SEQUENTIAL TESTING FOR THE STABILITY OF HIGH FREQUENCY PORTFOLIO BETAS*

Alexander Aue[1] **Siegfried Hörmann**[2] **Lajos Horváth**[3]
Marie Hušková[4] **Josef G. Steinebach**[5]

[1] Department of Statistics, University of California, Davis,
One Shields Avenue, Davis, CA 95616, USA

[2] Département de Mathémathique, Université Libre de Bruxelles,
Bd du Triomphe, B–1050 Brussels, Belgium

[3] Department of Mathematics, University of Utah,
155 South 1440 East, Salt Lake City, UT 84112, USA

[4] Charles University of Prague, Department of Statistics,
Sokolovská 83, CZ–186 00 Prague, Czech Republic

[5] Mathematical Institute, University of Cologne,
Weyertal 86–90, D–50 931 Köln, Germany

Abstract

Despite substantial criticism, variants of the capital asset pricing model (CAPM) remain until today the primary statistical tools for portfolio managers to assess the performance of financial assets. In the CAPM, the risk of an asset is expressed through its correlation with the market, widely known as the beta. There is now a general consensus among economists that these portfolio betas are time-varying and that, consequently, any appropriate analysis has to take this variability into account. Recent advances in data acquisition and processing techniques have led to an increased research output concerning high-frequency models. Within this framework, we introduce here a modified functional CAPM and sequential monitoring procedures to test for the constancy of the portfolio betas. As our main results we derive the large-sample properties of these monitoring procedures. In a simulation study and an application to S&P 100 data we show that our method performs well in finite samples.

1 Introduction

The traditional Capital Asset Pricing Model, henceforth abbreviated CAPM, of Sharpe (1964), Lintner (1965) and Merton (1973), and its subsequently introduced modifications and extensions have

*We would like to thank Oliver Linton and three anonymous referees for their constructive comments. This research was partially supported by NSF grant DMS 0905400, grant GACR 201/09/J006, grant MSM 0021620839, Banque nationale de Belgique and Communauté française de Belgique—Actions de Recherche Concertées (2010–2015).

enjoyed both countless applications in all areas of empirical finance and strong criticism from theoreticians and practitioners alike. One of the primary advantages of the CAPM is its simplicity. To describe an asset's sensitivity to market risk, only one parameter is needed in addition to the expected market return: the so-called beta measuring the correlation of the asset with the market. The formulation of the CAPM is also rooted in financial economics theory (see, for example, Markowitz, 1999). Criticism has focused on the role of the betas, in particular their potentially time-varying character (see Ghysels, 1998, and Harvey, 1991, among many others), and has led to conditional versions of the CAPM (see Andersen et al., 2006, for a comprehensive review). Ghysels (1998) has shown, however, that in many cases misspecified conditional CAPMs tend to produce larger pricing errors than a (piecewise) unconditional CAPM and has consequently advocated the use of the latter, taking into account what is known as structural breaks. Further contributions to the literature supporting this hypothesis will be addressed in Section 2.

In this paper we take the view of a portfolio manager who has to decide online whether to hold or sell the assets in his portfolio. To this end, we develop a new scheme to monitor the constancy of the portfolio betas over time. Our work draws both from Ghysels (1998) and Chu et al. (1996), yet goes beyond their contributions in that we account for high-frequency phenomena (see, for example, Barndorff-Nielsen and Shephard, 2004). While Ghysels (1998) provides the theoretical foundation for our approach, Chu et al. (1996) have introduced several monitoring procedures which we will adjust below for our purposes. These procedures, or stopping rules, can be briefly summarized as follows. Based on a training sample of size m, we define our procedures as first crossing times of a suitably constructed threshold function by a quadratic form detector built from least-squares estimates of the portfolio betas. Exceedance of the threshold indicates then the presence of a structural break, that is, time variability in the betas. A timely and accurate detection is paramount. We assess the quality of the procedures both through their theoretical large-sample properties, letting $m \to \infty$, and their empirical finite-sample performance, fixing m at moderate sizes.

The results presented in this paper are novel in several respects. First, we extend both the traditional CAPM model and the sequential monitoring procedures to a high-frequency setting. Second, since in many modern applications of finance and econometrics high-frequency observations can be taken on fine time grids, it appears natural to focus on a functional approach in which one observes a reasonably smooth function, or curve, on a compact interval $[0, S]$ instead of a potentially large collection of individual data points. The interval endpoint S is typically determined by practical considerations and can, for example, represent a trading day or some other appropriate unit of time. In view of the competing models available in the literature and the related model fitting and identification issues, we aim at providing flexible and nonrestrictive assumptions to ensure widespread applicability of the monitoring procedures. Third, to derive statements in the proposed framework, it is necessary to further advance the theory of functional, thus Hilbert space-valued, time series.

The remainder of this paper is organized as follows. In Section 2, we discuss the version of the CAPM that will be assumed throughout as well as the monitoring procedures for the portfolio betas. Section

3 contains the assumptions and the statement of the main results establishing the long-run behavior. Sections 4 and 5 are devoted to a number of examples and empirical results, respectively. Section 6 concludes and all proofs are given in a technical appendix.

2 Monitoring in a functional capital asset pricing model

We start with the definition of the multiple unconditional CAPM of Sharpe (1964), Lintner (1965) and Merton (1973). We consider in the following d risky assets and denote the vector of their daily log-returns by $\boldsymbol{r}_i \in \mathbb{R}^d$. Let moreover $r_{M,i} \in \mathbb{R}$ be the log-return of the observable market portfolio. The CAPM is then given by

$$\boldsymbol{r}_i = (\mathbf{1} - \boldsymbol{\beta})\gamma + \boldsymbol{\beta} r_{M,i} + \boldsymbol{\varepsilon}_i, \qquad i \in \mathbb{Z}, \tag{2.1}$$

where $\mathbb{Z}$ denotes the set of integers, $\mathbf{1} = (1, \ldots, 1)'$ is the d-dimensional vector whose components are all equal to one and $'$ signifies transposition, γ denotes the expected return on a zero-beta portfolio (the return on risk free assets) and $\boldsymbol{\varepsilon}_i \in \mathbb{R}^d$ is a random error. The d-dimensional vector $\boldsymbol{\beta}$ contains the portfolio betas which quantize the assets' market risk. Using model (2.1) for asset pricing purposes, in particular over longer time spans, assumes that the portfolio betas remain indeed unchanged during the observation period. Otherwise, assets will be mispriced and predictions based on (2.1) will be incorrect and misleading. These facts have been pointed out in a variety of works. The general consensus emerging from these papers is that the $\boldsymbol{\beta}$'s are time-varying. To account for this, one line of research advocated the use of conditional CAPM which explicitly model the dynamics of the portfolio betas; see Andersen et al. (2006) for a recent review. On the other hand, Ghysels (1998) compared several conditional CAPM with piecewise constant versions and found that the latter approach typically leads to superior performance measured in terms of pricing errors. The hypothesis of Ghysels (1998) has been supported by subsequent research. Garcia and Ghysels (1998) examine, for example, the effect of structural change on emerging markets. Huang and Chen (2005), and Wang et al. (2009) provided examples of sudden changes in the economic environment which in turn caused changes in the portfolio betas. Amsler and Schmidt (1985), MacKinley (1987), and Balduzzi and Robotti (2008) studied the performance of multiple CAPM via simulations and assess the power of tests to distinguish between CAPM and other asset pricing models. Santos (1998) assumed a regime switching CAPM with low, medium and high regimes of volatility and used his model to study the behavior of US financial markets during several decades. In keeping with the conclusion of these papers, we shall entertain here a model for which the portfolio betas are piecewise constant.

In today's financial applications it is common to deal with high-frequency, intra-day data. Many models have been put forward to incorporate the specific high-frequency stylized facts. We refer here to the contributions of Goodhart and O'Hara (1997), Engle (2000), and Barndorff-Nielsen and Shephard (2004), and to the references in these papers. In this paper, we assume $\boldsymbol{r}_i$ and $r_{M,i}$ to be random functions on $[0, S]$. Without loss of generality, we shall always let $S = 1$. Let $(\boldsymbol{P}_i(s) : s \in [0,1],\ i \in \mathbb{Z})$ be a sequence of price processes and define the log-returns $\boldsymbol{r}_i(s) = \log \boldsymbol{P}_i(s + H) - \log \boldsymbol{P}_i(s)$ for some fixed time parameter H. To incorporate both the possibility of unstable betas and the high frequency of the intra-day observations, we replace (2.1) now with the functional CAPM

$$\boldsymbol{r}_i(s) = (\mathbf{1} - \boldsymbol{\beta}_i)\gamma + \boldsymbol{\beta}_i r_{M,i}(s) + \boldsymbol{\varepsilon}_i(s), \qquad s \in [0,1],\ i \in \mathbb{Z}. \tag{2.2}$$

The quantities in (2.2) all have counterparts in the original CAPM (2.1), but are to be interpreted differently. The log-returns $\boldsymbol{r}_i(s)$ are now functions of two arguments, the first of which is the daily index i as before. The second argument, $s \in [0,1]$, is used to describe the intra-day behavior. Similar explanations apply to $r_{M,i}(s)$ and $\boldsymbol{\varepsilon}_i(s)$. Dependence in the functional CAPM can thus have two sources: serial dependence across trading days $i \in \mathbb{Z}$ and intra-day dependence across $s \in [0,1]$. Both structures are explicitly enabled. We will give additional model assumptions in Section 3 below. Other approaches to high-frequency data which model stock prices and instantaneous returns ($H \to 0$) as solutions to stochastic differential equations are discussed in Harrison et al. (1984), Zhang et al. (2005), Barndorff-Nielsen et al. (2008a, 2008b), and Jacod et al. (2009). These authors focus on issues related to quadratic variation and typically use a single stochastic process to model prices and returns over the whole observation period, while we describe the data through a sequence of stochastic processes, where one such process is used for one trading day, or any other appropriate time span.

