM}^2(s_\nu)\big) + O_P\big(m^{-\eta}(|a|^3+|b|^3)\big),
\end{aligned}
$$

for some $\eta > 0$ and uniformly in $|a| + |b| \le C$.

For the conditional variance we obtain

$$
\begin{aligned}
& var^*\big\{Z_m(a,b)\big\} = E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m}\int_0^{|d_i(s_\nu)|/\sqrt{m}} \big(g(\varepsilon_i(s_\nu), x, d_i(s_\nu)) - E^* g(\varepsilon_i(s_\nu), x, d_i(s_\nu))\big)dx\Big)^2 \\
& = \sum_{i_1=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\int_0^{|d_{i_1}(s_\nu)|/\sqrt{m}} \big(g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu)) - E^* g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu))\big)dx\Big)^2 \\
& \quad + 2E^* \sum_{1\le i_1 < i_2 \le m} \frac{1}{n}\sum_{\nu_1=1}^{n} \Big\{\Big(\int_0^{|d_{i_1}(s_{\nu_1}|/\sqrt{m}} \big(g(\varepsilon_{i_1}(s_{\nu_1}), x, d_{i_1}(s_{\nu_1})) - E^* g(\varepsilon_{i_1}(s_{\nu_1}), x, d_{i_1}(s_{\nu_1}))\big)dx\Big) \\
& \quad \times \Big(\frac{1}{n}\sum_{\nu_2=1}^{n}\int_0^{|d_{i_2}(s_{\nu_2})|/\sqrt{m}} \big(g(\varepsilon_{i_2}(s_{\nu_2}), y, d_{i_2}(s_{\nu_2})) - E^* g(\varepsilon_{i_2}(s_{\nu_2}), y, d_{i_2}(s_{\nu_2}))\big)dy\Big)\Big\} \\
& = I_1 + I_2 \quad \text{(say)}.
\end{aligned}
$$

Using Assumption (A.3) together with the Cauchy-Schwarz inequality, we get

$$
\begin{aligned}
I_1 = & \sum_{i_1=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\int_0^{|d_{i_1}(s_\nu)|/\sqrt{m}} \big(g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu)) - E^* g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu))\big)dx\Big)^2 \\
& \le D\sum_{i_1=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\int_0^{|d_{i_1}(s_\nu)|/\sqrt{m}} g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu))dx\Big)^2 \\
& \le D\sum_{i_1=1}^{m}\frac{1}{n}\sum_{\nu_1=1}^{n}\frac{1}{n}\sum_{\nu_2=1}^{n} E^*\Big[\Big(\Big|\int_0^{|d_{i_1}(s_{\nu_1})|/\sqrt{m}} g(\varepsilon_{i_1}(s_{\nu_1}), x, d_{i_1}(s_{\nu_1})dx\Big|\Big) \\
& \quad \times \Big(\Big|\int_0^{|d_{i_1}(s_{\nu_2})|/\sqrt{m}} g(\varepsilon_{i_1}(s_{\nu_2}), z, d_{i_1}(s_{\nu_2}))dz\Big|\Big)\Big] \\
& \le D\sum_{i_1=1}^{m}\frac{1}{n}\sum_{\nu_1=1}^{n}\frac{1}{n}\sum_{\nu_2=1}^{n}\Big[\Big(\big(|d_{i_1}(s_{\nu_1})|/\sqrt{m}\big)^3\Big(\big(|d_{i_1}(s_{\nu_2})|/\sqrt{m}\big)^3\Big)\Big]^{1/2} \\
& = D\big(\frac{1}{m}\big)^{3/2}\Big(|a|^3 m + |b|^3 \sum_{i_1=1}^{m}\Big(\frac{1}{n}\sum_{\nu_1=1}^{n}|r_{i_1M}(s_{\nu_1})|^{3/2}\Big)^2 O_P((|a|^3+|b|^3)m^{-\eta}),
\end{aligned}
$$

uniformly in $|a| + |b| \le C$.

Concerning I_2 we have, due to the independence of $\{r_{iM}(s)\}$ and $\{\boldsymbol{\varepsilon}_i(s)\}$,

$$I_2 \le 2\sum_{i_1=1}^{m-1}\sum_{i_2=1}^{m-i_1}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{n}\sum_{r_1=1}^{n}\int_0^{|d_{i_1}(s_{r_1})|/\sqrt{m}}\int_0^{|d_{i_1+i_2}(s_{r_2})|/\sqrt{m}}$$

$$\Big(E^*(g(\varepsilon_{i_1}(s_{r_1}),x,d_{i_1}(s_{r_1})))^2\big(E^*(-\psi(\varepsilon_{i_1+i_2}(s_{r_2})-y)+\psi(\varepsilon_{i_1+i_2}^{(i_2)}(s_{r_2})-y))^2$$

$$+E^*(-\psi(\varepsilon_{i_1+i_2}(s_{r_2}))+\psi(\varepsilon_{i_1+i_2}^{(i_2)}(s_{r_2})))^2\big)\Big)^{1/2}dxdy$$

$$\le D\sum_{i_1=1}^{m-1}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{n}\sum_{r_1=1}^{n}|d_{i_1}(s_{r_1})/\sqrt{m}|^{1/2+1}$$

$$\times\sum_{i_2=1}^{m-i_1}|d_{i_1+i_2}(s_{r_2})/\sqrt{m}|\sup_{|a|\le a_0}\Big(E^*\big(\psi(\varepsilon_{i_1+i_2}(s_{r_2})-a)-\psi(\varepsilon_{i_1+i_2}^{(i_2)}(s_{r_2})-a)\big)^2\Big)^{1/2}$$

$$\le D\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{m^{3/2}}\sum_{i_1=1}^{m-1}|d_{i_1}(s_{r_1}))|^{3/2}$$

$$\times\sup_{|a|\le a_0}\sum_{i_2=1}^{m-i_1}|d_{i_1+i_2}(s_{r_2})|\Big(E\big(\psi(\varepsilon_0(s_{r_2}))-a)-\psi(\varepsilon_0(s_{r_2}))^{(i_2)}-a)\big)^2\Big)^{1/2}$$

$$\le O_P(m^{-\eta}),$$

where we used the fact

$$E|d_{i_1+i_2}(s_{r_1})|^{3/2}\cdot|d_{i_1}(s_{r_2})|\le\Big(E|d_1(s_{r_1})|^3Ed_1^2(s_{r_2})\Big)^{1/2}.$$

On combining the above estimates for $E^*Z_m(a,b),I_1,I_2$, we conclude that Lemma 5.2 holds true. □

Lemma 5.3. *Let the assumptions of Theorem 2.1 be satisfied. Then, as* $m\to\infty$,

$$\sup_{|a|+|b|\le C}|\boldsymbol{M}_m(a,b)-E^*\boldsymbol{M}_m(a,b)|=O_P(m^{-\eta}),$$

$$E^*\boldsymbol{M}_m(a,b)=-\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(a+b\frac{1}{m}\sum_{i=1}^{m}r_{iM}(s_\nu),a\frac{1}{m}\sum_{i=1}^{m}r_{iM}(s_\nu)+b\frac{1}{m}\sum_{i=1}^{m}r_{iM}^2(s_\nu)\Big)^T+O_P\big(m^{-\eta}\big),$$

and

$$\sup_{|a|+|b|\le C}\Big|\boldsymbol{M}_m(a,b)+\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\lambda'(0,s_\nu)\Big(a\,m+b\sum_{i=1}^{m}r_{iM}(s_\nu),a\sum_{i=1}^{m}r_{iM}(s_\nu)+b\sum_{i=1}^{m}r_{iM}^2(s_\nu)\Big)^T\Big|=O_P\big(m^{-\eta}\big),$$

with some $\eta>0$, *where*

$$\boldsymbol{M}_m(a,b)=\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\big(1,r_{iM}(s_\nu)\big)^T\big(\psi(\varepsilon_i(s_\nu)-(a+br_{iM}(s_\nu))/\sqrt{m})-\psi(\varepsilon_i(s_\nu))\big).$$

Proof. Again one has to get suitable approximations for the conditional expectation $\boldsymbol{M}_m(a,b)$ and the conditional (2×2)-variance matrix

$$var^*\{\boldsymbol{M}_n(a,b)\}=E^*\big(\boldsymbol{M}_n(a,b)-E^*\boldsymbol{M}_n(a,b)\big)\big(\boldsymbol{M}_n(a,b)-E^*\boldsymbol{M}_n(a,b)\big)^T.$$

We start with the conditional expectation

$$E^*\boldsymbol{M}_m^T(a,b)=\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\big(1,r_{iM}(s_\nu)\big)\big(-\lambda(d_i(s_\nu)/\sqrt{m},s_\nu)\big)$$

$$=-\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\lambda'(0,s_\nu)\sum_{i=1}^{m}\big(1,r_{iM}(s_\nu)\big)d_i(s_\nu)$$

$$+O_P\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=1}^{m}\big(1,|r_{iM}(s_\nu)|\big)|d_i^2(s_\nu)|\Big)$$

$$=-\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\lambda'(0,s_\nu)\sum_{i=1}^{m}\big(a+br_{iM}(s_\nu),ar_{iM}(s_\nu)+br_{iM}^2(s_\nu)\big)$$

$$+O_P\Big(a^2m^{-1/2}+b^2\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=1}^{m}|r_{iM}(s_\nu)|^3\Big)=O_P\big((a^2+b^2)m^{-\eta}\big),$$

uniformly in $|a|+|b|\le C$, where the rates above are to be understood componentwise.

For the conditional variance matrix we only calculate one term. The calculation of the others is similar and will therefore be omitted. We have

$$
\begin{aligned}
& var^*\Big\{\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m} r_{iM}(s_\nu)\big(\psi_j(\varepsilon_{i,j}(s_\nu)-d_i(s_\nu)/\sqrt{m})-\psi_j(\varepsilon_{i,j}(s_\nu))\big)\Big\}\\
&= \frac{1}{m}\sum_{i=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\big(\psi(\varepsilon_i(s_\nu)-d_i(s_\nu)/\sqrt{m})-\psi(\varepsilon_i(s_\nu))+\lambda_j(d_i(s_\nu)/\sqrt{m},s_\nu)\big)\Big)^2\\
&\quad + 2\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m-i} r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})\\
&\quad \times E^*\big(\psi_j(\varepsilon_{i,j}(s_{r_1})-d_i(s_{r_1})/\sqrt{m})-\psi_j(\varepsilon_{i,j}(s_{r_1}))+\lambda_j(d_i(s_{r_1})/\sqrt{m},s_{r_1})\big)\big)\\
&\quad \times \big(\psi(\varepsilon_{i+j}(s_{r_2})-d_{i+j}(s_{r_2})/\sqrt{m})-\psi_j(\varepsilon_{i+j}(s_{r_2}))+\lambda_j(d_{i+j}(s_{r_2})/\sqrt{m},s_{r_2})\big)\\
&= J_1+2J_2 \quad \text{(say)}.
\end{aligned}
$$

In view of Assumption (A.2), a similar estimate as that for I_1 in the proof of Lemma 5.2 gives

$$
\begin{aligned}
J_1 &= \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{n}\sum_{r=1}^{n} r_{iM}(s_\nu)r_{iM}(s_r)\\
&\quad \times E^*\big(\psi(\varepsilon_i(s_\nu)-d_i(s_\nu)/\sqrt{m})-\psi(\varepsilon_i(s_\nu))+\lambda(d_i(s_\nu)/\sqrt{m},s_\nu)\big)\\
&\quad \times \big(\psi(\varepsilon_i(s_r)-d_i(s_r)/\sqrt{m})-\psi(\varepsilon_i(s_r))+\lambda(d_i(s_r)/\sqrt{m},s_r)\big)\\
&\le D\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{n}\sum_{r=1}^{n}|r_{iM}(s_\nu)||r_{iM}(s_r)|\\
&\quad \times \Big(E^*\big(\psi(\varepsilon_i(s_\nu)-d_i(s_\nu)/\sqrt{m})-\psi(\varepsilon_i(s_\nu)\big)^2 E^*\big(\psi(\varepsilon_i(s_r)-d_i(s_r)/\sqrt{m})-\psi(\varepsilon_i(s_r)\big)^2\Big)^{1/2}\\
&\le D\frac{1}{m}\sum_{i=1}^{m}\Big(\frac{1}{n}\sum_{\nu=1}^{n}|r_{iM}(s_\nu)||d_i(s_\nu)/\sqrt{m}|^{1/2}\Big)^2\\
&\le D\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\sum_{i=1}^{m}|r_{iM}(s_\nu)|^2||d_i(s_\nu)/\sqrt{m})| = O_P(m^{-\eta}).
\end{aligned}
$$

Concerning J_2 we obtain

$$
\begin{aligned}
J_2 &= \frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m-i} r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})\\
&\quad \times E^*\big(\psi(\varepsilon_i(s_{r_1})-d_i(s_{r_1})/\sqrt{m})-\psi(\varepsilon_i(s_{r_1}))+\lambda(d_i(s_{r_1})/\sqrt{m},s_{r_1})\big)\\
&\quad \times \big(\psi(\varepsilon_{i+j}(s_{r_2})-d_{i+j}(s_{r_2})/\sqrt{m})-\psi(\varepsilon_{i+j}(s_{r_2}))-(\psi(\varepsilon_{i+j}^{(j)}(s_{r_2})-d_{i+j}(s_{r_2})/\sqrt{m})-\psi(\varepsilon_{i+j}^{(j)}(s_{r_2}))))\big),
\end{aligned}
$$

and, uniformly in $|a|+|b|\le C$,

$$
\begin{aligned}
|J_2| &\le D\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m-i}|r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})|(|d_i(s_{r_1})|/\sqrt{m})^{1/2}\\
&\quad \times \big\{E^*\big(\psi(\varepsilon_{i+j}(s_{r_2})-d_{i+j}(s_{r_2})/\sqrt{m})-\psi(\varepsilon_{i+j}(s_{r_2}))\\
&\quad -(\psi(\varepsilon_{i+j}^{(j)}(s_{r_2})-d_{i+j}(s_{r_2})/\sqrt{m})-\psi(\varepsilon_{i+j}^{(j)}(s_{r_2})))\big)^2\big\}^{1/2}\\
&\le D\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m^{3/2}}\sum_{i=1}^{m}\sum_{j=1}^{m-i}\big(a^{1/2}+b^{1/2}|r_{iM}(s_{r_1}|\big)|r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})|\\
&\quad \times \sup_{|a|\le a_0}\big(E(\psi(\varepsilon_0(s_{r_2})-a)-\psi(\varepsilon_0^{(j)}(s_{r_2})-a))^2\big)^{1/2}\\
&= O_P\big((a^{1/2}+b^{1/2})m^{-\eta}\big).
\end{aligned}
$$

Now, a similar estimate as that for I_2 in the proof of Lemma 5.2 gives

$$
\sup_{|a|+|b|\le C}|J_2| = O_P\big(m^{-\eta}\big),
$$

with some $\eta>0$, so that altogether we have

$$
\sup_{|a|+|b|\le C} var^*\Big\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m} r_{iM}(s_r)\big(\psi(\varepsilon_i(s_r)-d_i(s_r)/\sqrt{m})-\psi(\varepsilon_i(s_r))\big)\Big\} = O_P\big(m^{-\eta}\big),
$$

for some $\eta>0$. □

Remark 5.1. Inserting $\widehat{\alpha}^*_{j,m}$ and $\widehat{\beta}^*_{j,m}$ (as defined in (4.3)) for a, b into the assertion of Lemma 5.3 and omitting the index j for the sake of smplicity, we receive

$$\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\big(1, r_{iM}(s_\nu)\big)^T\big(\psi(\varepsilon_i(s_\nu)-(\widehat{\alpha}^*_m+\widehat{\beta}^*_m r_{iM}(s_\nu))/\sqrt{m})-\psi(\varepsilon_i(s_\nu))\big)$$
$$+\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(\widehat{\alpha}_m+\widehat{\beta}^*_m\frac{1}{m}\sum_{i=1}^{m}r_{iM}(s_\nu),\widehat{\alpha}^*_m\frac{1}{m}\sum_{i=1}^{m}r_{iM}(s_\nu)+\widehat{\beta}_m\frac{1}{m}\sum_{i=1}^{m}r^2_{iM}(s_\nu)\Big)^T=O_P\big(m^{-\eta}\big).$$

Due to the definition of $\widehat{\alpha}^*_m$ and $\widehat{\beta}^*_m$ and by our assumptions, we have the following asymptotic representation:

$$\widehat{\alpha}^*_m=\frac{1}{\int_0^1\lambda'(0,z)dz}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1\psi(\varepsilon_i(s))ds-\widehat{\beta}^*_m\frac{\int_0^1\lambda'(0,z)Er^2_{0M}(z)dz}{\int_0^1\lambda'(0,z)dz}+o_P(m^{-\eta}),$$

$$\widehat{\beta}^*_m=\frac{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1\psi(\varepsilon_i(s))\Big(r_{iM}(s)-\frac{\int_0^1\lambda'(0,z)Er_{0M}(z)dz}{\int_0^1\lambda'(0,z)dz}\Big)ds}{\int_0^1\lambda'(0,z)Er^2_{0M}(z)dz-\frac{\left(\int_0^1\lambda'(0,z)Er_{0M}(z)dz\right)^2}{\int_0^1\lambda'(0,z)dz}}+o_P(m^{-\eta}).$$

The last two relations are important for getting the limit distribution of our test procedure.

The next lemma follows along the arguments of Lemma 5.4 in Chochola et al. [7], modified along the lines of the proofs of the previous lemmas. So the proof will only be sketched and not be given in detail.

Lemma 5.4. *Let the assumptions of Theorem 2.1 be satisfied. Then, for any $T > 0$, as $m \to \infty$,*

$$\max_{1\le k\le\lfloor mT\rfloor}\left(\frac{|\{N_{k,m}(a,b)-E^*N_{km}(a,b)\}_{a=\widehat{\alpha}^*_m,b=\widehat{\beta}^*_m}|}{(k/m)^\gamma}\right)=O_P(m^{-\eta}),$$

for some $\eta > 0$, where $\widehat{\alpha}^*_m, \widehat{\beta}^*_m$ are as in (4.3), and

$$N_{k,m}(a,b)=\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}r_{iM}(s_\nu)\big(\psi(\varepsilon_i(s_\nu)-a/\sqrt{m}-br_{iM}(s_\nu)/\sqrt{m})-\psi(\varepsilon_i(s_\nu))\big).$$

Proof. Lemma 5.4 is related to Lemma 5.3, but it is somewhat more complicated.
Direct calculations give

$$\begin{aligned}E^*N_{k,m}(a,b)=&-\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}r_{iM}(s_\nu)\lambda\big((a+br_{iM}(s_\nu),s_\nu)/\sqrt{m}\big)\\=&-\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\frac{1}{m}\Big(a\sum_{i=m+1}^{m+k}r_{iM}+b\sum_{i=m+1}^{m+k}r^2_{iM}(s_\nu)\Big)\\&+O_P\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}|r_{iM}(s_\nu)||(a+br_{iM}(s_\nu))/\sqrt{m}|^2\Big),\end{aligned}$$

uniformly for $|a| + |b| \le C$, with some $\eta > 0$. In fact we need to study more carefully the properly standardized remainder

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{(k/m)^\gamma}\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=m+1}^{m+k}|r_{iM}(s_\nu)|\Big)+\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{(k/m)^\gamma}\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=m+1}^{m+k}|r_{iM}(s_\nu)|^3\Big).$$

Both terms above are $O_P(m^{-\eta})$ for some $\eta > 0$.

Next, we try to get an upper bound for $var^*\{N_{k,m}(a, b)\}$. We have

$$\begin{aligned}var^*\{N_{k,m}(a,b)\}=&\ \frac{1}{m}\sum_{i=m+1}^{m+k}E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\\&\times\big(\psi(\varepsilon_i(s_\nu)-d_i(s_\nu)/\sqrt{m})-\psi(\varepsilon_i(s_\nu))-E^*\psi(\varepsilon_i(s_\nu)-d_i(s_\nu)/\sqrt{m})\big)\Big)^2\\&+2\frac{1}{m}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{r_1=1}^{n}r_{iM}(s_{r_1})\,E^*\Big(\psi(\varepsilon_i(s_{r_1})-d_i(s_{r_1})/\sqrt{m})-\psi(\varepsilon_i(s_{r_1}))-E^*\psi(\varepsilon_i(s_{r_1})-d_i(s_{r_1})/\sqrt{m})\Big)\\&\times\Big(\sum_{\nu=1}^{m+k-i}\frac{1}{n}\sum_{r_2=1}^{n}r_{i+\nu,M}(s_{r_2})\\&\times\big(\psi(\varepsilon^{(\nu)}_{i+\nu}(s_{r_2})-d_{i+\nu}(s_{r_2})/\sqrt{m})-\psi(\varepsilon^{(\nu)}_{i+\nu}(s_{r_2}))-E^*\psi(\varepsilon^{(\nu)}_{i+\nu}(s_{r_2})-d_{i+\nu}(s_{r_2})/\sqrt{m})\big)\Big)\\=&\ L_{1,k}+2L_{2,k}\quad\text{(say)},\end{aligned}$$

and, along the lines of the proof of Lemma 5.3 (see the estimation of the terms J_1, J_2 there), we get

$$L_{1,k} = \frac{k}{m} m^{-1/2}\big(|a| + |b|\big)O_P(1),$$
$$|L_{2,k}| = \frac{1}{m^{1+1/2}}(|a|^{1/2}k + |b|^{1/2}k)O_P(1),$$

uniformly in $|a| + |b| \leq C$ and in $1 \leq k \leq \lfloor mT \rfloor$. So, altogether we have

$$var^*\{N_{k,m}(a,b)\} = \frac{k}{m} m^{-\eta}(|a| + |b|)O_P(1),$$

uniformly in $|a| + |b| \leq C$ and in $1 \leq k \leq \lfloor mT \rfloor$, with some $\eta > 0$.

Quite similarly we get, for $\ell = 1, 2, \ldots$,

$$var^*\{N_{k+\ell,m}(a,b) - N_{k,m}(a,b)\} = \frac{\ell}{m} m^{-\eta}(|a| + |b|)O_P(1).$$

Then, on applying Theorem B.4 of Kirch [13],

$$m^{-1+2\gamma} E^* \max_{1\leq k\leq \lfloor mT\rfloor} \Big(\frac{1}{k^\gamma}|N_{m,k}(a,b) - E^* N_{m,k}(a,b)|\Big)^2$$
$$= m^{-1+2\gamma}(\log m)^2 \sum_{k=1}^{\lfloor mT \rfloor} \frac{1}{k^{2\gamma}} m^{-\eta}(|a| + |b|)O_P(1) \;=\; (\log m)^2 m^{-\eta}(|a| + |b|)O_P(1).$$

We need to replace a, b by the estimators $\widehat{\alpha}_m^*, \widehat{\beta}_m^*$. However our $N_{k,m}(a,b)$ depends on $\varepsilon_1(\cdot), \ldots, \varepsilon_m(\cdot)$. Therefore we try to replace $N_{k,m}(a,b)$ by something that is asymptotically equivalent, but does not depend on $\varepsilon_1(\cdot), \ldots, \varepsilon_m(\cdot)$.

Toward this note that

$$N_{k,m}^{(m)}(a,b) = \frac{1}{\sqrt{m}} \sum_{i=1}^{k} \frac{1}{n} \sum_{\nu=1}^{n} r_{i+m,M}(s_\nu)\big(\psi(\varepsilon_{i+m}^{(i)}(s_\nu) - d_i(s_\nu)/\sqrt{m}) - \psi(\varepsilon_{m+i}^{(i)}(s_\nu))\big)$$

has all the properties of $N_{k,m}(a,b)$ above, but it is independent of $\varepsilon_1(\cdot), \ldots, \varepsilon_m(\cdot)$. This together with the consistency of $\widehat{\alpha}_m^*$ and $\widehat{\beta}_m^*$ implies

$$m^{-1+2\gamma} \max_{1\leq k\leq \lfloor mT\rfloor} \Big(\frac{1}{k^\gamma}\Big|\{N_{m,k}^{(m)}(a,b) - E^*(N_{m,k}^{(m)}(a,b)|\}_{a=\widehat{\alpha}_m^*, b=\widehat{\beta}_m^*)}\Big|\Big)^2$$
$$= O_P\big((\log m)^2 m^{-\eta} \max(|\widehat{\alpha}_m^*| + |\widehat{\beta}_m^*|, |\widehat{\alpha}_m^*|^{1/2} + |\widehat{\beta}_m^*|^{1/2})\big) = O_P\big((\log m)^2 m^{-\eta}\big).$$

It is still necessary to show the closeness of $N_{k,m}(a,b)$ and $N_{k,m}^{(m)}(a,b)$. Clearly, $N_{k,m}^{(m)}(a,b)$ is independent of $\varepsilon_1(\cdot), \ldots, \varepsilon_m(\cdot)$ and

$$E^*\big(N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b)\big) = 0,$$

$$N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b) = \frac{1}{\sqrt{m}} \frac{1}{n} \sum_{\nu=1}^{n} \sum_{i=1}^{k} (r_{i+m,M}(s_\nu)$$
$$\times \Big(\big(\psi(\varepsilon_{i+m}(s_\nu) - \frac{d_{i+m}(s_\nu)}{\sqrt{m}}) - \psi(\varepsilon_{i+m}^{(i)}(s_\nu) - \frac{d_{i+m}(s_\nu)}{\sqrt{m}})\big)$$
$$- \big(\psi(\varepsilon_{i+m}(s_\nu)) - \psi(\varepsilon_{i+m}^{(i)}(s_\nu))\big)\Big),$$

$$E^*|N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b)| \leq \frac{D}{\sqrt{m}} \frac{1}{n} \sum_{\nu=1}^{n} \sum_{i=1}^{k} |r_{i+m,M}(s_\nu)|$$
$$\times \sup_{|a|\leq a_0} E\big(|\psi(\varepsilon_0(s_\nu) - a) - \psi(\varepsilon_0^{(i)}(s_\nu) - a)| + |\psi(\varepsilon_0)(s_\nu) - \psi(\varepsilon_0^{(i)}(s_\nu))|\big)$$
$$\leq \frac{D}{\sqrt{m}} \frac{1}{n} \sum_{\nu=1}^{n} \sum_{i=1}^{\lfloor mT \rfloor} |r_{i+m,M}(s_\nu)| \sup_{|a|\leq a_0} E|\psi(\varepsilon_0(s_\nu) - a) - \psi(\varepsilon_0^{(i)}(s_\nu) - a)|,$$

which holds for any $1 \leq k \leq \lfloor mT \rfloor$. So, in view of our assumptions,

$$\sup_{|a|+|b|\leq C} E^*|N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b,j)| = O_P(m^{-1/2}),$$

whence

$$\sup_{1\leq k\leq \lfloor mT\rfloor} \left(\frac{\sup_{|a|+|b|\leq C} E^*|N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b)|}{(k/m)^\gamma}\right) = O_P(m^{-\eta}),$$

for some $\eta > 0$. A combination of the above estimates completes the proof of Lemma 5.4. □

Remark 5.2. From Lemma 5.4 we get the following approximations:

$$\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\frac{1}{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu)-\widehat{\alpha}_m^*/\sqrt{m}-\widehat{\beta}_m^* r_{iM}(s_\nu)/\sqrt{m})$$
$$=\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))-\Big(\widehat{\alpha}_m\frac{1}{m}\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}(s_\nu)$$
$$+\widehat{\beta}_m^*\frac{1}{m}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}^2(s_\nu)\Big)+O_P(m^{-\eta}).$$

In view of Remark 5.1, we similarly get

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))$$
$$=\widehat{\alpha}_m^*\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}(s_\nu)+\widehat{\beta}_m^*\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}^2(s_\nu))+O_P\big(m^{-\eta}\big).$$

After some standard steps this results in

$$\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu)-\widehat{\alpha}_m^*/\sqrt{m}-\widehat{\beta}_m^* r_{iM}(s_\nu)/\sqrt{m})$$
$$=\frac{1}{\sqrt{m}}\Big(\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))-\frac{k}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))+O_P\big(m^{-\eta}\big)\Big).$$

Here we also used that

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m(k/m)^\gamma}\Big|\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(\frac{k}{m}\sum_{i=1}^{m}r_{iM}(s_\nu)-\sum_{i=m+1}^{m+k}r_{iM}(s_\nu)\Big)$$
$$+\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(\frac{k}{m}\sum_{i=1}^{m}r_{iM}^2(s_\nu)-\sum_{i=m+1}^{m+k}r_{iM}^2(s_\nu)\Big)\Big|=o_P(1)$$

as $m\to\infty$ (cf. Lemma 5.5 (iii)).

Lemma 5.5. *Let Assumptions* (B.1), (B.4), *and* (B.6)-(B.7) *be satisfied. Then,*

(i) *there is a constant $C>0$ such that*

$$E\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi_j(\varepsilon_{i,j}(s))ds\Big|$$
$$\le C\,(\log m)\,\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\nu)-r_{0M}(s_\nu-h)\|_2; \tag{5.7}$$

(ii) *for $j=1,\dots,d$, as $m\to\infty$,*

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\big(\widetilde{z}_{i,j}-z_{i,j}\big)\Big|$$
$$=\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\Big(\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi_j(\varepsilon_{i,j}(s_\nu))-\int_0^1 r_{iM}(s)\psi_j(\varepsilon_{i,j}(s))ds\Big)\Big|=o_P(1). \tag{5.8}$$

Moreover, due to strict stationarity, the above relations also hold with $\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\sum_{i=m+1}^{m+k}$ *being replaced by* $\frac{1}{\sqrt{m}}\sum_{i=1}^{m}$.

(iii) *for $j=1,\dots,d$, as $m\to\infty$,*

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}\lambda_j'(0,s_\nu)\big(r_{iM}(s_\nu)-Er_{0M}(s_\nu)\big)\Big|=o_P(1); \tag{5.9}$$

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}\lambda_j'(0,s_\nu)\big(r_{iM}^2(s_\nu)-Er_{0M}^2(s_\nu)\big)\Big|=o_P(1). \tag{5.10}$$

Moreover, the above relations also hold with $\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^\gamma}\sum_{i=m+1}^{m+k}$ *being replaced by* $\frac{1}{m}\sum_{i=1}^{m}$.

Proof. Again, for the sake of simplicity, we omit the index j in the sequel.

(i) On interchanging summation, expectation and integration, a similar argument as in the proof of (5.3) gives

$$\begin{aligned}
&E\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi(\varepsilon_i(s))ds\Big|\\
&\le\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\Big\{E\Big(\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi(\varepsilon_i(s))\Big|\Big)^2\Big\}^{1/2}ds\\
&\le D\,(\log m)\sup_{s\in[0,1]}\|\psi(\varepsilon_0(s))\|_2\,\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\nu)-r_{0M}(s_\nu-h)\|_2,
\end{aligned}$$

with some $D>0$, where in the last inequality we made use of the independence of the sequences $\{r_{iM}(\cdot)\}$ and $\{\varepsilon_i(\cdot)\}$. Since $\sup_{s\in[0,1]}\|\psi(\varepsilon_0(s))\|_2<\infty$, this proves (i).

(ii) Consider

$$\begin{aligned}
&\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\Big(\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))-\int_0^1 r_{iM}(s)\psi(\varepsilon_i(s))ds\Big)\Big|\\
&=\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\big(r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))-r_{iM}(s)\psi(\varepsilon_i(s))\big)ds\Big|\\
&\le C\Big(\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi(\varepsilon_i(s))\Big|ds\\
&\quad+\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}r_{iM}(s_\nu)\big(\psi(\varepsilon_i(s_\nu))-\psi(\varepsilon_i(s))\big)\Big|ds\Big)\\
&=\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\big(I_{1,m}(s_\nu,s)+I_{2,m}(s_\nu,s)\big)ds. \qquad (5.11)
\end{aligned}$$

According to (5.7) and Assumption (B.7a), as $m\to\infty$,

$$E\Big(\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}I_{1,m}(s_\nu,s)ds\Big)\le C\,(\log m)\,\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\nu)-r_{0M}(s_\nu-h)\|_2=o(1). \qquad (5.12)$$

By an analogous argument,

$$\begin{aligned}
&E\Big(\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}I_{2,m}(s_\nu,s)ds\Big)\\
&\le\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}E\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}r_{iM}(s_\nu)\big(\psi(\varepsilon_i(s_\nu))-\psi(\varepsilon_i(s))\big)\Big|ds\\
&\le C\,(\log m)\sup_{s\in[0,1]}\|r_{0M}(s)\|_2\,\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|\psi(\varepsilon_i(s_\nu))-\psi(\varepsilon_i(s_\nu-h))\|_2=o(1), \qquad (5.13)
\end{aligned}$$

where we have used the independence of the sequences $\{r_{iM}(\cdot)\}$ and $\{\varepsilon_i(\cdot)\}$ once again in combination with Assumptions (B.1) and (B.7b).

(iii) In a first step, we replace $\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\big(r_{iM}^q(s_\nu)-Er_{0M}^q(s_\nu)\big)$ in (5.9)-(5.10) with

$$x_i=\int_0^1\lambda'(0,s)\big(r_{iM}^q(s)-Er_{0M}^q(s)\big)ds,\quad i=1,2,\ldots;\ q=1,2.$$

This can be done in a similar way as in the proof of (5.8). We even have an additional multiplication by $1/\sqrt{m}$ here. Note that the sequence $\{x_i\}_{i=1,2\ldots}$ is again strictly stationary and ergodic with $Ex_0=0$.

Now, it suffices to prove (5.9) and (5.10) with $\max_{K<k\le\lfloor mT\rfloor}$ instead of $\max_{1\le k\le\lfloor mT\rfloor}$ and $\sum_{i=K+1}^{m+k}$ replacing $\sum_{i=m+1}^{m+k}$, where $K=K_m$ is such that $K\to\infty$, but $K/m^{1-\gamma}\to 0$, e.g., for $K=\log m$.