As indicated in the Introduction, we take the sequential point of view of a portfolio manager. Because the portfolio betas are assumed to be piecewise constant, we require first that there is a historical training period of length m for which $\boldsymbol{\beta} = \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m$. Based on these current values of the portfolio betas, we are interested in finding online time points that require an update of the market risks collected in the vector $\boldsymbol{\beta}$. This will be achieved by casting the problem into a sequential hypothesis testing procedure, with the null hypothesis given by

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

thus describing the constancy of the portfolio betas over time. In view of the results in Chu et al. (1996), we consider the alternative

$$H_A : \ \text{There is } k^* \geq 1 \ \text{ such that } \ \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m = \ldots = \boldsymbol{\beta}_{m+k^*-1} \neq \boldsymbol{\beta}_{m+k^*} = \boldsymbol{\beta}_{m+k^*+1} = \ldots,$$

where the unknown time lag k^* is commonly referred to as a break point. The alternative describes thus a situation in which the market risks are given by a vector $\boldsymbol{\beta}$ before k^* and by a different vector, say, $\boldsymbol{\beta}^*$ thereafer.

In the following, we will suggest sequential monitoring procedures, or detection rules, to distinguish between H_0 and H_A. To do so, we first compute the least squares estimator $\hat{\boldsymbol{\beta}}_m$ for the common portfolio betas from the historical sample. After this, as observations arrive sequentially, we can compute the least squares estimators $\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \ldots$. Each estimator $\tilde{\boldsymbol{\beta}}_k$ is based on

$\boldsymbol{r}_{m+1}, r_{M,m+1}, \ldots, \boldsymbol{r}_{m+k}, r_{M,m+k}$ (viewed as functions on $[0,1]$). Finally, we compare $\hat{\boldsymbol{\beta}}_m$ to the $\tilde{\boldsymbol{\beta}}_k$ and reject H_0 in favor of H_A at time lag k if the estimators are significantly different. To make this precise, we assume as in Barndorff-Nielsen and Shephard (2004) that, during day i, we have observed the random functions $\boldsymbol{r}_i$ and $r_{M,i}$ at J equally spaced intra-day time points $s_j = j/J$, $1 \le j \le J$. The least squares estimators $\hat{\boldsymbol{\beta}}_m$ and $\tilde{\boldsymbol{\beta}}_k$ are then given by

$$\hat{\boldsymbol{\beta}}_m = \left(\sum_{i=1}^{m} \sum_{j=1}^{J} \left[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j) \right]^2 \right)^{-1} \sum_{i=1}^{m} \sum_{j=1}^{J} \left[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j) \right] \left[\boldsymbol{r}_i(s_j) - \hat{\boldsymbol{r}}_m(s_j) \right],$$

$$\hat{r}_{M,m}(s) = \frac{1}{m} \sum_{i=1}^{m} r_{M,i}(s), \quad \hat{\boldsymbol{r}}_m(s) = \frac{1}{m} \sum_{i=1}^{m} \boldsymbol{r}_i(s),$$

and

$$\tilde{\boldsymbol{\beta}}_k = \left(\sum_{i=m+1}^{m+k} \sum_{j=1}^{J} \left[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j) \right]^2 \right)^{-1} \sum_{i=m+1}^{m+k} \sum_{j=1}^{J} \left[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j) \right] \left[\boldsymbol{r}_i(s_j) - \tilde{\boldsymbol{r}}_k(s_j) \right],$$

$$\tilde{r}_{M,k}(s) = \frac{1}{k} \sum_{i=m+1}^{m+k} r_{M,i}(s), \quad \tilde{\boldsymbol{r}}_k(s) = \frac{1}{k} \sum_{i=m+1}^{m+k} \boldsymbol{r}_i(s).$$

Under H_0, the difference $\hat{\boldsymbol{\beta}}_m - \tilde{\boldsymbol{\beta}}_k$, suitably normalized, is asymptotically normal with mean zero and some covariance matrix $\boldsymbol{D}$ if both $m \to \infty$ and $k \to \infty$. Since $\boldsymbol{D}$ is unknown in practice, it will be replaced with an estimator $\hat{\boldsymbol{D}}_m$ based on the historical sample. This issue is discussed in more detail in Section 4 below. The sequential procedure we propose is then based on the detector

$$V_k = \frac{1}{m} \tilde{U}_k^2 (\tilde{\boldsymbol{\beta}}_k - \hat{\boldsymbol{\beta}}_m)' \hat{\boldsymbol{D}}_m^{-1} (\tilde{\boldsymbol{\beta}}_k - \hat{\boldsymbol{\beta}}_m), \qquad k \in \mathbb{N},$$

where

$$\tilde{U}_k = \frac{1}{J} \sum_{i=m+1}^{m+k} \sum_{j=1}^{J} \left[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j) \right]^2. \tag{2.3}$$

Let $T > 0$. We stop and declare H_0 to be invalid at the first time k such that the detector V_k exceeds the value of a scaled threshold function w, therefore yielding the stopping rule

$$\tau_m = \min \left\{ k \le \lfloor mT \rfloor : V_k > cw\left(\frac{k}{m}\right) \right\},$$

with c being a suitably chosen constant (see below) and $\lfloor \cdot \rfloor$ denoting integer part. The procedure is called a closed-end method, since monitoring does not continue after $\lfloor mT \rfloor$ observations have been taken. We shall write $\tau_m < \infty$ to indicate that the monitoring has been terminated during the testing period, i.e., the detector V_k has crossed the boundary $cw(k/m)$ for some $k \le \lfloor mT \rfloor$. We shall write $\tau_m = \infty$ if the set $\{k \le \lfloor mT \rfloor : V_k > cw(k/m)\}$ is empty, i.e., the detector has not crossed the boundary during the testing period. The threshold function w is chosen such that a change can be detected as soon as possible. The only formal restrictions to be imposed on w are positivity and continuity. Depending on the application a more "sensitive" w can be chosen which is able to detect changes quickly, but inherits a higher false discovery rate. Vice versa, a less sensitive w can be chosen which will need longer to detect a change, but is therefore more stable and reliable. (We will elaborate more on these issues in two of the examples of Section 5.) The constant $c = c(\alpha)$ is determined by the equation

$$\lim_{m\to\infty} P\left(\tau_m < \infty\right) = \alpha \quad \text{under} \quad H_0, \tag{2.4}$$

where $\alpha \in (0,1)$ denotes the asymptotic significance level. Stopping rules of this type were originally discussed in Chu et al. (1996) in the context of linear regressions. Modifications have since been applied successfully in various settings, including autoregressive processes (Hušková et al., 2007), GARCH processes (Berkes et al., 2004), and multiple time series regressions (Aue et al., 2009).

We conclude this section with a couple of remarks concerning computational properties of the sequential procedure. It is worthwhile mentioning that $\hat{\boldsymbol{\beta}}_m$ and $\tilde{\boldsymbol{\beta}}_k$ are weighted sums of the residuals $(\boldsymbol{r}_i(s_j) - \hat{\boldsymbol{r}}_m(s_j)\colon 1 \le j \le J, 1 \le i \le m)$ and $(\boldsymbol{r}_i(s_j) - \tilde{\boldsymbol{r}}_k(s_j)\colon 1 \le j \le J, m+1 \le i \le m+k)$, respectively. Using the formulae

$$\sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\left[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j)\right]\left[\boldsymbol{r}_i(s_j) - \tilde{\boldsymbol{r}}_k(s_j)\right] = \sum_{i=m+1}^{m+k}\sum_{j=1}^{J} r_{M,i}(s_j)\boldsymbol{r}_i(s_j) - k\sum_{j=1}^{J}\tilde{r}_{M,k}(s_j)\tilde{\boldsymbol{r}}_k(s_j)$$

and

$$\sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\left[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j)\right]^2 = \sum_{i=m+1}^{m+k}\sum_{j=1}^{J} r^2_{M,i}(s_j) - k\sum_{j=1}^{J}\tilde{r}^2_{M,k}(s_j),$$

one can quickly compute $\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \ldots$ via sequential updating. Similarly, the detector V_k is a function of the weighted residuals. Here

$$V_k = \frac{1}{m}\boldsymbol{R}_k' \hat{\boldsymbol{D}}_m^{-1}\boldsymbol{R}_k,$$

where

$$\begin{aligned}
\boldsymbol{R}_k &= \sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\left[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j)\right]\left[\boldsymbol{r}_i(s_j) - \tilde{\boldsymbol{r}}_k(s_j)\right] \\
&\qquad - \frac{\tilde{U}_k}{\hat{U}_m}\sum_{i=1}^{m}\sum_{j=1}^{J}\left[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j)\right]\left[\boldsymbol{r}_i(s_j) - \hat{\boldsymbol{r}}_m(s_j)\right] \\
&= \sum_{i=m+1}^{m+k}\sum_{j=1}^{J} r_{M,i}(s_j)\boldsymbol{r}_i(s_j) - k\sum_{j=1}^{J}\tilde{r}_{M,k}(s_j)\tilde{\boldsymbol{r}}_k(s_j) \\
&\qquad - \frac{1}{\hat{U}_m}\frac{1}{J}\left(\sum_{i=m+1}^{m+k} r^2_{M,i}(s_j) - k\sum_{j=1}^{J}\tilde{r}^2_{M,k}(s_j)\right)\sum_{i=1}^{m}\sum_{j=1}^{J}\left[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j)\right]\boldsymbol{r}_i(s_j),
\end{aligned}$$

with

$$\hat{U}_m = \frac{1}{J}\sum_{i=1}^{m}\sum_{j=1}^{J}\left[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j)\right]^2. \tag{2.5}$$

These computational properties appear useful when dealing with large data sets now common in financial economics. In Section 5, we shall discuss the applicability of the sequential procedures introduced in this section. There, we report the results of an analysis of high-frequency tick-data from stocks in the S&P 100 index and a small simulation study. Before turning to the empirical evidence, we state the asymptotic behavior of τ_m in Section 3 and review several examples in Section 4.