In view of the strict stationarity and $Ex_0=0$, the ergodic theorem gives, as $m\to\infty$,

$$\max_{K<k\le\lfloor mT\rfloor}\frac{1}{k}\Big|\sum_{i=m+1}^{m+k}x_i\Big|=o_P(1).$$

On observing that

$$\max_{K<k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^{\gamma}}\left|\sum_{i=m+1}^{m+k}x_i\right|\le\; T^{1-\gamma}\max_{K<k\le\lfloor mT\rfloor}\frac{1}{k}\left|\sum_{i=m+1}^{m+k}x_i\right|,$$

this completes the proof of (5.9) and (5.10), respectively. □

Acknowledgements. The authors thank two anonymous referees for their constructive suggestions, which led to improve the presentation of their results. They also gratefully acknowledge support by grant GAČR P201/12/1277 (O. Chochola, M. Hušková and Z. Prášková) and DFG grant STE 306/22-1 (J.G. Steinebach).

References

[1] T.G. Andersen, T. Bollerslev, F.X. Diebold, J. Wu, Realized Beta: Persistence and predictability. In: T. Fomby and D. Terrell (eds.), Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger, vol. B, pp. 1–40. Elsevier, Amsterdam, 2006.

[2] D.W.K. Andrews, Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59 (1991), 817–858.

[3] A. Aue, S. Hörmann, L. Horváth, M. Hušková, J.G. Steinebach, Sequential testing for the stability of high-frequency portfolio betas. Econom. Theory 28 (2012) 804–837.

[4] I. Berkes, S. Hörmann, J. Schauer, Split invariance principles for stationary processes. Ann. Probab. 39 (2011) 2441–2473.

[5] P. Billingsley, Convergence of Probability Measures. Wiley, New York, 1968.

[6] O. Chochola, Robust Monitoring Procedures for Dependent Data. Ph.D. Thesis, Charles University of Prague, 2013.

[7] O. Chochola, M. Hušková, Z. Prášková, J.G. Steinebach, Robust monitoring of CAPM portfolio betas. J. Multivariate Analysis 118 (2013), 374–395.

[8] I. Fazekas, O. Klesov, A general approach to the strong law of large numbers. Theory Probab. Appl. 45 (2000) 436–448.

[9] E. Ghysels, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? J. Finance 53 (1998) 549–573.

[10] S. Hörmann, P. Kokoszka, Weakly dependent functional data. Ann. Statist. 38 (2010) 1845–1884.

[11] L. Horváth, P. Kokoszka, Inference for Functioal Data with Applications. Springer, New York, 2012.

[12] P.J. Huber, Robust Statistics. Wiley, New York, 1981.

[13] C. Kirch, Resampling Methods for the Change Analysis of Dependent Data. Ph.D. Thesis, University of Cologne, 2006.

[14] J. Lintner, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review Econom. Statist. 47 (1965) 13–37.

[15] R. Merton, An intertemporal asset pricing model. Econometrica 41 (1973) 867–880.

[16] F. Móricz, Moment inequalities and the strong laws of large numbers. Z. Wahrsch. verw. Geb. 35 (1976) 299–314.

[17] F. Móricz, R. Serfling, W. Stout, Moment and probability bounds with quasi-superadditive structure for the maximum partial sums. Ann. Probab. 10 (1982) 1032–1040.

[18] W.F. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk. J. Finance 19 (1964) 424–442.

[19] A. Zeileis, Econometric computing with HC and HAC covariance matrix estimators. J. Statist. Software 11 (2004), 1–17.

Econometric Reviews
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lecr20

Fourier-Type Tests Involving Martingale Difference Processes

Zdeněk Hlávka[a], Marie Hušková[a], Claudia Kirch[b] & Simos G. Meintanis[cd]

[a] Faculty of Mathematics and Physics, Department of Statistics, Charles University in Prague, Sokolovská, Czech Republic

[b] Karlsruhe Institute of Technology, Institute of Stochastics, Karlsruhe, Germany

[c] Department of Economics, National and Kapodistrian University of Athens, Athens, Greece

[d] Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa

Accepted author version posted online: 20 Oct 2014.

Fourier–Type Tests Involving Martingale Difference Processes

Zdeněk Hlávka[1,], Marie Hušková[1], Claudia Kirch[2], Simos G. Meintanis[3,4]

[1]Faculty of Mathematics and Physics, Department of Statistics, Charles University in Prague, Sokolovská, Czech Republic,

[2]Karlsruhe Institute of Technology, Institute of Stochastics, Karlsruhe, Germany,

[3]Department of Economics, National and Kapodistrian University of Athens, Athens, Greece,

[4]Unit for Business Mathematics and Informatics, North–West University, Potchefstroom, South Africa

E-mail: hlavka@karlin.mff.cuni.cz

On sabbatical leave from the University of Athens

Abstract

We develop testing procedures which detect if the observed time series is a martingale difference sequence. Furthermore, tests are developed that detect change–points in the conditional expectation of the series given its past. The test statistics are formulated following the approach of Fourier–type conditional expectations first proposed by Bierens (1982) and have the advantage of computational simplicity. The limit behavior of the test statistics is investigated under the null hypothesis as well as under alternatives. Since the asymptotic null distribution contains unknown parameters, a bootstrap procedure is proposed in order to actually perform the test. The performance of the bootstrap version of the test is compared in finite samples with other methods for the same problem. A real–data application is also included.

KEYWORDS: Martingale difference hypothesis; Change–point test; Bootstrap test; Empirical characteristic function

1 INTRODUCTION

The martingale difference hypothesis (MDH) is a property integrated in many statistical models popular in finance and economics. Such an assumption is standard with asset returns in an efficient market, with changes in consumption, as well as with disturbances in a correctly specified time–series regression model, among others. The basic idea underlying the MDH is the unpredictability of macro and financial series on the basis of currently available information. Hence the MDH is critically related to the efficient market hypothesis first put forward in seminal papers by Samuelson (1965) and Fama (1970). The efficient market hypothesis states that in efficient markets, prices follow a martingale and always fully and instantaneously reflect all available relevant information consisting of past prices and returns. Consequently, as price adjustments to new piece of information is instantaneous and accurate, no agent, however well informed, can use market information as a basis for superior forecasts or in order to accomplish trading profits beyond the level justified by the risk which he/she assumes. However, most efficiency studies on financial markets focus on a weak form of market efficiency through the MDH, whereby the profit expected from an asset which is forecasted to have its future price equal to its the current price is equal to zero. The MDH for exchange rates has been a major concern in the finance literature and many authors have investigated this hypothesis, often with mixed conclusions; see for instance, Belaire-Franch and Contreras (2011), Yilmaz (2003), Hong and Lee (2003), Fong et al. (1997), and Fong and Ouliaris

(1995). Less standard areas where the MDH has been put forward include electricity prices (Veka, 2013) and CO_2 emissions (Daskalakis et al., 2009, Charles et al., 2011a), among others.

The standard formulation of the MDH is

$$\mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = 0, \qquad t = 1, \ldots, \qquad (1)$$

where $\mathbb{I}_t$ denotes the information set available at time t, and Y_t represents first differences of a process which under this hypothesis forms a martingale sequence. The statistical problem of testing the null hypothesis in (1.1) has been addressed by many authors; see Escanciano and Lobato (2009a) for an excellent survey. One approach is to consider methods which test for lack of autocorrelation in Y_t, a condition which is a necessary (but not a sufficient) for the MDH. A standard method for testing autocorrelation of fixed finite order is via the Box–Pierce Portmanteau test commonly implemented via its Ljung and Box (1978) modification. Other methods include the robustified Box–Pierce statistic (or variance ratio test) of Lo and MacKinlay (1988), the automatic variance ratio test of Choi (1999), that was later modified for heteroscedasticity by Kim (2009), and the automatic Portmanteau test of Escanciano and Lobato (2009b), which is data–driven with respect to the order of autocorrelation. (Recently, a different data–driven test using neural networks and boosting has been put forward by Kapetanios and Blake, 2010). In this category we also mention the tests of Lobato et al. (2002) which works under general dependence structures and that of Francq et al. (2005) applied not on the original observations but on corresponding residuals. On the other hand, MDH tests involving an infinite set of autocorrelations are derived in the frequency domain via the spectral density or distribution. Such a test was first proposed by Durlauf (1991) and was later extended by Deo (2000) to account for conditional heteroscedasticity.

While the aforementioned methods only consider linear dependencies (of finite or infinite order), a number of testing procedures for (1.1) make use of the following equivalence relation

$$\mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = 0 \Leftrightarrow \mathbb{E}[Y_t v(\mathbb{I}_{t-1})] = 0,\ t = 1, \dots, \qquad (2)$$

where $v(\cdot)$ is a suitably chosen weight function. As before, eqn. (1.2) may involve a finite but also a potentially infinite set of time lags. An additional challenge with these tests is the choice of the weight function $v(\cdot)$ figuring in (1.2). In this connection a subclass of such tests uses indicator functions, while others use non–linear but smooth weight functions $v(\cdot)$. The tests of Domínguez and Lobato (2003), Escanciano and Velasco (2006a) and Escanciano and Mayoral (2010), belong to the first category. The second line of research was initiated by Bierens (1982), and includes de Jong (1996), Bierens and Ploberger (1997), Kuan and Lee (2004) and Bierens and Wang (2012). In these papers instead of using classical autocorrelation between Y_t and past values, the authors employ nonlinear exponential–type transformations of such past values.

In the present paper we follow the approach of Bierens and consider tests for the MDH involving an exponential–type weight function. Note that this approach is also followed by Hong (1999) and Escanciano and Velasco (2006b), who propose tests based on the 'generalized autocovariances' figuring in the equation in the right–hand side of (1.2).

However while these authors construct their tests based on L2–type distances involving spectral densities (Hong, 1999) or spectral distribution functions (Escanciano and Velasco, 2006b) corresponding to the general autocovariances, our tests are based on distances involving these autocovariances directly, and properly integrated over an argument u which will be specified below. As it will be seen, the new test has the advantage of computational simplicity. Moreover the tests proposed herein will be based on a fixed number of lags. In this connection, we note that it would be natural to allow the time lag to be time dependent, and possibly increase without bound with the sample size as in de Jong (1996). Although this issue is also discussed in the paper and shown that such an approach is still feasible both from the practical as well as from the theoretical point of view, our computations and proofs will be much more technical and therefore we shall not pursue this aspect of our tests in detail in this work.

However, the main contribution of the present paper is the development of change–point tests involving the MDH, an issue which, to the best of our knowledge, has not been addressed before in the literature. The relevance of this problem stems from the fact that despite the prominence of the MDH in the financial literature, market efficiency should be viewed as a dynamic state. Specifically the idea of market adaptivity goes back to Grossman and Stiglitz (1980), who postulate the rise of occasional profitable opportunities to compensate investors for the cost of analysing the market. (This is also compatible with the adaptive market efficiency hypothesis of Lo, 2004.) In other words markets experience structural changes, and at a certain time may pass from a state of market efficiency, where the MDH holds true, to a state where well–informed agents can systematically gain excess returns, and vice versa. (For recent empirical evidence in stock returns see Todea and Lazăr, 2012, and Kim et al., 2011.)

The remainder of this paper is as follows. In Section 2 we specify the null hypotheses considered and the corresponding test statistics, while Section 3 addresses the issue of computation of these statistics. Section 4 is devoted to the asymptotic behavior of the tests under the null hypothesis as well as alternatives while in Section 5, a bootstrap version is formulated and its asymptotic validity is shown. The finite–sample properties of the proposed methods are investigated in Section 6 where we also present a real–data example. Finally, the proofs are given in Section 7.

2 NULL HYPOTHESES AND TEST STATISTICS

We consider three different types of hypotheses and for each particular null hypothesis propose a corresponding test statistic. Specifically on the basis of observations $Y_1, ..., Y_n$, we first consider the null hypothesis of a martingale difference sequence (MDS)

$$H_0^{(1)}: \quad \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = 0$$

and test it against the hypothesis that the structure indeed depends on the past observations:

$$H_1^{(1)}: \quad \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = g(Y_{t-1}, \ldots, Y_{t-m}),$$

$$P(g(Y_{t-1}, \ldots, Y_{t-m}) = 0) < 1,$$

for an arbitrary unknown function g.

Based on the same methodology, we also propose two tests for changes in the structure of a given process. First, we test the MDH against the alternative that the process changes from a martingale structure to a non–martingale structure at some unknown point k_0.

Specifically the null and the alternative hypothesis are as follows:

$$
\begin{aligned}
&H_0^{(2)}: \quad \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = 0, \\
&H_1^{(2)}: \quad \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = 0, \quad t < k_0, \\
&\text{but } \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = g(Y_{t-1}, \ldots, Y_{t-m}), \quad t \geq k_0 \\
&\quad P(g(Y_{t-1}, \ldots, Y_{t-m}) = 0) < 1.
\end{aligned}
$$

The unknown point $1 < k_0 < n$ is called change–point.

The next test considers a generalized version of the MDH and a possible change in the conditional martingale difference structure:

$$
\begin{aligned}
&H_0^{(3)}: \quad \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = c, \\
&H_1^{(3)}: \quad \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = c, \quad t < k_0, \\
&\text{but } \mathbb{E}(Y_t \mid \mathbb{I}_{t-1}) = g(Y_{t-1}, \ldots, Y_{t-m}), \quad t \geq k_0, \\
&\quad P(h(Y_{t-1}, \ldots, Y_{t-m}) = c) < 1,
\end{aligned}
$$

where c is an unknown constant. As already mentioned, the parameter m is chosen in advance and therefore restricts the kind of alternatives that are detectable. Also, although the tests are formulated for general time t, it is clear that detection of the alternative hypotheses is considered only after the first m values of the process Y_t have been observed.

We formulate our procedure by using the following characterization of Bierens (1982): For real y and a given vector x of dimension m, $\mathbb{E}(y \mid x) = 0$ holds if and only if $\mathbb{E}(ye^{iu'x}) = 0$, for all $u \in \mathbb{R}^m$. In view of this characterization let

$$S_t^{(m)}(u) = \frac{1}{\sqrt{n}} \sum_{\tau=m+1}^{t} Y_\tau e^{iu'\mathbf{Y}_{\tau,m}}, \quad t = m+1,\ldots,n, \qquad (2.1)$$

$$S_t^{(m)}(u) = 0, \quad t = 0,1,\ldots,m,$$

where $\mathbf{Y}_{t,m} = (Y_{t-1}, Y_{t-2}, ..., Y_{t-m})'$ and $m > 0$ denotes a chosen time–lag, and consider the integrated process

$$Q_m(s) = \int_{\mathbb{R}^m} | S_{\lfloor sn \rfloor}^{(m)}(u) - S_n^{(m)}(u) |^2 \, w(u)du, \quad 0 \le s \le 1, \qquad (2.2)$$

where $w(\cdot)$ denotes a weight function the choice of which we shall discuss in the next paragraph. Notice that this approach amounts to choosing $v(\mathbb{I}_{t-1}) := v(\mathbb{I}_{t-1}, u) = e^{iu'\mathbf{Y}_{t,m}}$ in the original formulation in (1.2).

We suggest to reject the null hypothesis $H_0^{(1)}$ against alternative $H_1^{(1)}$ if

$$T_n^{(1)} := Q_m(0) \qquad (2.3)$$

is large. Notice that $T_n^{(1)}$ is related to the test statistic developed by Escanciano (2009) for a slightly different problem.

Likewise, the null hypothesis $H_0^{(2)}$ is rejected in favor of alternative $H_1^{(2)}$ if

$$T_n^{(2)}(\gamma) := \max_{m+1 \le k \le n} Q_m(k/n) / q(k/n, \gamma) \qquad (2.4)$$

is large, where

$$q(s,\gamma) = (1-s)^\gamma, s \in (0,1), \quad 0 \le \gamma < 1. \qquad (2.5)$$

In turn, the null hypothesis $H_0^{(3)}$ should be rejected against alternative $H_1^{(3)}$ if

$$T_n^{(3)}(\gamma) := \max_{m+1 \le k \le n} \tilde{Q}_m(k/n) / \tilde{q}(k/n, \gamma), \qquad (2.6)$$

is large, where

$$\tilde{Q}_m(s) = \int_{\mathbb{R}^m} | S^{(m)}_{\lfloor sn \rfloor}(u) - sS^{(m)}_n(u) |^2 \ w(u)du, \qquad (2.7)$$

$$\tilde{q}(t,\gamma) = (s(1-s))^{\gamma}, s \in (0,1), \quad 0 \le \gamma < 1. \qquad (2.8)$$

The choice of $\gamma \in [0,1)$ is made in concordance with the test procedures generally used in change point analysis (e.g. Horváth and Kokoszka, 1997), i.e., γ close to 1 leads to the procedures more sensitive w.r.t a change either at the beginning or at the end in comparison with those with γ close to 0.

As already noticed the tests may not be consistent against some specific alternatives because they take into account only a finite number m of lags. As a remedy one can formally replace , e.g., $\mathbf{u}'\mathbf{Y}_{\tau,m}$ by $\sum_{j=1}^{t-1} u_j Y_{t-j} a_j$ in $S_t^{(m)}(u)$ with a_j 's chosen in such way that they restrict influence of Y_{t-j} for large j. See de Jong (1996) for an idea. However, as a consequence we lose computational simplicity, assumptions in theorems in Sections 4 and 5 have to be much stronger, and the proofs become still more complex.

3 COMPUTATION OF TEST STATISTICS

As it will be seen in Section 4, the asymptotic theory of the tests developed in this paper is valid for a large class of weight functions $w(\cdot)$. Here however we will carry out the computation step–by–step, and show that certain classes of weight functions render the test statistics with the desirable property of computational simplicity.

Consider first the test statistic figuring in (2.3), and observe that straightforward calculation leads to:

$$Q_m(0) = \int_{\mathbb{R}^m} | S_n^{(m)}(u) |^2 \; w(\mathbf{u})d\mathbf{u}$$
$$= \frac{1}{n} \sum_{\tau=m+1}^{n} \sum_{\sigma=m+1}^{n} Y_\tau Y_\sigma I_w (Y_{\tau,m} - Y_{\sigma,m}), \qquad (3.1)$$

where.

Assuming that, for each argument $u := (u_1,...,u_m)' \in \mathbb{R}^m$, the weight function can be decomposed as $w(u) = \prod_{j=1}^{m} w_j(u_j)$, where $w_j(\cdot)$ satisfies $w_j(u) = w_j(-u)$ for each $u \in \mathbb{R}$, we obtain

$$I_w(Y_{\tau,m} - Y_{\sigma,m}) = \prod_{j=1}^{m} \int \cos\{u_j (Y_{\tau-j} - Y_{\sigma-j})\} w_j(u_j) du_j = \prod_{j=1}^{m} I_{w_j}(Y_{\tau-j} - Y_{\sigma-j}). \qquad (3.2)$$

For simplicity we take $w_1 \equiv w_2 \equiv \cdots \equiv w_m \equiv w$ (say).

It is now clear that there is no difficulty in choosing the weight function $w(\cdot)$ that leads to a test statistic $T_n^{(1)} = Q_m(0)$ which is easy to calculate. In fact, we may use any (even) weight function $w(u)$ as a building block (provided that the corresponding integral I_w is analytic), and define $w(u)$ by the product equation figuring above (12). Examples of such functions are the following: (i) $w(u) = e^{-|u|^b}, b = 1, 2$, (ii) $w(u) = 2(1-\cos(u))/u^2$ and (iii) $w(u) = 1, |u| \le 1$ and $w(u) = 0, |u| > 1$. We may also slightly generalize by making the transformation $u \mapsto au$ in $w(u)$, which for each choice of $w(\cdot)$ results in a corresponding parametric family of weight functions indexed by the parameter $a > 0$.

In the simulation study, we will be using the weight function $w(u) = e^{-a|u|}$ for which the integral $I_w(x)$ reduces to

$$I_w(x) = \int \cos(ux)w(u)du = \frac{2/a}{1+(x/a)^2}. \qquad (3.3)$$

At this point we shall investigate the role of the weight function $w(u)$ in the test statistic in (3.1). Our arguments will be heuristic somewhat, but nevertheless reveal interesting connections. First recall from (3.1) and (3.2) that

$$Q_m(0) = \frac{1}{n}\sum_{\tau,\sigma=m+1}^{n} Y_\tau Y_\sigma \prod_{j=1}^{m} I_w(Y_{\tau-j} - Y_{\sigma-j}).$$

Now assume that $\int u^\rho w(u)du < \infty, \rho > 0$, and use a Taylor expansion of $\cos(x)$ in $I_w(x)$ to obtain

$$Q_m(0) = \frac{1}{n}\sum_{\tau,\sigma=m+1}^{n} Y_\tau Y_\sigma v_{\tau\sigma},$$

where $v_{\tau\sigma} = \prod_{j=1}^{m} v_{\tau\sigma}^{(j)}$, with

$$v_{\tau\sigma}^{(j)} = \sum_{k=0}^{\infty}(-1)^k \frac{(Y_{\tau-j} - Y_{\sigma-j})^{2k}}{(2k)!}\lambda_{2k},$$

and $\lambda_\rho = \int u^\rho w(u)du$. It is clear from the previous equations that the test statistic $Q_m(0)$ comes in a form reminiscent of a weighted V –statistic where each product $Y_\tau Y_\sigma$ receives a total weight $v_{\tau\sigma}$ with (component) weights $v_{\tau\sigma}^{(j)}$ depending on past observations and on the weight function w in a complicated way. In $v_{\tau\sigma}^{(j)}$ specifically, the contribution of each pair of past observations (through $Y_{\tau-j} - Y_{\sigma-j}$) is determined by, among other things, the weight function w, via the quantities λ_ρ. To gain some extra insight, take $w(u) = e^{-a|u|}$ so that $\lambda_\rho = (2\rho!)/a^{1+\rho}, \rho = 2k$, and write Q_a for the resulting test statistic that is easily seen to simplify to

$$Q_a = \frac{1}{n} \sum_{\tau,\sigma=m+1}^{n} Y_\tau Y_\sigma \prod_{j=1}^{m} \sum_{k=0}^{\infty} (-1)^k \frac{2}{a^{1+2k}} (Y_{\tau-j} - Y_{\sigma-j})^{2k}.$$

From the last equation it is evident that while in the test statistic the weights depend on past observations, this dependence decreases as the value of a increases. In fact asymptotically we have

$$\lim_{a\to\infty} aQ_a = 2^m \frac{1}{n} \sum_{\tau,\sigma=m+1}^{n} Y_\tau Y_\sigma,$$

which shows that in the limit the contribution of the weights $v_{\tau\sigma}$ to the value of the test statistic is neutralized, as $a \to \infty$. On the other hand a value of a which is too small causes numerical instability in Q_a. (In this connection notice that for $a = 0$, Q_a is no longer finite.) Consequently the choice of the value of a comes down to a compromise which should, on the one hand avoid large values of a that diminish the influence of the weights on the test statistic, but on the other hand should also avoid values of this parameter near the origin that render the test statistic vulnerable to numerical error. (This reasoning was specific to the weight function $e^{-a|u|}$, but it can be easily seen to apply to other weight functions such as e^{-au^2}, for instance.) Moreover, in the context of goodness–of–fit testing with i.i.d. observations it has been documented that the choice of the specific parametric form of the weight function $w(u)$ is much less important than the choice of the value of the weight parameter a. In this respect the situation here is similar to nonparametric density estimation where the choice of the kernel is much less important than the choice of the bandwidth. In fact, there is an interesting connection between, on the one hand the weight function w and the weight parameter a, and on the other hand the kernel and the bandwidth, respectively, of nonparametric estimation of a corresponding density; see for instance Meintanis (2013) and Henze et al. (2005) for this connection in the context of goodness–of–fit testing with i.i.d. observations. To see this recall representation (11) for the test statistic $Q_m(0)$ and take $w_a(u) = w(au), a > 0$, as weight function. Then by a simple change of variables in the integral $I_w(x)$ we can write

$$Q_m(0) = \frac{1}{n} \sum_{\tau,\sigma=m+1}^{n} Y_\tau Y_\sigma K_{\tau\sigma},$$

where

$$K_{\tau\sigma} = \left(\frac{1}{a}\right)^m I_w\left(\frac{\mathbf{Y}_{\tau,m} - \mathbf{Y}_{\sigma,m}}{a}\right).$$

The last representation shows that the statistic $Q_m(0)$ is in the same spirit as the criterion used by Lavergne and Patilea (2013), where a similar statistic is used for estimation of parameters in parametric models by means of a method involving conditional moments. As in Lavergne and Patilea (2013) our kernel $I_w(\cdot)$ (defined below (3.1)) takes the form of the characteristic function of a measure $w(\cdot)$ (possibly non-normalized), the only difference being that here $w(\cdot)$ is assumed to be symmetric around the origin. By way of example take $w(u) = e^{-\|u\|^2}$, and notice that then

$$I_w(x) = \pi^{m/2} e^{-\|x\|^2/4},$$

which implies that the kernel coincides with a constant multiple of the density of $\mathcal{N}(0, 2\mathbb{I}_m)$, i.e. of a zero–mean multivariate normal distribution with covariance matrix equal to the identity matrix (of dimension m) multiplied by 2. Making a connection with the limit obtained above we note that, unlike the case of density estimation in which we typically assume that $a = a_n \to 0$, as $n \to \infty$, here we consider a fixed bandwidth as in

Lavergne and Patilea (2013). Also the interesting limit results for fixed sample size n and in the case of $a \to \infty$, and not as $a \to 0$, which in fact is not feasible. A finer analysis of the effect of the value of a on the power of tests was undertaken by Tenreiro (2009), but this is confined to the strict parametric context of i.i.d. testing for univariate normality, and even then the entire problem is highly non–trivial and involves a series of approximations, not to mention the need for apriori fixing specific deviations from the null hypothesis in order to determine an optimal value for a.

Concerning the test statistic $T_n^{(2)}(\gamma)$ defined in (6), it is most important to evaluate $Q_m(k/n)$. Proceeding similarly as in (11), it is easy to see that

$$\begin{aligned}
Q_m(k/n) &= \int_{\mathbb{R}^m} | S_k^{(m)}(u) - S_n^{(m)}(u) |^2 \; w(\mathrm{u}) d\mathbf{u} \\
&= \int_{\mathbb{R}^m} \left| \frac{1}{\sqrt{n}} \left\{ \sum_{m+1}^{k} Y_\tau e^{iu'Y_{\tau,m}} - \sum_{m+1}^{n} Y_\tau e^{iu'Y_{\tau,m}} \right\} \right|^2 w(\mathbf{u}) d\mathbf{u} \\
&= \int_{\mathbb{R}^m} \left| \frac{1}{\sqrt{n}} \sum_{k+1}^{n} Y_\tau e^{iu'Y_{\tau,m}} \right|^2 w(u) d\mathbf{u} \\
&= \frac{1}{n} \sum_{\tau=k+1}^{n} \sum_{\sigma=k+1}^{n} Y_\tau Y_\sigma I_w (Y_{\tau,m} - Y_{\sigma,m}) \\
&= \frac{1}{n} \sum_{\tau=k+1}^{n} \sum_{\sigma=k+1}^{n} Y_\tau Y_\sigma \prod_{j=1}^{m} I_w (Y_{\tau-j} - Y_{\sigma-j}) \\
&= \frac{(2/a)^m}{n} \sum_{\tau=k+1}^{n} \sum_{\sigma=k+1}^{n} Y_\tau Y_\sigma \prod_{j=1}^{m} \frac{1}{1+\{(Y_{\tau-j} - Y_{\sigma-j})/a\}^2},
\end{aligned} \tag{3.4}$$

where in the last step we choose $w(u) = e^{-a|u|}$. We also note that the terms $Q_m(k/n)$ may be calculated recursively from $k = n$ to $k = m+1$ and therefore the computational complexity of the sequential test statistic $T_n^{(2)}(\gamma)$ is obviously of the same order as that of the test statistic $T_n^{(1)}$.

Compared to $T_n^{(1)}$ and $T_n^{(2)}$, the calculation of the test statistic $T_n^{(3)}$ defined in (8) is more involved because we have to evaluate the term $\tilde{Q}_m(k/n)$ for which we have the following:

$$\begin{aligned}
\tilde{Q}_m(k/n) &= \int_{\mathbb{R}^m} \left| S_k^{(m)}(u) - \frac{k}{n} S_n^{(m)}(u) \right|^2 w(u) d\mathbf{u} \\
&= \frac{1}{n} \sum_{\tau=m+1}^{n} \sum_{\sigma=m+1}^{n} c_{\tau,k} c_{\sigma,k} Y_\tau Y_\sigma I_w(Y_{\tau,m} - Y_{\sigma,m}) \\
&= \frac{(2/a)^m}{n} \sum_{\tau=m+1}^{n} \sum_{\sigma=m+1}^{n} c_{\tau,k} c_{\sigma,k} Y_\tau Y_\sigma \prod_{j=1}^{m} \frac{1}{1+\{(Y_{\tau-j} - Y_{\sigma-j})/a\}^2},
\end{aligned} \tag{3.5}$$

where $c_{\tau,k} = I(\tau \le k) - k/n$.

We close this section by noting that our methods have a direct connection with methods involving the empirical characteristic function and that, apart from testing the MDH, such methods have been efficiently employed in the past for a variety of testing problems with dependent data. Earlier works are Epps (1988, 1987) and Feuerverger (1990), while the most recent literature includes Quessy and Éthier (2012), Leucht (2012), Hlávka et al. (2012), and Ghosh (2012).

4 ASYMPTOTIC BEHAVIOR OF THE TEST STATISTICS

4.1 Behavior Under The Null Hypothesis

The next theorem gives the asymptotics of the introduced statistics under the null hypothesis.

Theorem 4.1 *Assume that* $\{Y_t\}$ *is a martingale difference sequence as well as stationary and ergodic with* $\mathbb{E}|Y_1|^{2+\delta} < \infty$ *for some* $\delta > 0$ *and let* $w(\cdot)$ *be a measurable non-negative function on* $\mathbb{R}^m$ *such that*

$$w(\mathbf{t}) = w(-\mathbf{t}) > 0, \quad for \quad all \quad \mathbf{t} \in \mathbb{R}^m, \quad 0 < \int_{\mathbb{R}^m} w(\mathbf{t}) d\mathbf{t} < \infty. \tag{4.1}$$

Then as $n \to \infty$:

$a)\quad T_n^{(1)} \xrightarrow{\mathcal{L}} \int_{\mathbb{R}^m} |Z(0,\mathbf{u})|^2 \; w(\mathbf{u}) du,$

$b)\quad T_n^{(2)}(\gamma) \xrightarrow{\mathcal{L}} \sup_{0<s<1} \frac{1}{(1-s)^\gamma} \int_{\mathbb{R}^m} |Z(s,\mathbf{u}) - Z(1,\mathbf{u})|^2 \; w(\mathbf{u}) d\mathbf{u},$

$c)\quad T_n^{(3)}(\gamma) \xrightarrow{\mathcal{L}} \sup_{0<s<1} \frac{1}{(s(1-s))^\gamma} \int_{\mathbb{R}^m} |Z(s,\mathbf{u}) - sZ(1,\mathbf{u})|^2 \; w(\mathbf{u}) d\mathbf{u},$

where $0 \le \gamma < 1$, $\{Z(s,\mathbf{u}), s \in [0,1], \mathbf{u} \in \mathbb{R}^m\}$ is a Gaussian process with expectation zero and covariance ($0 \le s_1 \le s_2 \le 1$)

$$cov\{Z(s_1,\mathbf{u}_1), Z(s_2,\mathbf{u}_2)\} = s_1 E(Y_{m+1}^2 h(\mathbf{Y}_{m+1}, u_1) h(\mathbf{Y}_{m+1}, u_2)), \quad u_1, u_2,$$

$$h(\mathbf{Y}_m, \mathbf{u}) = \cos\left(\sum_{q=1}^m u_q Y_{m+1-q}\right) + \sin\left(\sum_{q=1}^m u_q Y_{m+1-q}\right), \tag{17}$$

Here $\mathbf{u} = (u_1, \ldots, u_m)', \mathbf{Y}_{m+1} = (Y_m, \ldots, Y_1)'$.

The proof is postponed to Section 7.

The assumption that the MDS is stationary and ergodic is only needed to apply the central limit theorem and other limit theorems for stationary and ergodic sequences. Otherwise a more general form of limit theorems for MDS have to be used and the proofs become still more technical and less transparent.

The assertion of our theorem remains true if $\mathrm{cov}\{Z(s_1,\mathbf{u}_1), Z(s_2,\mathbf{u}_2)\}$ are replaced by their consistent estimators. Then critical values can be obtained by simulating the limit distribution. But it is more convenient to get their approximation via a proper bootstrap as explained in Section 5.

At the end of Section 2 we have shortly discussed possible modifications of the test statistics to increasing number of lags m. We expect that under proper assumptions (much stronger than above) the limit behavior of the modified test statistics will have similar structure as in the above theorem but with a Gaussian process $\{Z(s,\mathbf{u}); s\in[0,1], \mathbf{u}\in\mathbb{R}^m\}$ with more complex dependence structure, and moreover the dimensions of both the process and the weight function $w(\cdot)$ tends to ∞ together with n. To prove an analog to the above theorems requires among others to derive an extension of crucial Lemma 7.1. We are losing among others the properties of stationarity and ergodicity, and the dimensions of both the process $\{Z(s,\mathbf{u}); s\in[0,1], \mathbf{u}\in\mathbb{R}^m\}$ and the weight function $w(\cdot)$ tend to ∞ together with n.

4.2 Behavior Under Alternatives

Here various results on the limit behavior of our test statistics are presented. We formulate alternatives through a sequence of martingale difference sequences $\{\xi_t\}$ perturbed by some function g. For $\{\xi_t\}$ and g we postulate

1. $\{\xi_t\}$ is a stationary and ergodic martingale difference sequence and g is a measurable function such that

$$P(g(\xi_{m+1})=0)<1, \qquad \mathbb{E}|\xi_1|^{2+\delta}<\infty, \quad \mathbb{E}|g(\xi_{m+1})|^2<\infty$$

for some $\delta>0$.

4.2.1 Alternative Of A Stationary Time Series That Is Not An MDS

In order to get an idea about the asymptotic power of statistic $T_n^{(1)}$ we consider the following types of alternatives: [**A1**a:]

1. **Fixed alternative:**

$$Y_k = \xi_k + g(\xi_k),$$

where $(\{\xi_t\}, g)$ satisfies AE.