3 Model assumptions and main results

Throughout this paper we assume that $(r_{M,i}(s)\colon s \in [0,1], i \in \mathbb{Z})$ and $(\boldsymbol{r}_i(s)\colon s \in [0,1], i \in \mathbb{Z})$ are function-valued time series. Both sequences are assumed to be generated by functional Bernoulli shifts. In traditional time series this approach has been used, for example, by Shao and Wu (2004), Wu (2005) and, in the context of multivariate financial time series, by Aue et al. (2009).

Assumption 3.1. *For any $i \in \mathbb{Z}$, we let $r_{M,i} = a(\boldsymbol{\delta}_i, \boldsymbol{\delta}_{i-1}, \ldots)$, where $a(\cdot)$ is a measurable real-valued functional and $(\boldsymbol{\delta}_i\colon i \in \mathbb{Z})$ is a sequence of independent, identically distributed (i.i.d.) random functions. For any $i \in \mathbb{Z}$, we let $E[\boldsymbol{\varepsilon}_i(s)] = \mathbf{0}$, $s \in [0,1]$, and $\boldsymbol{\varepsilon}_i = \boldsymbol{b}(\boldsymbol{\eta}_i, \boldsymbol{\eta}_{i-1}, \ldots)$, where $\boldsymbol{b}(\cdot)$ is a measurable $\mathbb{R}^d$-valued functional and $(\boldsymbol{\eta}_i\colon i \in \mathbb{Z})$ is a sequence of i.i.d. random functions. Moreover, $(\boldsymbol{\delta}_i\colon i \in \mathbb{Z})$ and $(\boldsymbol{\eta}_i\colon i \in \mathbb{Z})$ are assumed to be independent.*

The random functions $(\boldsymbol{\delta}_i\colon i \in \mathbb{Z})$ and $(\boldsymbol{\eta}_i\colon i \in \mathbb{Z})$ in Assumption 3.1 can be vector-valued with dimension not necessarily equal to d. The independence between these sequences can be dropped. It is used here mainly to simplify the presentation. Assumption 3.1 ensures that $(r_{M,i}(s)\colon s \in [0,1], i \in \mathbb{Z})$ and $(\boldsymbol{r}_i(s)\colon s \in [0,1], i \in \mathbb{Z})$ are stationary and ergodic sequences of random function innovations, possessing a causal representation. In addition we also need that both processes are smooth functions of the intra-day argument $s \in [0,1]$. To this end, let

$$\omega(x; h) = \sup_{0 \le t \le 1-h} \sup_{0 \le s \le h} |x(t+s) - x(t)|$$

denote the modulus of continuity of a function $x(\cdot)$ on $[0,1]$. For $p \ge 1$, let $\|\xi\|_p = (E[|\xi|^p])^{1/p}$ be the L_p-norm of a random variable ξ.

Assumption 3.2. *For any $i \in \mathbb{Z}$, let*

$$\Big\| \sup_{0 \le s \le 1} |r_{M,i}(s)| \Big\|_4 < \infty \qquad \text{and} \qquad \max_{1 \le \ell \le d} \Big\| \sup_{0 \le s \le 1} |\varepsilon_{i,\ell}(s)| \Big\|_4 < \infty$$

as well as

$$\lim_{h \to 0} \big\|\omega(r_{M,i}; h)\big\|_4 = 0 \qquad \text{and} \qquad \lim_{h \to 0} \max_{1 \le \ell \le d} \big\|\omega(\varepsilon_{i,\ell}; h)\big\|_4 = 0,$$

where, for $s \in [0,1]$, $\boldsymbol{\varepsilon}_i(s) = (\varepsilon_{i,1}(s), \ldots, \varepsilon_{i,d}(s))'$.

We allow weak dependence structures for the processes $r_{M,i}$ and $\boldsymbol{r}_i$ as functions of the daily index i. Roughly speaking we assume that both processes can be well approximated with L-dependent sequences with the rate of the approximation getting better as $L \to \infty$. We give the precise formulation in the following assumption.

Assumption 3.3. *For any $i \in \mathbb{Z}$, let*

$$\sum_{L=1}^{\infty} \sup_{0\le s\le 1} \big\| r_{M,i}(s) - r_{M,i}^{(L)}(s) \big\|_4 < \infty,$$

where $r_{M,i}^{(L)} = a(\boldsymbol{\delta}_i, \boldsymbol{\delta}_{i-1}, \ldots, \boldsymbol{\delta}_{i-L+1}, \boldsymbol{\delta}_{i-L}^{(i)}, \boldsymbol{\delta}_{i-L-1}^{(i)}, \ldots)$, is an approximating process built from a sequence $(\boldsymbol{\delta}_L^{(i)} \colon i, L \in \mathbb{Z})$ of i.i.d. copies of $\boldsymbol{\delta}_0$, independent of $(\boldsymbol{\delta}_i \colon i \in \mathbb{Z})$. Similarly, we assume that, for all $1 \le \ell \le d$,

$$\sum_{L=1}^{\infty} \sup_{0\le s\le 1} \big\| \varepsilon_{i,\ell}(s) - \varepsilon_{i,\ell}^{(L)}(s) \big\|_4 < \infty,$$

where $\boldsymbol{\varepsilon}_i^{(L)}$ denote the coordinates of the process $\boldsymbol{\varepsilon}_i^{(L)} = \boldsymbol{b}(\boldsymbol{\eta}_i, \boldsymbol{\eta}_{i-1}, \ldots, \boldsymbol{\eta}_{i-L+1}, \boldsymbol{\eta}_{i-L}^{(i)}, \boldsymbol{\eta}_{i-L-1}^{(i)}, \ldots)$, with $(\boldsymbol{\eta}_L^{(i)} \colon i, L \in \mathbb{Z})$ being i.i.d. copies of $\boldsymbol{\eta}_0$, independent of $(\boldsymbol{\eta}_i \colon i \in \mathbb{Z})$.

The approximations for $r_{M,i}$ and $\varepsilon_{i,\ell}$ in Assumption 3.3 are obtained through a coupling construction. Closeness between the original processes and their approximation counterparts $r_{M,i}^{(L)}$ and $\varepsilon_{i,\ell}^{(L)}$ is evaluated in terms of distance measures between the L_p-norm differences. Similar types of approximation schemes were first used in Ibragimov (1962) for univariate random variables, and lately in more general contexts also in Wu (2007) and Aue et al. (2009). Note that by stationarity the expected values of the processes appearing in Assumptions 3.2 and 3.3 do not depend on i. By Assumption 3.1 we can assume that the processes $(\boldsymbol{\delta}_L^{(i)} \colon i, L \in \mathbb{Z})$ and $(\boldsymbol{\eta}_L^{(i)} \colon i, L \in \mathbb{Z})$ are independent. Consequently, $(r_{M,i}^{(L)} \colon i, L \in \mathbb{Z})$ and $(\boldsymbol{\varepsilon}_i^{(L)} \colon i, L \in \mathbb{Z})$ are independent as well.

Assumption 3.4. *We let $J = J(m) \to \infty$ as $m \to \infty$.*

Since we work with tick-level data we assume that, for large m, we are able to obtain the intra-day behavior in high resolution. Under Assumption 3.4, the least squares estimators $\hat{\boldsymbol{\beta}}_m$ and $\tilde{\boldsymbol{\beta}}_k$ can be approximated with functionals of the stochastic integrals

$$\boldsymbol{z}_i = \int_0^1 \rho_i(s) \boldsymbol{\varepsilon}_i(s) ds, \qquad i \in \mathbb{Z}, \tag{3.1}$$

where $\rho_i(s) = r_{M,i}(s) - E[r_{M,i}(s)]$, $s \in [0,1]$. Assumption 3.2 implies that the functions ρ_i and $r_{M,i}$ have continuous sample paths with probability one, and $\boldsymbol{z}_i$ in (3.1) can be defined as a pathwise Riemann integral. We also define

$$\boldsymbol{D} = E\big[\boldsymbol{z}_0 \boldsymbol{z}_0'\big] + \sum_{i=1}^{\infty} E\big[\boldsymbol{z}_0 \boldsymbol{z}_i' + \boldsymbol{z}_i \boldsymbol{z}_0'\big]. \tag{3.2}$$

It will be shown below that, under our set of assumptions, $\boldsymbol{D}$ exists and that the sum in the definition of $\boldsymbol{D}$ is absolutely convergent. The covariance matrix $\boldsymbol{D}$ is unknown and it needs to be estimated from the sample. We do this using only the historical training period, since all processes are stationary prior to time lag m regardless if H_0 or H_A holds. We assume that $\boldsymbol{D}$ can be estimated by a weakly consistent estimator $\hat{\boldsymbol{D}}_m$. Let $\stackrel{P}{\rightarrow}$ signify convergence in probability.