2. **Local alternative:** We observe $Y_1,\dots,Y_n$ with

$$Y_k = \xi_k + g(\mathrm{x}_k)d_n,$$

where $d_n \to 0$ and $(\{\xi_t\}, g)$ satisfies AE.

Theorem 4.2 *Let (16) be satisfied.*

a) For the fixed alternative $A1a$ with

$$\int_{R^m}(\mathbb{E}g(\mathrm{x}_{n+1})h(\mathbf{Y}_{n+1},\mathbf{u}))^2 w(\mathbf{u})d\mathbf{u}>0 \qquad (4.3)$$

the following holds

$$T_n^{(1)} \xrightarrow{P} \infty. \qquad (4.4)$$

b) For the local alternative $A1b$ with $|d_n\sqrt{n}| \to \infty$ as $n \to \infty$ and

$$0 < \int_{R^m} r^2(\mathbf{u}) w(\mathbf{u}) d\mathbf{u} < \infty \qquad (4.5)$$

where

$$r(\mathbf{u}) = E(g(\mathrm{x}_{m+1}) h(\mathrm{x}_{m+1}, \mathbf{u})) \qquad (4.6)$$

(h as in (17)), equation (19) holds true.

c) For the local alternative $A1b$ with $d_n\sqrt{n} \to b \neq 0$ the following holds true

$$T_n^{(1)} \xrightarrow{\mathcal{L}} \int_{\mathbb{R}^m} |Z(0,\mathbf{u}) + br(\mathbf{u})|^2 \, w(\mathbf{u}) du,$$

where $\{Z(0,\mathbf{u}); \mathbf{u}) \in \mathbb{R}^m\}$ is a Gaussian process from Theorem 4.1 and (20) is satisfied.

4.2.2 Change-Point Alternative With An MDS Before The Change

In order to get an idea about the asymptotic power of statistic $T_n^{(2)}$ we consider the following types of alternatives:

[**A2**a]: **Fixed alternative:** We observe

$$Y_k = \xi_k + g(\xi_k) 1_{\{k > k_0\}}, \quad k_0 = \lfloor \lambda n \rfloor$$

for some $0 < \lambda < 1$, where $(\{\xi_t\}, g)$ fulfill assumption AE. Let (20) be satisfied.

[**A2**b]: **Local alternative:** We observe $Y_1, \ldots, Y_n$ with

$$Y_k = \xi_k + g(\xi_k) d_n 1_{\{k > k_0\}}, \quad k_0 = \lfloor \lambda n \rfloor$$

for some $0 < \lambda < 1$ and some constant c, where $\{d_n\}$ is a sequence of real numbers with $d_n \to 0$ and where $(\{\xi_t\}, g)$ satisfies AE.

Theorem 4.3 *Let (16) be satisfied. [a)]*

1. For the fixed alternative $A2a$ with (18) the following holds

$T_n^{(2)}(\gamma) \xrightarrow{P} \infty.$

2. In the situation of the local alternative $A2b$ with $|d_n\sqrt{n}| \to \infty$ and if and (20) are satisfied then the following holds

$T_n^{(2)}(\gamma) \xrightarrow{P} \infty.$

3. For the local alternative $A2b$ with $d_n\sqrt{n} \to b \neq 0$ it holds

$$T_n^{(2)}(\gamma) \xrightarrow{\mathcal{L}} \sup_{0<s<1} \frac{1}{(1-s)^\gamma} \int_{\mathbb{R}^m} |Z(s,\mathbf{u}) - Z(1,\mathbf{u}) - (1-\max(\lambda,s))br(\mathbf{u})|^2 \, w(\mathbf{u})du,$$

where $r(\mathbf{u})$ is defined in (21) and (20) is satisfied.

4.2.3 General Change-Point Alternative

In order to get an idea about the asymptotic power of statistic $T_n^{(3)}$ we consider the following types of alternatives:

[**A3**a:] **Fixed alternative:** We observe

$$Y_k = \xi_k + g(\xi_k)\,1_{\{k \geq k_0\}} + c, \quad k_0 = \lfloor \lambda n \rfloor$$

for some $0 < \lambda < 1$, where $(\{\xi_t\}, g)$ satisfies AE and (20), c is a positive constant.

[**A3**b]: 2. **Local alternative:** We observe $Y_1, \ldots, Y_n$ with

$$Y_k = \xi_k + c + g(\xi_k) d_n \, 1_{\{k > k_0\}}, \quad k_0 = \lfloor \lambda n \rfloor$$

for some $0 < \lambda < 1$, where $\{d_n\}$ is a sequence of real numbers with $d_n \to 0$ and $(\{\xi_t\}, g)$ satisfies AE and (20).

Theorem 4.4 *Let (16) be satisfied. [a)]*

a) For the fixed alternative $A3a$ with (18) the following holds

$$T_n^{(3)}(\gamma) \xrightarrow{P} \infty.$$

b) For local alternatives $A3b$ with $|d_n\sqrt{n}| \to \infty$ it holds

$$T_n^{(3)}(\gamma) \xrightarrow{P} \infty.$$

c). For the local alternative $A3$ with $d_n\sqrt{n} \to b \neq 0$ it holds

$$T_n^{(3)}(\gamma) \xrightarrow{\mathcal{L}} \sup_{0<s<1} \frac{1}{(s(1-s))^{\gamma}} \int_{\mathbb{R}^m} |Z(s,\mathbf{u}) - sZ(1,\mathbf{u}) - \min(\lambda, s)(1-\max(s,\lambda)br(\mathbf{u})|^2 \; w(\mathbf{u})du,$$

where $r(\mathbf{u})$ is defined in (21) and satisfies (20).

The proofs are postponed to Section 7.

One can infer from the above theorems that our tests are consistent for the respective fixed alternatives as well as for the local ones if $\sqrt{n}\,|d_n| \to \infty$. The case with $|d_n| \approx n^{-1/2}$ is a border line, where the limit distribution differs from the limit null one but is bounded in probability.

The test statistics $T_n^{(3)}(\gamma)$ can be also used for testing $H_0^{(2)}$ against $H_1^{(2)}$ however $T_n^{(2)}(\gamma)$ has a higher power (as seen from Theorems 4.3 and 4.4).

At the end of this section we shortly discuss the problem of estimation the change point k_0 for situations considered in Theorem 4.3 a and Theorem 4.4 a. Going through the proof of Theorem 4.2 (see (7.8) - (7.11)) we notice that under the assumptions of Theorem 4.3 a and Theorem 4.4 a

$$\frac{1}{n}\tilde{Q}_m(\lfloor ns \rfloor / n) \overset{P}{\to} (\min(\lambda, s)(1-\max(s,\lambda)))^2 \int_{\mathbb{R}^m} |r(u)|^2\, w(u) d\mathbf{u}, \quad s \in (0,1).$$

Therefore

$$k(\gamma) = \min\{m < k < n; \tilde{Q}_m(k/n)/\tilde{q}(k/n,\gamma) = \max_{m<j<n} \tilde{Q}_m(j/n)/\tilde{q}(j/n,\gamma)\}$$

can serve as an estimator of the change point k_0 in either situation. From this we have the consistency of the change-point estimator in the sense that

$$k(\gamma)/n \overset{P}{\to} \lambda$$

follows using standard arguments. Afterwards, one can split the observations into two parts $Y_1,\ldots,Y_{k(\gamma)}$ and $Y_{k(\gamma)+1},\ldots,Y_n$ and apply the test separately to each part in order to check the stability of the model in each part. To study the details though would be quite technical and complex and we will not pursue it here.

5 BOOTSTRAP APPROXIMATIONS FOR THE TEST STATISTICS

The asymptotic distribution of the test statistics as derived in the previous paragraph depends in a complicated matter on the unknown distribution of the observations. Consequently, the asymptotical critical values can neither be calculated, nor estimated or simulated. For this reason a bootstrap approximation cannot be avoided.

There exist several alternative types of resampling in the context of testing the MDH. For example, Horowitz et al. (2006) employ the blocks–of–blocks bootstrap method with autocorrelation–type tests, while in Whang and Kim (2003) the variance–ratio test is implemented by sub–sampling in overlapping time periods; more information on different types of resampling with various MDH tests may be found in Fan and Mills (2009). In the context of our test however we propose to use a version of the wild bootstrap procedure initially suggested by Wu (1986) and Mammen (1993), and which, in situations similar to the present one, has been put on a firm theoretical basis by Domínguez and Lobato (2003) and Escanciano and Velasco (2006a, 2006b). Note that the wild bootstrap procedure has proved to be most effective in finite–sample studies (see Charles et al., 2011b, Fan and Mills, 2009), yielding reliable empirical levels as well as good power across many alternatives, and it is probably for this reason that it is often invoked with real–data applications; see for instance Veka (2013), Todea and Lazăr (2012), Kim et al. (2011), and Charles et al. (2011a).

We assume: [(B.1):]

(B.1) $\{\eta_i\}_i$ are i.i.d. with mean zero, unit variance and $E|\eta_1|^{2+\delta} < \infty$ for some $\delta > 0$,

(B.2) $\{\eta_i\}_i$ and $\{Y_i\}_i$ are independent sequences of random variables.

Then consider

$$S_t^{(m)*}(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{\tau=m+1}^{t} Y_\tau \exp\left(i\mathbf{u}'\mathbf{Y}_{\tau,m} \right) \eta_\tau,$$

and define $T_n^{(j)*}$, $j = 1,2,3$, analogously to $T_n^{(j)}$ with $S_t^{(m)}(\mathbf{u})$ replaced by $S_t^{(m)*}(\mathbf{u})$.

Theorem 5.1 *Let (B.1) and (B.2) be satisfied.*

1. Under the assumptions of Theorem 4.1, i.e., under the null hypothesis, it holds

$$P(T_n^{(j)*} \le x \mid Y_1,\ldots,Y_n) - P(T_n^{(j)} \le x) \xrightarrow{P} 0, \quad j = 1,2,3, \quad x \in \mathbb{R}^1.$$

These assertions remain true even under local alternatives, i.e., under the assumptions of Theorem 4.2c (for $T_n^{(1)}$), Theorem 4.3.c (for $T_n^{(2)}$), or Theorem 4.4c (for $T_n^{(3)}$).

2. Under the assumptions of Theorem 4.2a for all x

$$|P(T_n^{(1)*} \le x \mid Y_1,\ldots,Y_n) - P(\int_{\mathbb{R}^m} |Z^0(0,\mathbf{u})|^2 \, w(\mathbf{u})du \le x)| \xrightarrow{P} 0,$$

where $0 \le \gamma < 1$, $\{Z^0(s,\mathbf{u}), s \in [0,1], \mathbf{u} \in \mathbb{R}^m\}$ is a Gaussian process with expectation zero and covariance ($0 \le s_1 \le s_2 \le 1$)

$$cov\{Z^0(s_1,\mathbf{u}_1), Z(s_2,\mathbf{u}_2)\} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1+m}^{\lfloor ns_1 \rfloor} E(Y_j^2 h(\mathbf{Y}_j,\mathbf{u}_1) h(\mathbf{Y}_j,\mathbf{u}_2)), \quad \mathbf{u}_1, \mathbf{u}_2.$$

3. Under the assumptions of Theorem 4.3a for all x

$$|P(T_n^{(2)*}(\gamma) \mid Y_1,\ldots,Y_n)$$

$$- P(\sup_{0<s<1} \frac{1}{(1-s)^\gamma} \int_{\mathbb{R}^m} |Z^0(s,\mathbf{u}) - Z^0(1,\mathbf{u})|^2 \, w(\mathbf{u})d\mathbf{u} \le x)| \xrightarrow{P} 0,$$

4. Under the assumptions of Theorem 4.4a for all x

$$|P(T_n^{(3)*}(\gamma) \le x \mid Y_1,\ldots,Y_n)$$

$$- P(\sup_{0<s<1} \frac{1}{(s(1-s))^\gamma} \int_{\mathbb{R}^m} |Z^0(s,\mathbf{u}) - sZ^0(1,\mathbf{u})|^2 \, w(\mathbf{u})d\mathbf{u} \le x)| \xrightarrow{P} 0.$$

Due to the stationarity assumption the covariance $cov\{Z^0(s_1,\mathbf{u}_1), Z(s_2,\mathbf{u}_2)\}$ can be expressed more explicitly for all alternatives. Particularly, under the assumptions of Theorem 3.2a we get that this covariance equals

$$s_1\mathbb{E}(\zeta_j^2 h(\mathrm{z}_j,\mathbf{u}_1)h(\mathrm{z}_j,\mathbf{u}_2)),$$

where $\zeta_j = \xi_j + g(\mathrm{x}_j)$, under the assumptions of Theorem 3.3a it equals

$$\min(s_1,\lambda)\mathbb{E}(\xi_j^2 h(\mathrm{x}_j,\mathbf{u}_1)h(\mathrm{x}_j,\mathbf{u}_2)) + (s_1-\lambda)_+\mathbb{E}(\zeta_j^2 h(\mathrm{z}_j,\mathbf{u}_1)h(\mathrm{z}_j,\mathbf{u}_2))$$

and under the assumptions of Theorem 4.4a it equals

$$\min(s_1,\lambda)\mathbb{E}(\tilde{\xi}_j^2 h(\tilde{\mathrm{x}}_j,\mathbf{u}_1)h(\tilde{\mathrm{x}}_j,\mathbf{u}_2)) + (s_1-\lambda)_+\mathbb{E}(\tilde{\zeta}_j^2 h(\tilde{\mathrm{z}}_j,\mathbf{u}_1)h(\tilde{\mathrm{z}}_j,\mathbf{u}_2)),$$

where $\tilde{\xi}_j = \xi_j + c$ and $\tilde{\zeta}_j = \zeta_j + c$.

The above theorem shows that under the null hypothesis the bootstrap critical values are asymptotically equivalent to the asymptotic critical values while they are at least bounded under alternatives. For a fixed alternative the covariance structure of the limit process of $S_t^{(m)*}$ is close to a null hypothesis limit in the sense that the limit process is centered with a similar covariance structure as under alternatives, while it is not centered under alternatives. A more detailed connection can not be made in this context as the fixed alternative is not clearly linked to a specific time series under the null hypothesis since the limit distribution is not pivotal under the null hypothesis. In the contiguous case the bootstrap critical values are close to the critical values for observations ξ following the null hypothesis.

6 SIMULATION AND DATA EXAMPLE

6.1 Simulations

We investigate the small sample properties of the proposed test by Monte Carlo. Earlier small–sample studies for MDH tests were carried out by Lupi (1996), Fan and Mills (2009) and Charles et al. (2011b). Here we compare the new test with the spectral test of Escanciano and Velasco (2006b) which was shown to be one of the most powerful tests in the simulation studies presented by Escanciano and Velasco (2006b) and Charles et al. (2011b). To facilitate comparison we use the same processes as in Escanciano and Velasco (2006b). More precisely, with ε_t and u_t denoting two independent sequences of i.i.d. $N(0,1)$ random variables, we consider the following processes:

IID independent and identically distributed $N(0,1)$ variates.

GARCH(1,1) process $Y_t = \varepsilon_t \sigma_t$ with $\sigma_t^2 = 0.001 + 0.01Y_{t-1}^2 + 0.97\sigma_{t-1}^2$.

stochastic volatility model $Y_t = \varepsilon_t \exp(\sigma_t)$ with $\sigma_t = 0.936\sigma_{t-1} + 0.32u_t$.

non-linear moving average $Y_t = \varepsilon_{t-1}\varepsilon_{t-2}(\varepsilon_{t-2} + \varepsilon_t + 1)$.

bilinear process $Y_t = \varepsilon_t + b_1\varepsilon_{t-1}Y_{t-1} + b_2\varepsilon_{t-1}Y_{t-2}$ with $b_1 = 0.15$ and $b_2 = 0.05$.

bilinear process with $b_1 = 0.25$ and $b_2 = 0.15$.

the sum of a white noise and the first difference of a stationary AR(1) process $Y_t = \varepsilon_t + X_t - X_{t-1}$ with $X_t = 0.85X_{t-1} + u_t$.

threshold autoregressive model $Y_t = -0.5Y_{t-1} + \varepsilon_t$ if $Y_{t-1} \geq 1$ and $Y_t = 0.4Y_{t-1} + \varepsilon_t$ if $Y_{t-1} < 1$.

first order exponential autoregressive model $Y_t = 0.6Y_{t-1}\exp(-0.5Y_{t-1}^2) + \varepsilon_t$.

fractional integrated model ARFIMA$(0, 0.3, 0)$, i.e., $(1-L)^{0.3}Y_t = \varepsilon_t$ with L denoting the usual back shift operator.

The behavior of the non-sequential test statistic $T_n^{(1)}$ is investigated in Table 1 for various values of the tuning parameters a and m for three MDS processes (IID, G1, and SV) and for seven non-MDS processes. The empirical significance level seems to be reasonably close to $\alpha = 0.05$ and the empirical power is largest, in most cases, for $m = 1$ and $a = 1$. This power, although mostly reasonably high, it appears to be low for NDAR and NLMA alternatives and in certain cases with EXP(1). Compared to the state–of–art test by Escanciano and Velasco (2006b, Tables 1–3), the empirical power of our test (with $a = 1$ and $m = 1$) is mostly slightly lower but, on the other hand, the test statistic $T_n^{(1)}$ is computationally very simple and the proposed test may be carried out even for larger sample sizes. We also generalize the proposed test statistic to the change–point setup.