Assumption 3.5. *Let $\boldsymbol{D}$ be non-singular and let $\hat{\boldsymbol{D}}_m$ be such that $\hat{\boldsymbol{D}}_m \stackrel{P}{\rightarrow} \boldsymbol{D}$ as $m \to \infty$.*

Since the integrals $\boldsymbol{z}_i$ are unknown, HAC estimators for $\boldsymbol{D}$ are constructed from the residual vectors

$$\hat{\boldsymbol{z}}_i = \frac{1}{J}\sum_{j=1}^{J}\big[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j)\big]\hat{\boldsymbol{\varepsilon}}_i(s_j), \qquad 1 \le i \le m, \tag{3.3}$$

where

$$\hat{\boldsymbol{\varepsilon}}_i(s) = \boldsymbol{r}_i(s) - \hat{\boldsymbol{r}}_{i,m}(s) - \hat{\boldsymbol{\beta}}_m\big[r_{M,i}(s) - \hat{r}_{M,m}(s)\big], \qquad 1 \le i \le m.$$

Standard estimation methods employ lag kernel smoothing using weighted autocovariance estimators. Properties of such estimators were analyzed in Andrews (1991), who also gave recommendations on how to choose the bandwidth parameter. Extensions and modifications of these methods may be found in Newey and West (1987), Andrews and Monahan (1992), Jansson (2002), Phillips (2005), Smith (2005), and Liu and Wu (2010). In all cases, consistent estimators can be obtained under additional regularity conditions on the specific kernel and bandwidth, provided that both m and J tend to infinity. We discuss one such estimator for $\boldsymbol{D}$ which satisfies Assumption 3.5 in Section 5 below. Details on implementation and finite sample properties may be found in Zeileis (2004). Estimators for the integrated volatility based on smoothing, subsampling and averaging are studied in Zhang et al. (2005), Barndorff-Nielsen et al. (2008a, 2008b), and Jacod et al. (2009). Our final assumption details the requirements on the threshold function.

Assumption 3.6. *Let $w(\cdot)$ be continuous on $[0,T]$ and let $w(t) > 0$ for all $t \in [0,1]$.*

We are now ready to state the main result of our paper.

Theorem 3.1. *If Assumptions 3.1–3.6 are satisfied, then under H_0 and for any $T > 0$,*

$$\lim_{m\to\infty} P\big(\tau_m < \infty\big) = P\Big(\sup_{0\le t\le T}\frac{\Gamma_d(t)}{w(t)} > c\Big),$$

where, for $t \in [0,T]$, $\Gamma_d(t) = \sum_{1\le\ell\le d} B_\ell^2(t)$ and $B_1,\ldots,B_d$ denoting i.i.d. Gaussian processes with $E[B_\ell(t)] = 0$ and $E[B_\ell(t)B_\ell(u)] = \min(t,u) + tu$.

Several remarks are in order. Theorem 3.1 establishes the large sample behavior of the monitoring procedure based on the stopping rule τ_m. The limit distribution can be utilized for finite sample statistical inference. To this end, note that it can be shown that, for any $1 \le \ell \le d$,

$$\big(B_\ell(t)\colon t \in [0,T]\big) \stackrel{\mathcal{D}}{=} \big(W(t) + t\xi\colon t \in [0,T]\big),$$

where $(W(t)\colon t \geq 0)$ is a standard Brownian motion (Wiener process) independent of the the standard normal random variable ξ. Observe also that a comparison of the respective covariance functions leads to

$$\big(B_\ell(t)\colon t \geq 0\big) \stackrel{\mathcal{D}}{=} \Big((1+t)W\Big(\frac{t}{1+t}\Big)\colon t \geq 0\Big).$$

Hence

$$\sup_{0\leq t\leq T}\frac{\Gamma_d(t)}{w(t)} \stackrel{\mathcal{D}}{=} \sup_{0\leq t\leq T}\sum_{\ell=1}^{d}\frac{(1+t)^2}{w(t)}W_\ell^2\Big(\frac{t}{1+t}\Big) \stackrel{\mathcal{D}}{=} \sup_{0\leq u\leq T/(1+T)}\sum_{\ell=1}^{d}\frac{W_\ell^2(u)}{(1-u)^2 w\big(u/(1-u)\big)},$$

where $(W_\ell(t)\colon t \geq 0)$, $1 \leq \ell \leq d$, are independent standard Brownian motions. The foregoing leads to a natural selection for the threshold function: Choosing $w(t) = (1+t)^2$, the previous display implies that

$$\sup_{0\leq t\leq T}\frac{\Gamma_d(t)}{(1+t)^2} \stackrel{\mathcal{D}}{=} \sup_{0\leq u\leq T/(1+T)}\sum_{\ell=1}^{d}W_\ell^2(u) \stackrel{\mathcal{D}}{=} \frac{T}{1+T}\sup_{0\leq s\leq 1}\sum_{\ell=1}^{d}W_\ell^2(s),$$

so that quantiles of interest can be computed with the help of tables. The limiting distribution function in Theorem 3.1 is generally parameter-free, so that Monte Carlo simulations can be utilized in a straightforward way to obtain, for any $\alpha \in (0,1)$, the critical constants $c = c(\alpha)$ such that

$$P\Big(\sup_{0\leq t\leq T}\frac{\Gamma_d(t)}{w(t)} > c\Big) = \alpha,$$

thus ensuring that (2.4) holds, that is, the probability of a false alarm is approximately α if m is sufficiently large. The choice of the threshold w clearly affects the reaction time to a potential break. If there is knowledge on the suspected break time, one can choose w accordingly. For example, if a break is to occur soon after the monitoring commences, then the threshold should take its smallest value at or immediately following $m+1$. The situation is reverse if a break is suspected towards the end of monitoring. The precise derivation of the limit distribution of τ_m under the alternative is beyond the scope of this paper. This non-trivial task has been carried out, however, for simpler models in Aue and Horváth (2004), Aue et al. (2006a, 2006b, 2009) and Černiková et al. (2011).

Next, we study the behavior of the stopping rule under a one-change alternative. To this end, we allow $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$, the pre- and post-break regression parameters, to depend on m. Let $\Delta_m = \|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|$ measure the break size.

Theorem 3.2. *If Assumptions 3.1–3.6 are satisfied, then under H_A with $m < k^* = \lfloor mT^* \rfloor$ for some $0 < T^* < T$,*

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

provided that $m\Delta_m^2 \to \infty$ as $m \to \infty$.

Under the conditions of Theorem 3.2, a change in the regression parameters will be detected with high probability if the training sample is large. We note in passing that the result of Theorem 3.2 remains true if the regression parameters were to change smoothly in the following way. Assume that, starting at time lag $m + k^*$, the regression parameters start gradually shifting from the initial $\boldsymbol{\beta}$ to a new level $\boldsymbol{\beta}^{**}$ which is reached at time lag $m + k^{**}$ and assume that thereafter no further changes occur. If now $k^{**} = \lfloor mT^{**} \rfloor$ with some $T^* < T^{**} < T$ and $m\|\boldsymbol{\beta} - \boldsymbol{\beta}^{**}\|^2 \to \infty$ as $m \to \infty$, then, Theorem 3.2 continues to hold.

The model (2.2) can be extended to a functional factor model. For a presentation of popular factor models in time series we refer to Bauwens et al. (2006), Engle (2002), Engle and Kroner (1995), Engle et al. (1990), Hafner and Preminger (2009), and Hafner and Preminger (2009+). For the extension, we replace (2.2) with

$$\boldsymbol{r}_i(s) = (\mathbf{1} - \boldsymbol{\beta}_i)\boldsymbol{\gamma} + \boldsymbol{\beta}_i \boldsymbol{r}_{M,i}(s) + \boldsymbol{\varepsilon}_i(s), \qquad s \in [0,1],\ i \in \mathbb{Z},$$

where $\mathbf{1}$ now denotes a $d \times d'$ matrix whose elements are all equal to 1, $\boldsymbol{\gamma}$ is a d'-dimensional vector, $\boldsymbol{\beta}_i$ are $d \times d'$ matrices and $\boldsymbol{r}_{M,i}(s) = \big(r_{M(1),i}(s), \ldots, r_{M(d'),i}(s)\big)'$. Roughly speaking, d returns are explained by d' indices, and it is assumed that d' is much smaller than d. Very few factors can be used to explain the returns on a large number of assets. Using our method, we can find a sequential test for the hypothesis $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \ldots$, that is, for the stability of the portfolio beta matrices. The test is based on the same principles we discuss in this paper but involves more difficult notation.

4 Examples

In this section we introduce several examples and detail conditions under which Assumptions 3.1–3.3 are satisfied. We provide examples only for $\boldsymbol{\varepsilon}_i$ but the corresponding expressions for $r_{M,i}$ follow along the same arguments.

Example 4.1. If $(\boldsymbol{\varepsilon}_i = (\varepsilon_{i,1}, \ldots, \varepsilon_{i,d})' \colon i \in \mathbb{Z})$ are i.i.d. random processes satisfying $E[\boldsymbol{\varepsilon}_i(s)] = \mathbf{0}$,

$$\Big\| \sup_{0 \le s \le 1} |\varepsilon_{k,i}(s)| \Big\|_4 < \infty \qquad \text{and} \qquad \lim_{h \to 0} \big\|\omega(\varepsilon_{i,\ell}; h)\big\|_4 = 0,$$

then the conditions for $\boldsymbol{\varepsilon}_i$ are satisfied. If, for example, $(\boldsymbol{\varepsilon}_i \colon i \in \mathbb{Z})$ is a Gaussian process with almost surely continuous sample paths, then these assumptions hold.

To obtain processes exhibiting a more complex functional dependence structure, we will use in all of the following examples the i.i.d. processes of Example 4.1 as input to which we apply functional filters to obtain the output processes.

Example 4.2. Let $(\boldsymbol{\eta}_i = (\eta_{i,1}, \ldots, \eta_{i,d})' \colon i \in \mathbb{Z})$ be i.i.d. random processes with $E[\boldsymbol{\eta}_i(s)] = \mathbf{0}$ and $\|\sup_s |\eta_{i,\ell}(s)\|_4 < \infty$. Here and subsequently, we shall use the short notation $\sup_s$ to indicate that the supremum is taken over all $s \in [0,1]$. We also require that $\|\omega(\eta_{i,\ell}; h)\|_4 \to 0$ as $h \to 0$. Utilizing the functions $\boldsymbol{K}_n(s,t) = (K_{n,1}(s,t), \ldots, K_{n,d}(s,t))'$, $1 \le n \le q$, we define the functional MA(q) process $(\boldsymbol{\varepsilon}_i \colon i \in \mathbb{Z})$ via its coordinates by

$$\varepsilon_{i,\ell}(s) = \sum_{n=1}^{q} \int_0^1 K_{n,\ell}(s,t)\eta_{i-n,\ell}(t)dt + \eta_{i,\ell}(s). \tag{4.1}$$

To prove that $\| \sup_s |\varepsilon_{i,\ell}(s)| \|_4 < \infty$, we further impose the regularity condition

$$\max_{1\le\ell\le d} \max_{1\le n\le q} \sup_{0\le s\le 1} \int_0^1 |K_{n,\ell}(s,t)| dt < \infty. \tag{4.2}$$

The latter implies that each of the d coordinates of the q integrals in the definition of $\boldsymbol{\varepsilon}_i$ in (4.1) can be estimated as follows. We have

$$\sup_{0\le s\le 1} \Big| \int_0^1 K_{n,\ell}(s,t)\eta_{i-n,\ell}(t)dt \Big| \le \sup_{0\le t\le 1} |\eta_{i-n,\ell}(t)| \sup_{0\le s\le 1} \int_0^1 |K_{n,\ell}(s,t)|dt,$$

so that an application of the triangle inequality in combination with $\| \sup_s |\eta_{i,\ell}(s)| \|_4 < \infty$ and (4.2) leads to $\| \sup_s |\varepsilon_{i,\ell}(s)| \|_4 < \infty$. To ensure that the modulus of continuity $\omega(\varepsilon_{i,\ell}; h)$ satisfies the requirements of Assumption 3.2, we assume the smoothness condition

$$\max_{1\le\ell\le d} \max_{1\le n\le q} \sup_{0\le s\le 1-h} \sup_{0\le u\le h} \int_0^1 \big| K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t) \big| dt \to 0 \qquad (h\to 0). \tag{4.3}$$

We now have, for all $1\le \ell\le d$ and $1\le n\le q$, that

$$\begin{aligned} &\sup_{0\le s\le 1-h} \sup_{0\le u\le h} \Big| \int_0^1 \big[K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t)\big] \eta_{i-n,\ell}(t)dt \Big| \\ &\qquad \le \sup_{0\le t\le 1} |\eta_{i-n,\ell}(t)| \sup_{0\le s\le 1-h} \sup_{0\le u\le h} \int_0^1 \big| K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t) \big| dt, \end{aligned}$$

so that $\|\omega(\varepsilon_{i,\ell}; h)\|_4 \to 0$ as $h\to 0$ follows from assumption on $\eta_{i,\ell}$ and (4.3) via the triangle inequality. The processes $(\boldsymbol{\varepsilon}_i \colon i\in\mathbb{Z})$ are q-dependent by definition in (4.1). Assumption 3.3 holds consequently automatically with $L=q$.

Next, we extend Example 4.2 to the case $q=\infty$, that is, to linear processes in Hilbert spaces. A comprehensive presentation of functional linear processes can be found in Bosq (2000).

Example 4.3. Let the random functions $(\boldsymbol{\eta}_i \colon i\in\mathbb{Z})$ satisfy the assumptions of Example 4.2. Utilizing the functions $\boldsymbol{K}_n(s,t) = (K_{n,1}(s,t),\ldots,K_{n,d}(s,t))'$, $n\ge 1$, we define the functional linear process $(\boldsymbol{\varepsilon}_i \colon i\in\mathbb{Z})$ by

$$\varepsilon_{i,\ell}(s) = \sum_{n=1}^{\infty} \int_0^1 K_{n,\ell}(s,t)\eta_{i-n,\ell}(t)dt + \eta_{i,\ell}(s).$$

In the context of linear processes, condition (4.2) needs to be strengthened in order to obtain $\| \sup_s |\varepsilon_{i,\ell}(s)| \|_4 < \infty$. We assume here that the functions $\boldsymbol{K}_n(s,t)$ satisfy the summability condition

$$\sum_{n=1}^{\infty} n \sup_{0\le s\le 1} \int_0^1 |K_{n,\ell}(s,t)| dt < \infty. \tag{4.4}$$

It follows now from an application of the triangle inequality and (4.4) that, for any $1\le\ell\le d$,

$$\Big\| \sup_{0\le s\le 1} \Big| \sum_{n=1}^{\infty} \int_0^1 K_{n,\ell}(s,t)\eta_{i-n,\ell}(t)dt \Big| \Big\|_4 \le \sum_{n=1}^{\infty} \Big\| \sup_{0\le s\le 1} \Big| \int_0^1 K_{n,\ell}(s,t)\eta_{i-n,\ell}(t)dt \Big| \Big\|_4$$

$$\leq \sum_{n=1}^{\infty} \Big\| \sup_{0\leq t\leq 1} \big|\eta_{i-n,\ell}(t)\big| \sup_{0\leq s\leq 1} \int_0^1 \big|K_{n,\ell}(s,t)\big|dt \Big\|_4$$
$$= \big\| \sup_{0\leq t\leq 1} |\eta_{0,\ell}(t)| \big\|_4 \sum_{n=1}^{\infty} \sup_{0\leq s\leq 1} \int_0^1 \big|K_{n,\ell}(s,t)\big|dt < \infty,$$

and therefore $\|\sup_s |\varepsilon_{i,\ell}(s)|\|_4 < \infty$. In order to verify the conditions imposed on the modulus of continuity, we require additionally that the functions $\boldsymbol{K}_n(s,t)$ satisfy

$$\sum_{n=1}^{\infty} \sup_{0\leq s\leq 1-h} \sup_{0\leq u\leq h} \int_0^1 \big|K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t)\big|dt \to 0 \qquad (h\to 0). \tag{4.5}$$

By similar arguments as in display (4.5) we get

$$\Big\| \sup_{0\leq s\leq 1-h} \sup_{0\leq u\leq h} \Big| \sum_{n=1}^{\infty} \int_0^1 \big[K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t)\big] \eta_{i-n,\ell}(t)dt \Big| \Big\|_4$$
$$\leq \big\| \sup_{0\leq t\leq 1} |\eta_{0,\ell}(t)| \big\|_4 \sup_{0\leq s\leq 1-h} \sup_{0\leq u\leq h} \sum_{n=1}^{\infty} \int_0^1 \big|K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t)\big|dt,$$

and consequently $\|\omega(\varepsilon_{i,\ell};h)\|_4 \to 0$ as $h \to 0$, so that Assumption 3.2 is verified. It remains to be shown that Assumption 3.3 is also satisfied. It is easy to see that

$$\sup_{0\leq s\leq 1} \Big\| \sum_{j=n+1}^{\infty} \int_0^1 K_{j,\ell}(s,t)\eta_{i-j,\ell}(t)dt \Big\|_4 \leq \big\| \sup_{0\leq t\leq 1} |\eta_{0,\ell}(t)| \big\|_4 \sum_{j=n+1}^{\infty} \sup_{0\leq s\leq 1} \int_0^1 |K_{j,\ell}(s,t)|dt,$$

and therefore, by (4.4),

$$\sum_{n=1}^{\infty} \sup_{0\leq s\leq 1} \Big\| \sum_{j=n+1}^{\infty} \int_0^1 K_{j,\ell}(s,t)\eta_{i-j,\ell}(t)dt \Big\|_4$$
$$\leq \big\| \sup_{0\leq t\leq 1} |\eta_{0,\ell}(t)| \big\|_4 \sum_{n=1}^{\infty} \sum_{j=n+1}^{\infty} \sup_{0\leq s\leq 1} \int_0^1 |K_{j,\ell}(s,t)|\, dt$$
$$= \big\| \sup_{0\leq t\leq 1} |\eta_{0,\ell}(t)| \big\|_4 \sum_{n=1}^{\infty} (n+1) \sup_{0\leq s\leq 1} \int_0^1 |K_{n,\ell}(s,t)|\, dt < \infty,$$

which establishes Assumption 3.3.

In the next example we show that the AR(1) process in a Hilbert space satisfies Assumptions 3.1–3.3 if appropriate conditions are imposed which ensure that the AR(1) process can be represented as a functional linear process.