In Tables 2 and 3, we investigate the behavior of the change-point test statistic $T_n^{(2)}(\gamma)$ for $\gamma \in \{0, 0.5\}$. The empirical significance level lies close to $\alpha = 0.05$. Also the power pattern observed in Table 1 mostly persists here too (as well as in Table 4 below), with the tests having reasonable power, but again missing certain types of alternatives such as those involving the NDAR and NLMA, and in certain cases the SV process.

The simulation results concerning the change-point hypothesis $H_0^{(3)}$ are summarized in Table 4. The empirical powers for $H_0^{(3)}$ are slightly lower than empirical powers observed for $H_0^{(2)}$ in Table 2. This behavior is not surprising because the hypothesis $H_0^{(3)}$ is more general than the hypothesis $H_0^{(2)}$. Notice also that the test statistic $T_n^{(3)}(\gamma)$ is computationally more intensive than $T_n^{(2)}(\gamma)$.

As an illustration of the difference between hypotheses $H_0^{(2)}$ and $H_0^{(3)}$ and the corresponding test statistics $T_n^{(2)}(\gamma)$ and $T_n^{(3)}(\gamma)$, we include Table 5 where the process P1 consists of IID N(0,1) observations shifted by μ_1 and the process P2 are IID N(0,1) observations shifted by μ_2. This model with a constant shift is actually a slight simplification of one of the models used in Bao and Lee (2006). Using our method, we test the existence of the change-point k_0 without any prior knowledge concerning its location and we obtain the empirical rejection rates given in Table 5. The hypothesis $H_0^{(2)}$ is satisfied only for $\mu_1 = \mu_2 = 0$ and this is reflected in the left part of Table 5, where we can also see that the test statistics $T_n^{(2)}(\gamma)$ detects also the situation with $\mu_1 \neq 0$ that is not covered by $H_0^{(2)}$. The results for $H_0^{(3)}$ that may be found in the right-hand side of Table 5 show that $H_0^{(3)}$ is satisfied whenever $\mu_1 = \mu_2$, although in this case the corresponding test appears to be conservative somewhat.

6.2 Data Example: S&P 500

In this section we apply the suggested procedures on the log returns of the S&P 500 stock index. Since this is a commonly used index in the context of testing the MDH, we briefly mention some of the earlier findings: Hong and Lee (2005) find strong evidence against the MDH for both the raw data as well as for the residuals after removing linear dependence. On the other hand, Escanciano and Mayoral (2010) use several test statistics, some of them being significant for the MDH while others non-significant. Likewise Escanciano and Velasco (2006b), having apriori subdivided their data–set into three different sub–periods, are led to mixed conclusions with the MDH being accepted for certain periods with specific tests while being rejected for other periods with the same or other tests. The MDH is also rejected for 40% of the individual stocks in the S&P 500, by the test of Kapetanios and Blake (2010). While these findings correspond to different time periods with daily, but also with weekly data, it appears reasonable to suggest that the MDH for this particular series conforms nicely with the adaptive market hypothesis that market efficiency varies over time and profitable opportunities do appear episodically in an intrinsically dynamic fashion. In this connection Bao and Lee (2006), working in the context of density forecasts find that despite the fact that the entire distribution of the S&P 500 series is not predictable, some of its tail characteristics are better predicted by certain non–linear models, thus rejecting the MDH for this part of the distribution. This last finding is compatible with a collection of empirical results suggesting that time periods of predictability appear to coincide with a certain amount of higher market uncertainty and volatility; see for instance Kim et al. (2011), Veka (2013) and Charles et al. (2011a).

Here we shall initially follow the exposition of Escanciano and Velasco (2006b) who analyze three sample periods for log returns of the S&P 500 stock index; see Figure 1. (Note that these three sample periods were determined apriori and without any justification). The authors conclude that the MDH is not rejected for the first period (Jan1990–Dec1993), it is rejected for the second period (Jan1994–Dec1997), while it is questionable for the third period (Jan1998–Aug2002). These findings correspond to the spectral test in Escanciano and Velasco (2006b), and as already noted above vary with other MDH tests. Repeating their analysis using the test statistic (5), we are led to the same conclusion for the first two time periods. Specifically we obtain the p-value 0.476 for the first sample and the p-value 0.005 for the second sample.

Unfortunately, this approach heavily depends on the subdivision of the data set into two different time–periods, which is in fact done arbitrarily. However within this apriori fixed sample framework, one needs to perform repeated testing in order to locate a possible change point, that would eventually result to non–trivial problems which are often encountered with repeated testing. Therefore we now turn to the new test statistic (6) in the context of which the change point is not fixed but instead it is part of the output of the procedure. Specifically when the test for the change-point hypothesis $H_0^{(2)}$ is applied to the joint data set (Jan1990–Dec1997) it leads to a p-value 0.003 ($a=1$, $m=1$, $\gamma=0.5$). Next, proceeding as described in Section 4.2.3, we obtain the change–point estimate $k=1250$ corresponding to a change occurring on December 8th, 1994. Note that this date precedes by almost one year the arbitrarily chosen change-point in Escanciano and Velasco (2006b). Using, once more, the test statistic (5), we obtain p-value 0.649 for data observed until December 7th, 1994, and p-value 0.000 for data observed from December 8th, 1994, which implies that the MDH is not rejected for the first period (Jan1990–Dec7, 1994), while it is rejected for the second perriod (Dec8, 1994–Dec1997). To confirm that there is no further change in the first period we use the statistic (6) to test the change-point hypothesis $H_0^{(2)}$ and obtained a p-value of 0.526.

Hence, using the methods proposed in this paper we arrive at the following conclusions:

1. We confirm the results of Escanciano and Velasco (2006b) that the MDH is not rejected from January 1990 until December 1993, and it is rejected from January 1994 until December 1997.

2. The hypothesis $H_0^{(2)}$ of no change in the martingale difference structure between January 1990 and December 1997 is rejected with overall type I error equal to 0.05. However the change in the martingale difference structure of the S&P 500 log returns occurred in December 1994, almost one year later than the change-point considered previously in Escanciano and Velasco (2006b). This finding is corroborated by the fact that the MDH $H_0^{(1)}$ is not rejected for log returns until December 7th, 1994, and it is rejected for log returns observed after December 8th, 1994. As a further confirmation we note that the hypothesis $H_0^{(2)}$ of no change in the martingale difference structure is not rejected using the data between January 1990 and December 7th, 1994.

In order to further investigate possible differences in the two time periods which were determined by our change–point analysis we have computed several descriptive statistics corresponding to the S & P 500 data in these two periods. While most of these statistics were somewhat similar, there seems to be a considerable difference in the kurtosis of the data with the excess kurtosis being equal to 2.17 for the first period (Jan1990–Dec7, 1994) and equal to 9.21 for the second period (Dec8, 1994–Dec1997). Hence there appears to be a significantly different tail–behavior before and after the change–point, and although this is certainly not conclusive evidence it would be reasonable to suggest that in the second time period the stock market was more volatile, a fact that as already mentioned has been associated to market predictability and the rejection of the MDH.

7 PROOFS

In order to prove Theorem 4.1 the following lemma is essential.

Lemma 7.1 *Let (16) be satisfied and let* $\{Y_t\}$ *be a martingale difference sequence as well as stationary and ergodic with* $\mathbb{E}|Y_1|^{2+\delta} < \infty$ *for some* $\delta > 0$ *and define*

$$Z_n(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{k=m+1}^{\lfloor sn \rfloor} Y_k h(\mathbf{Y}_k, \mathbf{u}), \quad \mathbf{u} \in \mathbb{R}^m, \quad s \in (0,1)$$

where $h(\mathbf{Y}_k, \mathbf{u})$ is defined in (17), $\mathbf{Y}_k = (Y_{k-1}, \dots, Y_{k-m})^T$ and $\mathbf{u} = (u_1, \dots, u_m)^T$. Then [a)]

1. For any compact subset of $\mathbb{R}^m$ and any $0 \le s \le 1$ it holds

$$\sup_n \mathbb{E} \int_F (Z_n(s,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u} < \infty \qquad (7.1)$$

2. There exists an $a > 0$, $0 < D < \infty$ such that for any $0 \le s \le 1$ it holds

$$\sup_n \mathbb{E} \left| Z_n^2(s,\mathbf{u}_1) - Z_n^2(s,\mathbf{u}_2) \right| \le D \| \mathbf{u}_1 - \mathbf{u}_2 \|^a \qquad (7.2)$$

3. The marginal distributions of $\{Z_n(s,\mathbf{u})\}$ converge to the marginal distributions of a Gaussian process $\{Z(s,\mathbf{u})\}$ with covariance structure ($0 \le s_1 \le s_2 \le 1$)

$$cov\{Z(s_1,\mathbf{u}_1),Z(s_2,\mathbf{u}_2)\} = s_1 E(Y_{m+1}^2 h(\sum_{q=1}^{m} u_{q1}Y_{m-q})h(\sum_{v=1}^{m} u_{q2}Y_{m-v})), \quad \mathbf{u}_1,\mathbf{u}_2.$$

Proof. For each $\mathbf{u}$ the process $Z_n(s,\mathbf{u})$ is a sum of martingale differences so that

$$\mathbb{E}Z_n(s,\mathbf{u}) = 0,$$

$$\mathbb{E}\{Z_n(s,\mathbf{u})\}^2 = \frac{1}{n}\sum_{k=m+1}^{\lfloor ns \rfloor}\mathbb{E}(Y_k^2 h^2(\mathbf{Y}_k,\mathbf{u})) \le 4\mathbb{E}Y_{m+1}^2 \tag{7.3}$$

as $h(.)$ is bounded and $\{Y_k\}$ is stationary. From this, assertion a) follows.

In the following we will use D as a generic constant which may vary from line to line. First notice that by the boundedness of sine and cosine and the mean value theorem it holds

$$\mathbb{E}\,|\,h(\mathbf{Y}_k,\mathbf{u}_1)-h(\mathbf{Y}_k,\mathbf{u}_2)\,|^{2(2+\delta)/\delta} \le D\mathbb{E}\,|\,h(\mathbf{Y}_k,\mathbf{u}_1)-h(\mathbf{Y}_k,\mathbf{u}_2)\,|^{\min(2+\delta,2(2+\delta)/\delta)}$$

$$\le D\mathbb{E}\,|\langle \mathbf{Y}_k,\mathbf{u_1}-\mathbf{u_2}\rangle|^{\min(2+\delta,2(2+\delta)/\delta)} \le D\,||\,\mathbf{u_1}-\mathbf{u_2}\,||^{\min(2+\delta,2(2+\delta)/\delta)},$$

hence by the Hölder inequality and (24) it follows

$$\mathbb{E}\,|\,Z_n^2(s,\mathbf{u}_1)-Z_n^2(s,\mathbf{u}_2)\,| \le (\mathbb{E}\,|\,Z_n(s,\mathbf{u}_1)-Z_n(s,\mathbf{u}_2)\,|^2\,\mathbb{E}\,|\,Z_n(s,\mathbf{u}_1)+Z_n(s,\mathbf{u}_2)\,|^2)^{1/2}$$

$$\le D(\mathbb{E}\,|\,Z_n(s,\mathbf{u}_1)-Z_n(s,\mathbf{u}_2)\,|^2)^{1/2} = D\left(\frac{1}{n}\mathbb{E}(\sum_{k=m+1}^{\lfloor sn \rfloor} Y_k(h(\mathbf{Y}_k,\mathbf{u}_2)-h(\mathbf{Y}_k,\mathbf{u}_1))^2\right)^{1/2}$$

$$\le D\left(\frac{1}{n}\sum_{k=m+1}^{n}\mathbb{E}\left[(Y_k)^2(h(\mathbf{Y}_k,\mathbf{u}_2)-h(\mathbf{Y}_k,\mathbf{u}_1))^2\right]\right)^{1/2}$$

$$\le \left(\frac{1}{n}\sum_{k=m+1}^{n}(\mathbb{E}\,|\,Y_k\,|^{2+\delta})^{2/(2+\delta)}(\mathbb{E}\,|\,h(\mathbf{Y}_k,\mathbf{u}_1)-h(\mathbf{Y}_k,\mathbf{u}_2)\,|^{2(2+\delta)/\delta})^{\delta/(2+\delta)}\right)^{1/2}$$

$$\le D\,||\,\mathbf{u_1}-\mathbf{u_2}\,||^a$$

for some $a > 0$, hence b).

Concerning c) we apply Theorem 6.21, p.113, in Breiman (1968). We need to verify the assumptions. Since $\{Y_j\}$ is a stationary ergodic sequence it holds

$$\frac{1}{n}\sum_{j=1}^{n} h^2(\mathbf{Y}_j,\mathbf{u})\mathbb{E}(Y_j^2 \mid Y_{j-1},\ldots) \xrightarrow{P} v = Eh^2(\mathbf{Y}_j,\mathbf{u})Y_j^2,$$

it holds even a.s. Additionally, for all $\varepsilon > 0$

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(h^2(\mathbf{Y}_j,\mathbf{u})\mathbb{E}(Y_j^2 1_{\{|h(\mathbf{Y}_j,\mathbf{u})Y_j|>\varepsilon\sqrt{n}\}} \mid Y_{j-1},\ldots) \le \frac{1}{\varepsilon^{\delta} n^{1+\delta/2}}\sum_{j=1}^{n} |Y_j|^{2+\delta} \xrightarrow{P} 0.$$

Proof of Theorem. By assumption (16) we have

$$\int_{\mathbb{R}^m} \cos\left(\sum_{q=1}^{m} u_q x_q\right) \sin\left(\sum_{q=1}^{m} u_q x_q\right) w(\mathbf{u})d\mathbf{u} = 0,$$

which immediately implies

$$T_n^{(1)} = \frac{1}{n}\int_{\mathbb{R}^m}\left(\sum_{k=1+m}^{n} Y_k\left(\sin\left(\sum_{q=1}^{m} u_q Y_{k-q}\right) + \cos\left(\sum_{q=1}^{m} u_q Y_{k-q}\right)\right)\right)^2 w(\mathbf{u})d\mathbf{u}. \tag{7.4}$$

From Lemma 7.1 ($s=1$) in addition to Ibragimov and Chasminskij (1981), Theorem 22 (pages 380, 381) we get that

$$\int_F (Z_n(1,\mathbf{u}))^2 w(\mathbf{u})d\mathbf{u} \xrightarrow{\mathcal{L}} \int_F (Z(1,\mathbf{u}))^2 w(\mathbf{u})d\mathbf{u} \qquad (7.5)$$

for any compact subset F of $\mathbb{R}^m$. Since $w(\cdot)$ is integrable there exist for all $\eta > 0$ a compact set F_η such that

$$\mathbb{E}\int_{\mathbb{R}^m \setminus F_\eta} (Z_n(1,\mathbf{u}))^2 w(\mathbf{u})d\mathbf{u} \le D\int_{\mathbb{R}^m \setminus F_\eta} w(\mathbf{u})d\mathbf{u} < \eta, \quad \text{forall} n$$

and an analogous argument if Z_n is replaced by Z which together with (26) show

$$\int_{\mathbb{R}^m} (Z_n(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u} \xrightarrow{\mathcal{L}} \int_{\mathbb{R}^m} (Z(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}$$

proving a).

To prove b) we consider the process

$$X_n(s) = \sqrt{\int_{\mathbb{R}^m} (Z_n(s,\mathbf{u}) - Z_n(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}}, \quad s \in (0,1).$$

First, we prove the convergence of the finite dimensional distribution. To this end let $0 < s_1 < \ldots < s_r < 1$ and $b_1, \ldots, b_r \in \mathbb{R}$. Then analogously to the proof of a) we obtain

$$\int_{R^m} \sum_{r=1}^{\ell} b_r (Z_n(s_r,\mathbf{u}) - Z_n(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u} \xrightarrow{\mathcal{L}} \int_{R^m} \sum_{r=1}^{\ell} b_r (Z(s_r,\mathbf{u}) - Z(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u},$$

which by the continuous mapping theorem shows the convergence of the finite dimensional distributions of $X_n(\cdot)$ towards those of $\sqrt{\int_{\mathbb{R}^m} (Z(s,\mathbf{u}) - Z(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}}$.