Example 4.4. Let $(\boldsymbol{\eta}_i \colon i \in \mathbb{Z})$ be random processes satisfying the assumptions of Examples 4.2. The functional first-order autoregressive, AR(1), process $(\boldsymbol{\varepsilon}_i \colon i \in \mathbb{Z})$ is defined as the solution of the coordinate-wise recursive equations

$$\varepsilon_{i,\ell}(s) = \int_0^1 K_\ell(s,t)\varepsilon_{i-1,\ell}(t)dt + \eta_{i,\ell}(s), \qquad i \in \mathbb{Z}.$$

Defining $K_{1,\ell}(s,t) = K_\ell(s,t)$ and

$$K_{n,\ell}(s,t) = \int_0^1 K_\ell(s,u) K_{n-1,\ell}(u,t)du, \qquad n \geq 2, \tag{4.6}$$

we obtain the functional linear process representation

$$\varepsilon_{i,\ell}(s) = \sum_{n=1}^{\infty} \int_0^1 K_{n,\ell}(s,t)\eta_{i-n,\ell}(t)dt + \eta_{i,\ell}(s).$$

To establish Assumptions 3.1–3.3 for the functional AR(1) process, it suffices consequently to provide conditions so that the functions in (4.6) satisfy the requirements in (4.4) and (4.5). These conditions are the contraction

$$0 < \rho = \max_{1\leq \ell\leq d} \sup_{0\leq s\leq 1} \int_0^1 \big|K_\ell(s,t)\big|dt < 1 \tag{4.7}$$

and the smoothness assumption

$$\max_{1\leq \ell\leq d} \sup_{0\leq s\leq 1-h} \sup_{0\leq u\leq h} \int_0^1 \big|K_\ell(s+u,t) - K_\ell(s,t)\big|dt \to 0 \quad (h \to 0). \tag{4.8}$$

Using (4.6) and (4.7) we get $\sup_s \int_0^1 |K_{n,\ell}(s,t)|dt \leq \rho^n$, so (4.4) holds. Combining the latter inequality with (4.8) we can further estimate

$$\begin{aligned}\max_{1\leq \ell\leq d} \sup_{0\leq s\leq 1-h} \sup_{0\leq u\leq h} &\int_0^1 \big|K_{n,\ell}(s+u,t) - K_{n,\ell}(s,t)\big|dt \\ &\leq \rho^{n-1} \max_{1\leq \ell\leq d} \sup_{0\leq s\leq 1-h} \sup_{0\leq u\leq h} \int_0^1 \big|K_\ell(s+u,t) - K_\ell(s,t)\big|dt,\end{aligned}$$

and therefore (4.5) is also established. To conclude this example, we would like to mention that the contraction condition in (4.7) can be relaxed to assuming that there is a positive integer n_0 such that $0 < \max_\ell \sup_s \int_0^1 |K_{n_0,\ell}(s,t)|dt < 1$.

As our final example, we introduce the functional version of the multivariate constant correlation ARCH model as introduced in Jeantheau (1998). These processes allow to model heteroskedasticity in the intra-day components.

Example 4.5. Let $(\boldsymbol{\eta}_i : i \in \mathbb{Z})$ be random processes satisfying the assumptions of Examples 4.2. We call $(\boldsymbol{\varepsilon}_i(s) : s \in [0,1], i \in \mathbb{Z})$ a functional constant correlation ARCH process if it is the solution of the coordinate-wise recursions

$$\varepsilon_{i,\ell}(s) = \sigma_{i,\ell}(s)\eta_{i,\ell}(s), \tag{4.9}$$

$$\sigma_{i,\ell}^2(s) = \delta_\ell(s) + \int_0^1 K_\ell(s,t)\varepsilon_{i-1,\ell}^2(t)dt, \tag{4.10}$$

where $\delta_\ell(s) > 0$ and $K_\ell(s,t) \geq 0$. If $K_\ell(s,t) = 0$ if $s \neq t$, then $(\boldsymbol{\varepsilon}_i(s) : s \in [0,1], i \in \mathbb{Z})$ is the functional counterpart of the constant correlation ARCH model proposed by Bollerslev (1990). Hörmann, Horváth and Reeder (2010+) showed that (4.9) and (4.10) have a unique, strictly stationary and non-anticipative (that is, future independent) solution under the moment contraction

$$\max_{1\le\ell\le d} E\left[\Big(\sup_{0\le s\le 1}\Big|\int_0^1 K_\ell(s,t)\eta_{0,\ell}^2(t)dt\Big|\Big)^\alpha\right] < 1 \qquad \text{for some } \alpha > 0.$$

Assumptions 3.1–3.3 are now satisfied if, for all $1 \le \ell \le d$, $\delta_\ell(\cdot)$ is continuous on $[0,1]$ and we have $E[\sup_s |\eta_{0,\ell}(s)|^4] < \infty$ and $\lim_h \|\,\omega(\eta_{0,i};h)\|_4 = 0$. Furthermore it is required that the functions $K_\ell(s,t)$ satisfy $E[(\sup_s \int_0^1 K_\ell(s,t)\eta_{0,\ell}^2(t)dt)^2] < 1$ and

$$\sup_{0\le t\le 1}\ \sup_{0\le s\le 1-h}\ \sup_{0\le u\le h}\big|K_\ell(s+u,t) - K_\ell(s,t)\big| \to 0 \qquad (h \to 0).$$

For details we refer to Hörmann, Horváth and Reeder (2010+).

5 An Application to high-frequency S&P 100 data

In order to illustrate the applicability of the sequential procedures discussed in this paper to high-frequency data, we discuss in this section the monitoring of portfolio betas of several stocks in the S&P 100 over two periods in the years 2001 and 2002. As the market portfolio, we have utilized the S&P 100 index itself. Out of the S&P 100 index, we included for the sake of simplicity only five stocks (hence $d = 5$) from different sectors in our portfolio, namely Boeing (BA), Bank of America (BAC), Microsoft (MSFT), AT&T (T), and Exxon Mobile (XOM).

In the first application, 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 (which was checked with moving windows estimates). Notice that a price change of an asset occurs only in multiples of a preset tick size. The choice of this time span is motivated by the fact that, prior to January 29, 2001, the tick size was one sixteenth of a dollar for NYSE stocks, while all NYSE and AMEX stocks started to traded in decimals thereafter (see, for example, page 181 of Tsay, 2002). The monitoring horizon for the closed-end procedure was selected as 360 days, corresponding to $T = 3$. The intra-day behavior of the process $(\boldsymbol{r}_i(s)\colon s \in [0,S], i \in \mathbb{Z})$, which is at time s defined 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 $r_{M,i}(t)$ is defined analogously.

To apply the monitoring procedure based on τ_m as described in Section 3 involves choosing an appropriate threshold function $w(t)$. We have picked here

$$w_1(t) = (1+t)^2,$$
$$w_2(t) = t\sqrt{3(1+t^2)+t} + \epsilon,$$

where here we work with $\epsilon = 0.1$. As pointed out after Theorem 3.1, the choice of $w_1(t)$ is motivated by the fact that, in this case, the critical constant $c = c(\alpha, d; T)$ can be computed explicitly. The second threshold w_2 was chosen in an effort to minimize the delay between the occurrence of a change in the portfolio betas and its detection through τ_m. To detect a change (which happens soon after the monitoring commences) as early as possible, it is desirable to utilize a weight function which increases only slowly in the right neighborhood of zero. Here, the boundary function $w_1(t)$ is essentially 1 close to zero, while $w_2(t)$ takes approximately the value ϵ. The values of the critical constants $c(\alpha, 5; 3)$ with $\epsilon = 0.1$ for w_2, as required for our data, were determined by Monte Carlo simulations based on 10,000 iterations and are reported in Table 1.

α	0.1	0.05	0.01
w_1	7.85	9.18	12.27
w_2	9.72	11.09	13.99

Table 1: Simulated critical values $c(\alpha, 5; 3)$ based on 10,000 iterations.

The final choice to be made by the analyst concerns the estimation of the long-run covariance matrix $\boldsymbol{D}$ in (3.2), which enters the monitoring procedure through the quadratic form detector V_k. Recall that the stochastic integrals $\boldsymbol{z}_i$ appearing in the definition of $\boldsymbol{D}$ are defined in (3.1). We use here the functional Bartlett estimator $\hat{\boldsymbol{D}}_m$. To precisely give its definition, we first note that the $\boldsymbol{z}_i$ are unobservable, since they depend on the innovations $\boldsymbol{\varepsilon}_i(t)$. Consequently, we replace them with the corresponding residuals in (3.3). The Bartlett estimator is then given by

$$\hat{\boldsymbol{D}}_m = \boldsymbol{S}_m(0) - \sum_{k=1}^{m} H\Big(\frac{k}{q(m)+1}\Big)\big[\boldsymbol{S}_m(k) + \boldsymbol{S}_m(-k)\big],$$

where

$$\boldsymbol{S}_m(k) = \frac{1}{m}\sum_{i=1}^{m-k} \hat{\boldsymbol{z}}_i\, \hat{\boldsymbol{z}}'_{i+k}, \qquad k \geq 0,$$

$\boldsymbol{S}_m(k) = \boldsymbol{S}_m(-k)$, $k < 0$, and $H(u) = (1-|u|)I_{\{|u|\leq 1\}}$ is the Bartlett kernel. Other specifications of $H(u)$ may be entertained as well. For alternative kernel selections as well as background information on the Bartlett estimator we refer to Andrews (1991). Under the conditions of Theorem 3.1, $\hat{\boldsymbol{D}}_m$ satisfies Assumption 3.5. The general requirements on the bandwidth parameter $q(m)$ are the growth conditions $q(m) \to \infty$ and $q(m)/m \to 0$ as $m \to \infty$. In this application we work with $q = 5$.

The results of the thus specified monitoring procedure are the following. As pointed out, we used a 120 day training period, starting January 29, 2001. If no changes in the portfolio betas were detected after 360 additional trading days, the procedure would be terminated. We have found the portfolio betas to be stable prior to the September 11, 2001 terroristic attacks on the World Trade Center. The monitoring procedures based on the threshold functions w_1 and w_2 have, however, rejected the null hypothesis H_0 on September 17, 2001, the first trading day after the attack. The results are displayed graphically in Figure 1, which shows that the detector values (dotted) have clearly crossed the two threshold functions. It is visible from this plot that w_2 is more sensitive than w_1. To further

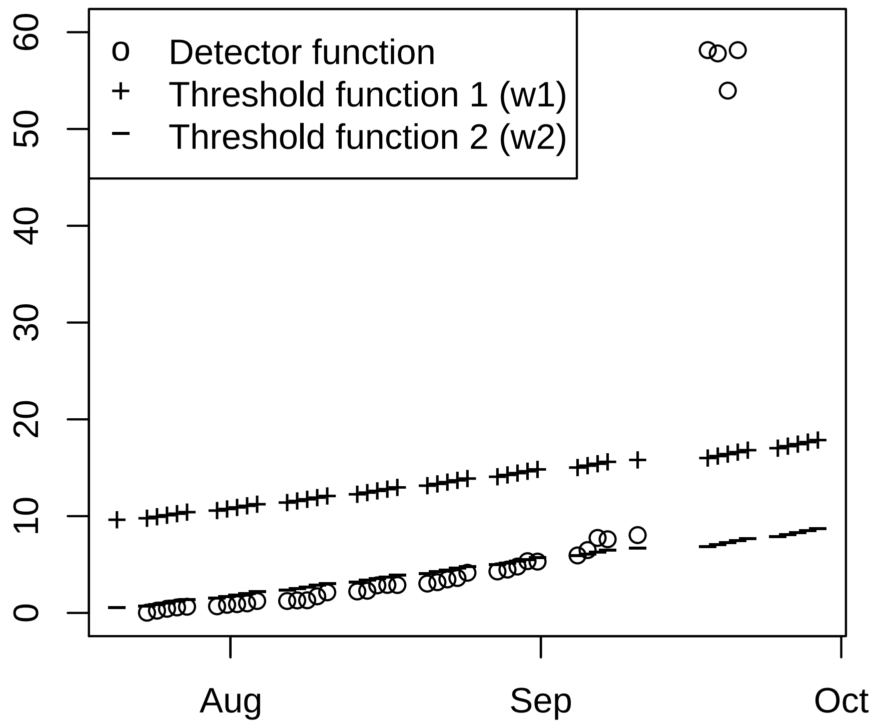

Figure 1: Detector V_k ($\circ\circ\circ$) and threshold functions w_1 ($+++$) and w_2 ($---$) for the monitoring procedure commencing on July 20, 2001. The significance level is set to $\alpha = 0.05$.

analyze the sensitivity we have conducted a second round of monitoring. To avoid the turmoil after the September 11, 2001 attacks, we now used as training period the $m = 80$ days between 9 October, 2001 and 5 February, 2002. The monitoring process would terminate if H_0 had not been rejected after 240 additional trading days (so that still $T = 3$). It can be seen from Figure 2, that the threshold function w_2 is more sensitive to early changes, as instability in the betas was indicated for the corresponding procedure already in the middle of March, 2002, while the procedure based on w_1 terminated about two months thereafter, in the middle of May, 2002. However, removing one (potential outlier) trading day in March from the sample, the first crossing of w_2 happens only in the beginning of May. This indicates that w_2, while being well suited to detect early changes, might be more susceptible to the effect of outliers. The particular type of threshold function has to be determined according to the goals of the sequential testing.

In order to further quantify the finite sample properties of the testing procedure, we have conducted a small simulation study without attempting to give an in-depth analysis. We simulated according

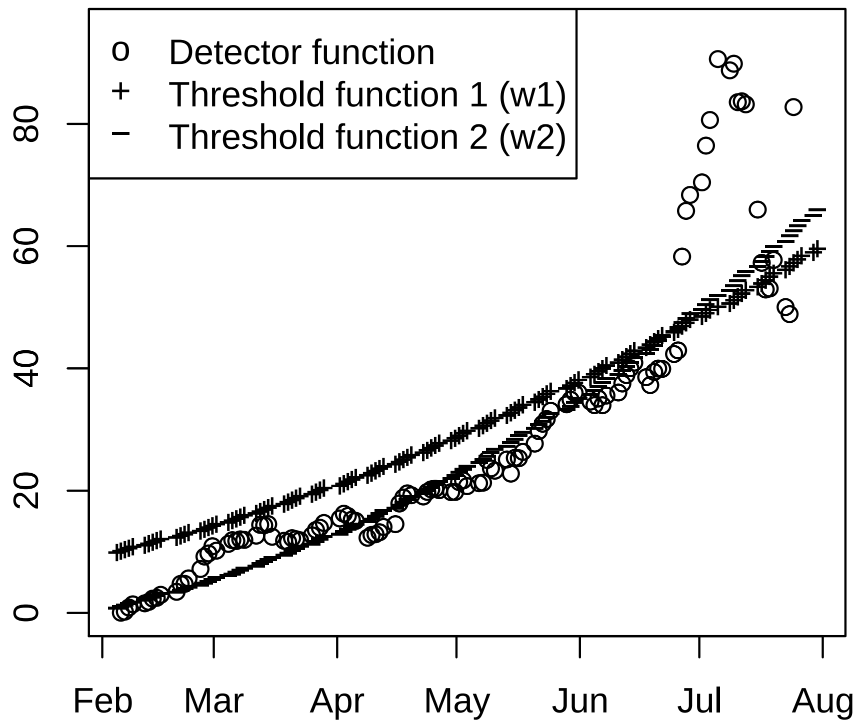

Figure 2: Detector V_k ($\circ\circ\circ$) and threshold functions w_1 ($+++$) and w_2 ($---$) for the monitoring procedure commencing on February 5, 2002. The significance level is set to $\alpha = 0.05$.

to the following version of the functional CAPM (2.2). We let the market portfolio be described by a functional AR(1) process

$$r_{M,i}(s) = \rho \int_0^1 K(s,t) r_{M,i-1}(t) dt + \eta_i(s) \tag{5.1}$$

(see Example 4.4), where $(\eta_i \colon i \in \mathbb{Z})$ denotes a sequence of independent standard Brownian motions and $K(s,t) = c\exp(-|t-s|)$ with c such that the norm of K is equal to one. We chose $\gamma = 0.05$ as the rate of return on the risk-free asset and $d = 5$ different assets with corresponding (null hypothesis) $\boldsymbol{\beta} = (0.8, 1.2, -0.1, 1, 0.6)'$. In addition, we specified the innovations $(\boldsymbol{\varepsilon}_i \colon i \in \mathbb{Z})$ as independent Brownian motions with standard deviation parameter σ controlling the noise level of the CAPM (that is $\boldsymbol{\varepsilon}_i = \sigma B_i$ and $(B_i \colon i \in \mathbb{Z})$ independent standard Brownian motions). The training period is of length $m = 80$ and monitoring is stopped after $n = 350$ observations if no change has been indicated by the stopping rule τ_m which was based on the weight function w_1 from above. The specifications $\rho = 0.1$ and $\rho = 0.4$ were used in (5.1) and in combination with several values of the bandwidth parameter q needed to compute the long-run covariance matrix $\boldsymbol{D}$.

$\rho = 0.1$						
q	1	2	3	4	5	6
$\alpha = 0.01$	0.012	0.007	0.017	0.019	0.024	0.025
$\alpha = 0.05$	0.025	0.034	0.034	0.041	0.064	0.078
$\alpha = 0.10$	0.041	0.073	0.059	0.076	0.088	0.122
$\rho = 0.4$						
q	1	2	3	4	5	6
$\alpha = 0.01$	0.012	0.018	0.024	0.025	0.029	0.055
$\alpha = 0.05$	0.055	0.051	0.051	0.073	0.086	0.136
$\alpha = 0.10$	0.068	0.086	0.099	0.113	0.135	0.193

Table 2: Empirical levels across two dependence levels ρ and various bandwidths q.

Some results for $\sigma = 1$ are displayed in Table 2. All entries are based on $N = 1{,}000$ repetitions. Other specifications of σ were tested as well but produced comparable outcomes and are hence omitted to conserve space. The results indicate that the procedure can be calibrated to the simulation model under consideration. In particular, the bandwidth selection $q = 4$ seems to provide a compromise between both dependence levels. As expected from the application result, simulations run with w_2 instead of w_1 lead to empirical levels that are slightly above the nominal levels.