To obtain tightness, we get by the Minkowski inequality

$$| X_n(s) - X_n(t) | \le D \sqrt{\int | Z_n(s,\mathbf{u}) - Z_n(t,\mathbf{u}) |^2 \, w(\mathbf{u}) d\mathbf{u}}.$$

Hence, by the Jensen inequality, for any $\varepsilon > 0$,

$$P(| X_n(s) - X_n(t) | \ge \varepsilon) \le P\left(D \int | Z_n(s,\mathbf{u}) - Z_n(t,\mathbf{u}) |^2 \, w(\mathbf{u}) d\mathbf{u} \ge \varepsilon \right)$$

$$\le DE\left(\int | Z_n(s,\mathbf{u}) - Z_n(t,\mathbf{u}) |^2 \, w(\mathbf{u}) d\mathbf{u} \right)^{1+\delta/2}$$

$$\le D \int E | Z_n(s,\mathbf{u}) - Z_n(t,\mathbf{u}) |^{2+\delta} \, w(\mathbf{u}) d\mathbf{u} \le D | s - t |^{1+\delta/2},$$

where the last line follows from Stout (1974), Theorem 3.7.8. By Billingsley (1968), Theorem 15.6., we obtain that

$$X_n(\cdot) \xrightarrow{D[0,1]} \sqrt{\int_{\mathbb{R}^m} (Z(\cdot,\mathbf{u}) - Z(1,\mathbf{u}))^2 w(\mathbf{u}) d\mathbf{u}}, \tag{7.6}$$

which concludes the proof of b) for $\gamma = 0$. For $0 < \gamma < 1/2$ it holds for arbitrary

$0 < a < 1/2$

$$\max_{0<s<1-a} Q_m(s)/(1-s)^\gamma \xrightarrow{\mathcal{L}} \sup_{0<s<1-a} \frac{1}{(1-s)^\gamma} \int_{\mathbb{R}^m} |Z(s,\mathbf{u}) - Z(1,\mathbf{u})|^2 \, w(\mathbf{u}) d\mathbf{u},$$

Hence it remains to study $\max_{1-a<s<1} Q_m(s)/(1-s)^\gamma$. This is equivalent to treating

$$\max_{(1-a)n<k<n} \frac{1}{((n-k)/n)^\gamma} \int \Big(\frac{1}{\sqrt{n}} \sum_{j=k+1}^{n} Y_j h(\mathbf{Y}_j, \mathbf{u})\Big)^2 w(\mathbf{u})\mathbf{u}. \tag{7.7}$$

By Stout (1974), Theorem 3.7.8. it holds

$$E \Big| \sum_{j=k+1}^{n} Y_j h(\mathbf{Y}_j, \mathbf{u}) \Big|^{2+\Delta} \le D(n-k)^{1+\Delta/2}$$

where $D > 0$ is a generic constant depending on neither k nor $\mathbf{u}$.

Therefore by Theorem B.3 (p. 184) in Kirch (2006) it holds

$$E\Big(\max_{(1-a)n<k<n} ((n-k)/n)^{-\gamma} \Big| \frac{1}{\sqrt{n}} \sum_{j=k+1}^{n} Y_j h(\mathbf{Y}_j, \mathbf{u}) \Big|^2\Big)^{(2+\Delta)/2}$$

$$\le D \sum_{k=n(1-a)}^{n} n^{(\gamma-1)(2+\Delta)/2} (n-k)^{-\gamma(2+\Delta)/2} (n-k)^{1+\Delta/2-1}$$

$$\le D n^{(\gamma-1)(2+\Delta)/2} \sum_{j=1}^{an} j^{\Delta/2-\gamma(2+\Delta)/2}$$

$$\le D(an)^{(1-\gamma)(2+\Delta)/2} n^{(\gamma-1)(2+\Delta)/2} \le D a^{(1-\gamma)(2+\Delta)/2}$$

Since $(1-\gamma)(2+\Delta)/2$ is positive choosing $a > 0$ small enough also the right side of the last expression can be made small. Therefore we have that choosing $a > 0$ small also (28) is small in probability. We can proceed similarly with limit process

$$\sup_{0<s<1} \frac{1}{(1-s)^{\gamma}} \int_{\mathbb{R}^m} |Z(s,\mathbf{u}) - Z(1,\mathbf{u})|^2 \ w(\mathbf{u})d\mathbf{u}.$$

Then we can conclude that the assertion b) holds even for $0 < \gamma < 1$.

It remains to show that the assertion of Theorem 4.1 c). The proof is quite parallel to the case of Theorem 4.1 b) and therefore omitted.

Proof of Theorem 4.2. We will use a more general formulation than is needed for Theorem 4.2 which will be of use in the proof of Theorems 4.3 and 4.4. To this end consider an arbitrary but fixed $s \in (0,1]$.

The sequence $\{\xi_k\}$ satisfies the assumptions of Theorem 3.1a) and we can write

$$\begin{aligned} Z_n(s,u) &= \frac{1}{\sqrt{n}} \sum_{k=m+1}^{sn} \xi_k h(\mathrm{x}_k, u) + \frac{d_n}{\sqrt{n}} \sum_{k=m+1}^{sn} g(\mathrm{x}_k) h(\mathrm{x}_k, u) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{k=m+1}^{sn} \xi_k (h(\mathbf{Y}_k, u) - h(\mathrm{x}_k, u)) \\ &\quad + \frac{d_n}{\sqrt{n}} \sum_{k=m+1}^{sn} g(\mathrm{x}_k)(h(\mathbf{Y}_k, u) - h(\mathrm{x}_k, u)) \\ &= J_{n1}(s,\mathbf{u}) + J_{n2}(s,u) + J_{n3}(s,u) + J_{n4}(s,u) \end{aligned}$$

By Theorem 4.1 a) it holds

$$\int J_{n1}^2(s,\mathbf{u}) w(\mathbf{u}) d\mathbf{u} = O_P(1). \tag{7. 8}$$

Next we deal with $J_{n2}(s,\mathbf{u}), J_{n3}(s,\mathbf{u}), J_{n4}^2(s,\mathbf{u})$ for the fixed alternative in a), i.e. $d_n = 1$.

By the uniform ergodic theorem of Ranga Rao (1962) we get that

$$\frac{1}{n} \sup_{u \in K} |J_{n3}^2(s,u)| \xrightarrow{P} 0 \tag{7.9}$$

for any compact subset K. Together with (16) and because $\mathbb{E}|n^{-1}J_{n,3}^2(s,u)| \le D$ uniformly in s and u this shows that

$$\frac{1}{n}\int J_{n3}^2(s,u)w(u)du \xrightarrow{P} 0. \tag{7.10}$$

Analogously it holds

$$\frac{1}{n}\int (J_{n2}(s,\mathbf{u})+J_{n4}(s,\mathbf{u}))^2 w(\mathbf{u})d\mathbf{u} \xrightarrow{P} s\int \left| \mathbb{E}g(\mathrm{x}_{n+1})h(\mathbf{Y}_{n+1},u) \right|^2 w(\mathbf{u})d\mathbf{u} > 0. \tag{7.11}$$

Putting (29) – (32) together proves assertion a).

For local alternatives we get similar to (32) that

$$\frac{1}{d_n^2 n}\int_{\mathbb{R}^m} \left| J_{n2}(s,u) \right|^2 w(u)du \xrightarrow{P} s\int |r(u)|^2 w(u)du \ne 0, \tag{7.12}$$

Similarly to (29) we get for local alternatives with $d_n\sqrt{n} \to b \to 0$

$$\int_{\mathbb{R}^m} \left| J_{n1}(s,u)+J_{n2}(s,u) \right|^2 w(u)du \xrightarrow{\mathcal{L}} s\int |Z(0,u)+br(u)|^2 du. \tag{7.13}$$

To show the negligibility of J_{n3} and J_{n4} for local alternatives, first note that for any $0<\kappa<4$ and for any $\varepsilon>0$ there exists a compact set $K \subset \mathbb{R}^m$ (depending on ε) such that by (16) and the mean value theorem

$$\begin{aligned}
&\int_{\mathbb{R}^m} |h(Y_k,u)-h(\xi_k,u)|^2 w(u)du \le D\int_K |h(Y_k,u)-h(\xi_k,u)|^{\kappa/2} w(u)du+\varepsilon \\
&\le Dd_n^{\kappa} \,\|(g(\xi_{k-1}),\ldots,g(\xi_{k-m}))^T\|^{\kappa} \int_K \|u\|^{\kappa} w(u)du+\varepsilon \\
&\le d_n^{\kappa} D_K \,\|(g(\xi_{k-1}),\ldots,g(\xi_{k-m}))^T\|^{\kappa}+\varepsilon.
\end{aligned}$$

This yields by the Cauchy-Schwarz inequality and the martingale difference property of $\{\xi_k(h(Y_k,u)-h(\xi_k,u))\}$

$$\mathbb{E}\left[\int_{\mathbb{R}^m} J_{n3}^2(s,u)w(u)du\right] = \int_{\mathbb{R}^m}\mathbb{E}\left[\xi_m (h(Y_m,u)-h(\xi_m,u))\right]^2 w(u)du$$

$$= \mathbb{E}\left[\xi_m^2 \int_{\mathbb{R}^m}(h(Y_m,u)-h(\xi_m,u))^2 w(u)du\right]$$

$$\leq \left(\mathbb{E}|\xi_m|^{2+\delta}\right)^{2/(2+\delta)}\left(\mathbb{E}\left[\int_{\mathbb{R}^m}(h(Y_m,u)-h(\xi_m,u))^2 w(u)du\right]^{(2+\delta)/\delta}\right)^{\delta/(2+\delta)}$$

$$\leq o(1)+D\varepsilon,$$

which in particular shows that

$$\int_{\mathbb{R}^m} J_{n3}^2(s,u)w(u)du = o_P(1). \tag{7.14}$$

Similarly, one gets by the Cauchy-Schwarz inequality

$$\mathbb{E}\frac{1}{d_n^2 n}\int_{\mathbb{R}^m} J_{n4}^2(s,u)w(u)du = \int_{\mathbb{R}^m}\mathbb{E}\left(\frac{1}{n}\sum_{k=m+1}^{sn} g(\xi_k)(h(Y_k,u)-h(\xi_k,u))\right)^2 w(u)du$$

$$\leq \int_{\mathbb{R}^m}\mathbb{E}\left[g(\xi_m)(h(Y_m,u)-h(\xi_m,u))\right]^2 w(u)du \leq o(1)+D\varepsilon,$$

hence

$$\frac{1}{d_n^2 n}\int_{\mathbb{R}^m} J_{n4}^2(s,u)w(u)du = o_P(1). \tag{7.15}$$

Together with (33), (34) and (35), this implies the assertion for local changes.

Proof of Theorem 4.3. To show a) and b) it suffices to show that as $n\to\infty$

$$Q_m(k_0/n) = \int_{\mathbb{R}^m}|S_{k_0}^{(m)}(u)-S_n^{(m)}(u)|^2\, w(u)du \xrightarrow{P} \infty.$$

with $k_0 = \lfloor n\lambda \rfloor$, $\lambda \in (0,1)$ and we can proceed as in the proof of Theorem 4.2 treating

$$Z_n(1,\mathbf{u}) - Z_n(k_0/n,\mathbf{u})$$

instead of $Z_n(s,u)$. Therefore the rest of the proof is omitted. The proof of Theorem 4.3 c) follows the lines of Theorem 4.1 c) and is skipped also. This completes the proof.

Proof of Theorem 4.4. The proof is quite parallel to the proofs of Theorem 4.2 and 4.3 and therefore is omitted.

Proof of Theorem 5.1. We follow the line of the proof of Theorem 4.1, but have to get the properties of

$$Z_n^*(s,u) = \frac{1}{\sqrt{n}} \sum_{k=m+1}^{\lfloor ns \rfloor} Y_k h(Y_k,u)\eta_k, \quad s \in (0,1), \quad u \in \mathbb{R}^m$$

given $\{Y_j\}_j$, instead of $Z_n(s,u), u \in \mathbb{R}^m$. The present situation is slightly simpler since, given $\{Y_j\}_j$, $Z_n^*(s,u)$ is the sum of independent variables with conditional mean zero and conditional variance

$$var\{Z_n^*(s,u) \,|\, \{Y_j\}_j\} = \frac{1}{n} \sum_{j=m+1}^{\lfloor ns \rfloor} Y_k^2 h^2(Y_k,u).$$

Further calculations give also

$$\sum_{k=m+1}^{\lfloor ns \rfloor} E(| \frac{1}{\sqrt{n}} Y_k h(Y_k,u)\eta_k |^{2+\delta} | \{Y_j\}_j) = \frac{1}{n^{(2+\delta)/2}} \sum_{k=m+1}^{\lfloor ns \rfloor} | Y_k h(Y_k,u) |^{2+\delta} \, E \,|\, \eta_1 |^{2+\delta} .$$

First, we prove the result under the null hypothesis as well as alternative A1a. Going through the proof of Lemma 7.1 we realize that conditional versions of a) and b) follow by the law of large numbers for stationary and ergodic sequences in combination with the non-conditional results, which hold not only for the null hypothesis but also for alternatives A1a as can be checked easily.

Hence, the crucial problem is to prove a conditional version of Lemma 7.1 c) which is the conditional asymptotic normality of $Z_n^*(s,u)$. We show that, as $n \to \infty$,

$$\frac{\frac{1}{n^{(2+\delta)/2}}\sum_{k=m+1}^{\lfloor ns\rfloor}|Y_k h(Y_k,u)|^{2+\delta}}{(\frac{1}{n}\sum_{j=m+1}^{\lfloor ns\rfloor}Y_k^2h^2(Y_k,u))^{(2+\delta)/2}}\to 0, \quad a.s. \qquad (7.16)$$

that ensures validity of the Lyapunov type condition for CLT. Under H_0 and alternative A1a using stationarity and ergodicity of $\{Y_j\}_j$ we obtain

$$\frac{1}{n}\sum_{k=m+1}^{\lfloor ns\rfloor}|Y_k h(Y_k,u)|^{2+\delta}\to sE|Y_k h(Y_k,u)|^{2+\delta}, \quad a.s.,$$
$$\frac{1}{n}\sum_{j=m+1}^{\lfloor ns\rfloor}Y_k^2h^2(Y_k,u)\to sEY_k^2h^2(Y_k,u), \quad a.s.$$

By the assumptions $E|Y_k h(Y_k,u)|^{2+\delta}<\infty$ and $0<EY_k^2h^2(Y_k,u)$. Therefore (37) holds true given $\{Y_j\}_j$.

Here the conditional limiting process is $Z^0((s,u),$ is Gaussian process with zero mean and covariance structure

$$cov(Z^0((s_1,u_1),Z^0(s_2,u_2))=\lim_{n\to\infty}\frac{1}{n}\sum_{k=m+1}^{\lfloor ns\rfloor}EY_t^2h(Y_t,u_1)h(Y_t,u_2), \quad s_1\le s_2,u_1,u_2.$$

Under the null hypothesis this is the same process as in Theorem 4.1 – under the alternative A1a it is similar but not the same. Then we proceed along the line of the proof of Theorem 4.1 where we again use stationarity and ergodicity of $\{Y_k\}_k$.

The assertions under A2a and A3a can be derived analogously splitting all relevant sums into the part before the change and the part after, since both are sums over stationary and ergodic sequences.

ACKNOWLEDGEMENT

The authors wish to sincerely thank the AE and two anonymous referees for the constructive criticism that led to substantial improvement of the presentation of this work. The work of Zdeněk Hlávka was supported by grant GACR P201/12/0083. The work of Marie Hušková was supported by grant GACR 201/09/0755. The work of Simos Meintanis was supported by research grant no. 11699 of the Special Account for Research Grants of the National and Kapodistrian University of Athens.

REFERENCES

BAO, Y., and LEE, T.H.: Asymmetric predictive abilities of nonlinear models for stock returns: Evidence from density forecast comparison. In T.B. Fomby and D. Terrell (Eds.) *Advances in Econometrics, Elsevier Ltd., 2006, Vol 20, 41–62.*

BELAIRE–FRANCH, J. and CONTRERAS, D.: Testing the martingale property of exchange rates: a replication. *Stud.Nonlin.Dynam.Econometr. 15 (2011) 1–-17*

BIERENS, H.J.: Consistent model specification tests. *J. Econometr., 20 (1982), 105–134.*

BIERENS, H.J. and PLOBERGER, W.: Asymptotic theory of integrated conditional moment test. *Econometrica, 65 (1997), 1129–1151.*

BIERENS, H.J. and WANG, L.: Integrated conditional moment tests for parametric conditional distributions. *Econometr.Theor., 28 (2012), 328–362.*

BILLINGSLEY, P. (1968): Convergence of Probability Measures. Wiley and Sons.

BREIMAN, L. (1968): Probability. Addison–Wesley. and Sons.

CHARLES A., DARNÉ, O., and FOUILLOUX, J.: Testing the martingale difference hypothesis in CO_2 emission allowances . *Econom.Modell.* 28 (2011a), 27–53.

CHARLES A., DARNÉ, O., and KIM, J.H.: Small sample properties of alternative tests for the martingale difference hypothesis. *Econom.Lett.* 110 (2011b), 151–154.

CHOI, I.: Testing the random walk hypothesis for exchange rates. *J.Appl.Econom.* 14 (1999), 293-308.

DASKALAKIS G., PSYCHOYIOS, D., and MARKELLOS, R.N.: Modelling CO_2 emission allowance prices and derivative: Evidence from the European trading scheme . *J.Bank.Finan.* 33 (2009), 1230–1240.

DE JONG, R.M.: Bierens' test under data dependence. *J. Econometr., 72 (1996), 1–32.*

DEO, R.S.: Spectral tests of the martingale hypothesis under conditional heteroscedasticity. *J. Econometr. 99 (2000), 291–315.*

DOMÍNGUEZ, M.A. and LOBATO, I.N.: Testing the martingale difference hypothesis. *Econometr.Rev. 22 (2003) 351–-377.*

DURLAUF, S.N.: Spectral based testing of martingale hypothesis. *J.Econometr.* 50 (1991), 355–376.

EPPS, T.W.: Testing that a stationary time series is Gaussian. *Ann. Statist., 15 (1987), 1683–1698.*

EPPS, T.W.: Testing that a Gaussian process is stationary. *Ann. Statist., 16 (1988), 1667–1683.*

ESCANCIANO, J.C., and LOBATO, I.N.: Testing the martingale hypothesis. In T.C. Mills and K. Patterson (Eds.) *Palgrave Handbook of Econometrics, MacMillan, Palgrave, 2009a, Vol 2, 972–1003.*

ESCANCIANO, J.C. and LOBATO, I.N.: An automatic Portmanteau test for serial correlation. *J.Econometr. 151 (2009b) 140–-149.*

ESCANCIANO, J.C., and MAYORAL, S.: Data–driven smooth tests for the martingale difference hypothesis. *Comput. Statist. Dat. Anal., 54 (2010), 1983–1998.*

ESCANCIANO, J.C. and VELASCO, C.: Testing the martingale difference hypothesis using integrated regression functions. *Comput.Statist.Dat.Anal. 51 (2006a) 2278–-2294*

ESCANCIANO, J.C. and VELASCO, C.: Generalized spectral tests for the martingale difference hypothesis. *J.Econometr. 134 (2006b) 151–-185*

ESCANCIANO, J.C.: On the lack of power of omnibus specification tests. *Econometric Theory 25 (2009) 162–-194*

FAMA, E.: Efficient capital markets: a review of theory and empirical work. *J. Finan., 25 (1970), 383–417.*

FAN, L., and MILLS, T.C.: Size and power properties of tests of the martingale difference hypothesis: a Monte Carlo study. *Inter. J. Econom. Econometr., 1 (2009), 48–63.*

FEUERVERGER, A.: An efficiency result for the empirical characteristic function in stationary time–series models. *Canad. J. Statist., 18 (1990), 155–161.*

FONG W. and OULIARIS, S.: Spectral tests of the martingale hypothesis for exchange rates. *J.Appl.Econometr.* 10 (1995), 255–271.

FONG W., KOH, S., and OULIARIS, S.: Joint variance–ratio test of the martingale hypothesis for exchange rates. *J.Buss.Econom.Statist.* 15 (1997), 51–59.

FRANCQ C., ROY, R., and ZAKOIAN, J.-M.: Diagnostic checking in ARMA models with uncorrelated errors. *J.Amer.Statist.Assoc.* 13 (2005), 532–544.

GHOSH, S.: Normality testing for a long-memory sequence using the empirical moment generating function. *J. Statist. Plann. Infer.* 143 (2013), 944–954.

GROSSMAN, S.J. and STIGLITZ, J.E.: On the impossibility of informationally efficient markets. *Amer.Econ.Rev. 70 (1980), 393–408.*

HENZE N., KLAR, B., and ZHU, L.X.: Checking the adequacy of the multivariate semiparametric location shift model. *J.Multivar.Anal.* 93 (2005), 238–256.

HLÁVKA, Z., HUŠKOVÁ, M., KIRCH, C., and MEINTANIS, S.G.: Monitoring changes in the error distribution of autoregressive models based on Fourier methods. *Test* 21 (2012), 605–634.

HONG, Y.: Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *J. Amer. Statist. Assoc., 94 (1999), 1201–1220.*

HONG Y. and LEE, T.H.: Inference on predictability of foreign exchange rate changes via generalized spectrum and nonlinear time series models. *Rev.Econom.Statist.* 85 (2003), 1048–1062.

HOROWITZ J.L., LOBATO, I.N., NANKERVIS, J.C. and SAVIN, N.E.: Bootstrapping the Box–Pierce Q test: a robust test of uncorrelatedness. *J.Econometr.* 133 (2006), 841–862.

HORVÁTH, L., and KOKOSZKA, P.: The effect of long–range dependence on change–point estimators. *J.Statist.Plann.Infer.* 64 (1997), 57–81.

IBRAGIMOV, I. and CHASMINSKIJ, R. (1981): Statistical Estimation; Asymptotic Theory. Springer Verlag, New York.

KAPETANIOS, G., and BLAKE, A.P.: Tests of the martingale difference hypothesis using boosting and RBF neural network approximations. *Econometr. Theor., 26 (2010), 1363–1397.*

KIM, J.H.: Automatic variance ratio test under conditional heteroscedasticity. *Finan.Resear.Lett.* 6 (2009), 179-185.

KIM J.H., SHAMSUDDIN, A., and LIM, K.P.: Stock return predictability and the adaptive markets hypothesis: Evidence from century–long U.S. data. *J.Empir.Finan.* 18 (2011), 868–879.

KIRCH, C. (2006): Resampling Methods for the Change Analysis of Dependent Data. Dissertation, Cologne, Germany.

KUAN, C.M., and LEE, W.M.: A new test of the martingale difference hypothesis. *Stud. NonLinear. Dynam. Econometr., 8 (2004), Issue 4, Article 1.*

LAVERGNE, P. and PATILEA, V.: Smooth minimum distance estimation and testing with conditional estimating equations: Uniform in bandwidth theory. *J.Econometr. 177 (2013) 47–-59.*

LEUCHT, A.: Characteristic function–based tests under weak dependence. *J.Multivar.Anal.* 108 (2012), 67-89.

LJUNG, G.M. and BOX, G.E.P.: On a measure of lack of fit in time series models. *Biometrika 65 (1978) 297–-303.*

LO, A.W.: The adaptive markets hypothesis. *J.Potrfol.Manag.* 30 (2004), 15-29.

LO, A.W. and MACKINLAY, A.C.: Stock market prices do not follow random walks: evidence from a simple specification test. *Rev.Finan.Stud.* 1 (1988), 41-66.

LOBATO I.N., NANKERVIS, J.C., and SAVIN, N.E.: Testing for zero autocorrelation in the presence of statistical dependence. *Econometr.Theor.* 18 (2002), 730–743.

LUPI, C.: A Monte Carlo analysis of two spectral tests of the martingale hypothesis. *J. Ital.Statist.Soc. 5 (1996), 335–360.*

MAMMEN, E.: Bootstrap and wild bootstrap for high–dimensional models. *Ann.Statist.* 21 (1993), 255–285.

MEINTANIS, S.G.: Comments on: An updated review of Goodness-of-Fit tests for regression models. *Test* 22 (2013), 432-436.

QUESSY, J.-F., and ÉTHIER, F.: Cramér–von Mises and characteristic function tests for the two and k–sample problems with dependent data. *Comput.Statist.Dat.Anal.* 56 (2012), 2097–2111.

RAO, R.: Relations between weak and uniform convergence of measures with applications *Ann.Math.Stat.* 33 (1962), 659-680

SAMUELSON, P.: Proof that properly anticipated prices fluctuate randomly. *Industr.Manag.Rev.* 6 (1965), 41-49.

STOUT, W.F. (1974): Almost Sure Convergence, Academic Press, New York.

TENREIRO, C.: On the choice of the smoothing parameter for the BHEP goodness–of–fit test. *Comput.Statist.Dat.Anal.* 53 (2009), 1038-1053.

TODEA, A. and LAZ Ă R, D.: Global crisis and relative efficiency: empirical evidence from central and eastern European stock markets. *Review of Finance and Banking* 4 (2012), 45–53.

VEKA, S.: Testing the martingale difference hypothesis for the Nordic power derivatives market. *J. Energy Markets , 6 (2013), 141–157.*

WHANG, Y. and KIM, J.: A multiple variance ratio test using subsampling. *Econom. Lett.* 79 (2003), 225–230.

WU, C.–F.J.: Jackknife, bootstrap and other resampling methods in regression analysis. *Ann.Statist.* 14 (1986), 1261–1350.

YILMAZ, K.: Martingale property of exchange rates and central bank interventions. *J.Buss.Econom.Statist. 21 (2003), 383–395.*

Table 1 Percentage of rejection of martingale difference hypothesis ($H_0^{(1)}$). 1000 simulations with 1000 wild bootstrap replicates, $\alpha = 0.05$.

			IID	G1	SV	NLMA	BIL-I	BIL-II	NDAR	TAR(1)	EXP(1)	ARFIMA
		m=1	4.9	5.1	2.1	10.4	18.0	44.2	2.7	68.8	47.5	80.3
	a=0.5	m=2	4.0	4.6	1.1	3.2	11.4	30.9	1.1	47.2	31.3	74.7
		m=1	5.5	5.6	3.9	6.5	18.9	46.8	1.9	68.1	37.3	79.6
	a=1	m=2	4.2	4.9	2.8	4.3	14.8	42.3	2.0	56.7	29.9	78.1
		m=1	4.8	5.3	3.5	6.8	23.7	49.1	0.8	67.1	27.0	76.2
	a=1.5	m=2	4.9	6.2	2.4	3.7	20.0	50.7	1.0	55.2	25.8	80.2
n=100		m=1	4.4	5.9	4.6	6.3	22.7	50.2	0.8	63.2	22.5	75.1
	a=2	m=2	5.0	4.6	2.9	4.6	21.1	48.8	0.7	59.7	20.0	78.1
		m=1	4.4	6.9	5.7	3.7	26.3	53.5	0.4	61.4	15.7	71.0
	a=3	m=2	3.2	5.6	4.0	3.9	25.1	55.7	1.0	61.1	14.0	73.7
		m=1	5.4	5.3	5.2	3.7	26.6	51.0	0.5	66.0	13.9	65.9
	a=4	m=2	5.5	5.0	5.2	4.3	28.1	55.4	0.9	62.4	13.9	71.1
		m=1	4.6	5.3	2.8	42.5	53.5	95.8	2.3	99.9	96.6	99.1
	a=0.5	m=2	4.3	4.6	0.2	18.3	41.8	92.8	2.1	98.6	87.8	99.2
		m=1	5.9	5.2	3.7	32.1	55.2	97.2	1.7	99.6	90.1	99.7
	a=1	m=2	5.8	4.9	1.7	16.8	53.4	97.4	2.2	99.3	84.4	99.5
		m=1	6.0	5.2	3.5	20.2	60.2	97.4	0.5	99.4	78.7	99.3
	a=1.5	m=2	4.6	5.4	3.0	14.4	54.7	96.6	1.9	99.0	77.3	99.6

n=300		m=1	5.8	5.3	5.3	14.1	63.2	95.2	1.0	98.6	62.3	98.7
	a=2	m=2	5.4	3.9	2.8	13.1	58.7	96.5	1.3	97.8	61.0	99.2
		m=1	6.8	5.3	2.9	8.3	64.4	96.6	0.8	98.4	33.0	95.8
	a=3	m=2	4.7	5.3	3.6	7.2	63.6	95.8	0.7	98.2	31.9	98.6
		m=1	4.3	5.8	4.6	7.1	67.2	96.1	0.5	96.9	18.7	90.6
	a=4	m=2	4.2	4.0	4.9	5.8	66.6	97.0	1.1	97.9	22.9	95.2

Table 2 Percentage of rejection of martingale difference hypothesis ($H_0^{(2)}$) for change from process P1 to process P2 with change-point k_0 . 1000 simulations with 1000 wild bootstrap replicates, $\alpha = 0.05$, $m = 1$, $a = 1$, $n = 300$.

			P2									
P1	k_0	γ	IID	G1	SV	NLMA	BIL-I	BIL-II	NDAR	TAR(1)	EXP(1)	ARFIMA
		0	6.0	5.6	4.9	9.5	18.8	42.6	1.4	63.1	30.3	75.8
	150	0.5	4.4	6.1	3.5	7.4	22.0	51.5	1.2	71.7	34.6	83.5
IID		0	5.2	4.6	4.2	4.5	9.1	14.3	3.6	20.4	8.0	40.9
	225	0.5	5.3	6.2	4.0	3.4	9.2	19.6	2.7	31.4	13.6	54.4
		0	5.1	5.1	4.1	9.3	25.8	62.7	1.4	80.9	48.3	89.1
	150	0.5	5.1	5.5	2.7	6.7	25.4	60.7	0.9	79.9	46.0	89.2
G1		0	5.6	4.0	3.9	4.2	14.9	30.6	0.9	47.4	19.4	63.3
	225	0.5	4.7	5.1	2.2	3.4	14.1	35.0	0.9	49.7	24.6	69.7
		0	5.1	3.9	3.5	8.6	14.3	32.4	2.1	48.5	18.7	66.7
	150	0.5	5.3	3.2	3.7	7.4	15.7	36.2	2.4	53.0	21.0	69.3
SV		0	4.7	4.1	4.3	3.5	4.9	10.7	3.4	13.6	6.6	28.2
	225	0.5	3.6	4.1	4.2	4.0	6.0	10.5	2.8	14.9	7.4	31.8

Table 3 Percentage of rejection of martingale difference hypothesis ($H_0^{(2)}$) for change from process P1 to process P2 with change-point k_0. 1000 simulations with 1000 wild bootstrap replicates, $\alpha = 0.05$, $m = 1$, $a = 1$, $n = 600$.

			P2									
P1	k_0	γ	IID	G1	SV	NLMA	BIL-I	BIL-II	NDAR	TAR(1)	EXP(1)	ARFIMA
		0	5.3	4.1	3.2	22.1	35.3	81.7	2.1	96.3	60.5	96.2
	300	0.5	5.4	6.1	3.2	21.1	40.3	86.7	2.3	97.5	66.4	97.1
IID		0	4.8	4.1	4.8	8.4	11.2	26.7	3.1	38.7	17.2	56.3
	450	0.5	4.4	5.2	2.4	7.5	15.0	38.1	2.1	56.4	19.4	72.2
		0	5.2	5.0	3.4	23.9	52.9	93.7	1.6	99.4	82.8	99.1
	300	0.5	4.9	5.0	3.6	19.0	50.4	93.4	0.8	99.2	82.2	99.5
G1		0	5.2	5.2	4.4	11.6	25.5	61.2	1.7	77.2	42.7	84.3
	450	0.5	4.0	4.2	3.6	7.4	27.2	59.0	0.5	81.2	43.4	89.3
		0	4.4	3.1	3.3	16.1	24.4	59.7	1.9	76.1	35.5	86.2
	300	0.5	4.7	4.2	4.1	18.8	26.7	66.4	2.9	81.2	41.5	89.3
SV		0	3.4	3.9	5.5	5.4	7.5	14.1	3.5	20.6	9.5	43.9
	450	0.5	3.9	4.3	3.8	6.6	6.5	19.0	2.4	28.2	10.6	46.4

Table 4 Percentage of rejection of martingale difference hypothesis ($H_0^{(3)}$) for change from process P1 to process P2 with change-point k_0. 1000 simulations with 1000 wild bootstrap replicates, $\alpha = 0.05$, $m = 1$, $a = 1$, $n = 300$.

			P2									
P1	k_0	γ	IID	G1	SV	NLMA	BIL-I	BIL-II	NDAR	TAR(1)	EXP(1)	ARFIMA
		0	6.7	6.4	3.9	5.8	13.2	34.2	1.3	50.4	20.3	72.3
	150	0.5	5.2	6.0	2.4	2.8	10.4	27.5	1.7	45.4	16.9	73.6
IID		0	5.9	6.6	3.6	4.0	11.0	18.2	3.6	27.4	10.9	46.0
	225	0.5	4.9	3.9	3.5	2.6	8.4	17.2	2.4	30.5	14.8	53.9
		0	5.0	6.0	2.9	4.0	15.7	42.6	0.6	60.0	25.8	84.3
	150	0.5	5.3	5.3	2.4	2.8	15.8	38.9	0.4	56.6	22.8	82.1
G1		0	5.2	4.5	2.9	2.6	11.7	32.5	0.7	44.9	21.9	69.9
	225	0.5	4.4	4.6	3.1	1.4	12.0	26.9	0.9	40.9	19.6	70.5
		0	4.1	3.4	3.3	4.8	8.0	20.2	2.1	29.1	11.8	49.9
	150	0.5	3.0	3.5	3.4	3.3	8.7	16.9	1.2	24.5	10.0	48.7
SV		0	3.4	3.7	3.3	3.4	5.7	9.6	3.0	13.1	4.7	29.6
	225	0.5	3.4	2.8	2.7	3.0	5.3	8.2	2.0	14.1	5.3	29.9

Table 5 Percentage of rejection of martingale difference hypotheses $H_0^{(2)}$ and $H_0^{(3)}$ for change from process P1 to process P2 with change-point $k_0 = 50$. 1000 simulations with 1000 wild bootstrap replicates, $\alpha = 0.05$, $m = 1$, $a = 1$, $n = 300$.

		$H_0^{(2)}$					$H_0^{(3)}$				
		μ_2					μ_2				
	μ_1	0.0	0.1	0.2	0.5	1.0	0.0	0.1	0.2	0.5	1.0
	0.0	4.1	32.0	82.7	100	100	3.5	6.6	12.0	52.2	99.0
	0.1	4.8	33.4	85.6	100	100	6.5	4.5	4.2	33.6	97.3
P1	0.2	5.5	40.1	87.9	100	100	13.9	5.6	4.4	19.9	88.7
	0.5	18.4	67.9	97.0	100	100	63.8	39.8	21.3	2.5	29.1
	1.0	65.3	95.2	100	100	100	99.9	99.1	96.8	46.5	1.1

Figure 1 Daily scaled log returns of S&P 500 from January 1993 until December 1997 (source: Yahoo! Finance, http://finance.yahoo.com.) Dashed line denotes January 1st, 1994, solid line denotes December 8th, 1994.