To assess the behavior under the alternative H_A, we conducted further simulation experiments using the same pre-break parameters as before, but introduced a change-point at $k^* = 10$ after which the post-break regression parameters were given by

$$\boldsymbol{\beta}^* = \delta\boldsymbol{\beta}, \qquad \delta = 1.5, 1.75, 2.$$

The magnitude of the break is therefore $\Delta = \|\boldsymbol{\beta} - \boldsymbol{\beta}^*\| = 1.86|1-\delta|$ and ranges between .930 for $\delta = 1.5$ and 1.86 for $\delta = 2$. Results based on $N = 1{,}000$ repetitions for a significance level of $\alpha = 0.10$ are reported in Table 3. Using the results from Table 2, the bandwidth $q = 5$ ($q = 3$) has been used for the dependence level $\rho = 0.1$ ($\rho = 0.4$). The results show that the procedure has nontrivial power against all break scenarios. As expected, the power increases with the break magnitudes under consideration.

δ	1.50	1.75	2.00
$\rho = 0.1$	0.400	0.741	0.910
$\rho = 0.4$	0.487	0.807	0.956

Table 3: Empirical rejections across two dependence levels ρ and various break sizes δ. The asymptotic level is set to $\alpha = 0.10$.

The simulations reveal a rough picture of the proposed testing procedure, yet much still needs to be explored. That, however, is beyond the scope of the present paper and we leave the investigation of further details to future research.

6 Conclusions

In this paper, we have introduced a flexible functional version of the CAPM that can take into account high-frequency data now often observed in practice. In this framework, we have developed sequential monitoring procedures that allow to test for the constancy of the functional portfolio betas, a proxy for the correlation of an asset with the market. For these monitoring procedures, large sample results have been derived that (a) are applicable to a range of processes of practical interest and (b) can be utilized to tune the monitoring procedures to a pre-specified false alarm level. The performance of the thus obtained procedures has been shown to be satisfactory in a simulation study and an application to intra-day S&P 100 data. Future research may focus on evaluating in more detail the finite sample behavior in order to calibrate the procedures for real-world applications.

A Mathematical proofs

In this section, we provide the proof of Theorem 3.1. It is assumed throughout that the null hypothesis H_0 holds. Denoting with $\boldsymbol{\beta}$ the common values of the portfolio betas under H_0, it follows from the definitions of the least-squares estimators $\hat{\boldsymbol{\beta}}_m$ and $\tilde{\boldsymbol{\beta}}_k$ that

$$\hat{\boldsymbol{\beta}}_m = \boldsymbol{\beta} + \left(\sum_{i=1}^{m} \sum_{j=1}^{J} \big[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j) \big]^2 \right)^{-1} \sum_{i=1}^{m} \sum_{j=1}^{J} \big[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j) \big] \boldsymbol{\varepsilon}_i(s_j), \tag{A.1}$$

$$\tilde{\boldsymbol{\beta}}_k = \boldsymbol{\beta} + \left(\sum_{i=m+1}^{m+k} \sum_{j=1}^{J} \big[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j) \big]^2 \right)^{-1} \sum_{i=m+1}^{m+k} \sum_{j=1}^{J} \big[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j) \big] \boldsymbol{\varepsilon}_i(s_j). \tag{A.2}$$

Recall that $\rho_i(s) = r_{M,i}(s) - E[r_{M,i}(s)]$ for $s \in [0,1]$ and $i \in \mathbb{Z}$. To analyze the fluctuations of the estimators around the true portfolio beta values, we introduce and further decompose the quantities

$$\begin{aligned} \boldsymbol{A}_m &= \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{1}{J} \sum_{j=1}^{J} \big[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j) \big] \boldsymbol{\varepsilon}_i(s_j) \\ &= \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{1}{J} \sum_{j=1}^{J} \rho_i(s_j) \boldsymbol{\varepsilon}_i(s_j) + \frac{1}{m^{3/2}} \frac{1}{J} \sum_{j=1}^{J} \sum_{i=1}^{m} \big[\hat{r}_{M,m}(s_j) - E[r_{M,i}(s_j)] \big] \sum_{n=1}^{m} \boldsymbol{\varepsilon}_n(s_j) \\ &= \boldsymbol{A}_{m,1} + \boldsymbol{A}_{m,2} \end{aligned}$$

and, for $t \in [0,1]$,

$$\boldsymbol{B}_m(t) = \frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+\lfloor mt \rfloor} \frac{1}{J} \sum_{j=1}^{J} \big[r_{M,i}(s_j) - \tilde{r}_{M,\lfloor mt \rfloor}(s_j) \big] \boldsymbol{\varepsilon}_i(s_j)$$

$$= \frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+\lfloor mt \rfloor} \frac{1}{J} \sum_{j=1}^{J} \rho_i(s_j) \boldsymbol{\varepsilon}_i(s_j) + \frac{1}{m^{3/2}} \frac{1}{J} \sum_{j=1}^{J} \sum_{i=m+1}^{m+\lfloor mt \rfloor} \left[\tilde{r}_{M,\lfloor mt \rfloor}(s_j) - E[r_{M,i}(s_j)]\right] \sum_{n=m+1}^{m+\lfloor mt \rfloor} \boldsymbol{\varepsilon}_n(s_j)$$

$$= \boldsymbol{B}_{m,1}(t) + \boldsymbol{B}_{m,2}(t).$$

We start with deriving the large sample behavior of the terms $\boldsymbol{A}_{m,1}$ and $\boldsymbol{B}_{m,1}(t)$. For these we write

$$\boldsymbol{A}_{m,1} = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \int_0^1 \rho_i(s) \boldsymbol{\varepsilon}_i(s) ds + \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \boldsymbol{\xi}_i(J) = \boldsymbol{A}_{m,1,1} + \boldsymbol{A}_{m,1,2},$$

$$\boldsymbol{B}_{m,1}(t) = \frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+\lfloor mt \rfloor} \int_0^1 \rho_i(s) \boldsymbol{\varepsilon}_i(s) ds + \frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+\lfloor mt \rfloor} \boldsymbol{\xi}_i(J) = \boldsymbol{B}_{m,1,1}(t) + \boldsymbol{B}_{m,1,2}(t),$$

where the process $(\boldsymbol{\xi}_i(J)\colon i \in \mathbb{Z})$ is defined by

$$\boldsymbol{\xi}_i(J) = \sum_{j=1}^{J} \int_{s_{j-1}}^{s_j} \left[\rho_i(s_j) \boldsymbol{\varepsilon}_i(s_j) - \rho_i(s) \boldsymbol{\varepsilon}_i(s)\right] ds.$$

In the first lemma, we derive the asymptotics for partial sums of the $\boldsymbol{\xi}_i(J)$'s.

Lemma A.1. *If Assumptions 3.1–3.4 are satisfied, then, for all $T > 0$,*

$$\frac{1}{\sqrt{m}} \max_{1 \le k \le mT} \Big| \sum_{i=1}^{k} \boldsymbol{\xi}_i(J) \Big| \xrightarrow{P} 0 \qquad (m \to \infty).$$

Proof. We establish the assertion of the lemma coordinate-wise. To simplify notation, let ε_i denote any of the d coordinate processes of $\boldsymbol{\varepsilon}_i = (\varepsilon_{i,1}, \ldots, \varepsilon_{i,d})'$ and adopt similar notations for any other d-dimensional quantity. Now write

$$\xi_i(J) = \sum_{j=1}^{J} \int_{s_{j-1}}^{s_j} \rho_i(s) \left[\varepsilon_i(s_j) - \varepsilon_i(s)\right] ds + \sum_{j=1}^{J} \int_{s_{j-1}}^{s_j} \varepsilon_i(s_j) \left[\rho_i(s_j) - \rho_i(s)\right] ds = \xi_{i,1}(J) + \xi_{i,2}(J).$$

Let $1 \le N < K < M \le \lfloor mT \rfloor$. Recall that, by Assumption 3.4, the number of intra-day observations J tends to infinity. In the following, we will show that there is a sequence $b(J) \to 0$ as $J \to \infty$, depending only on J, such that

$$E\left[\Big(\sum_{i=N+1}^{K} \xi_{i,1}(J)\Big)^2 \Big(\sum_{i=K+1}^{M} \xi_{i,1}(J)\Big)^2\right] \le b(J)(K-N)(M-K). \tag{A.3}$$

As a first step towards the verification of (A.3) we introduce the approximating sequence

$$\xi_{i,1}^{(L)}(J) = \sum_{j=1}^{J} \int_{s_{j-1}}^{s_j} \rho_i^{(L)}(s) \left[\varepsilon_i^{(L)}(s_j) - \varepsilon_i^{(L)}(s)\right] ds,$$

where $\rho_i^{(L)}(s) = r_{M,i}^{(L)}(s) - E[r_{M,i}(s)]$, $s \in [0,1]$. It follows from the definition of $r_{M,i}^{(L)}$ that $E[r_{M,i}^{(L)}(s)] = E[r_{M,0}(s)]$ for all $s \in [0,1]$ and $i, L \in \mathbb{Z}$. We assume without loss of generality that $K - N \le M - K$. It is easy to see that

$$E\left[\Big(\sum_{i=N+1}^{K} \xi_{i,1}(J)\Big)^2 \Big(\sum_{i=K+1}^{M} \xi_{i,1}(J)\Big)^2\right]$$

$$= E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\Big(\sum_{i=K+1}^{M}\big[\xi_{i,1}(J)-\xi_{i,1}^{(i-K)}(J)\big]+\sum_{i=K+1}^{M}\xi_{i,1}^{(i-K)}(J)\Big)^2\right]$$

$$\le 2E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\Big(\sum_{i=K+1}^{M}\big[\xi_{i,1}(J)-\xi_{i,1}^{(i-K)}(J)\big]\Big)^2\right]$$

$$+ 2E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\Big(\sum_{i=K+1}^{M}\xi_{i,1}^{(i-K)}(J)\Big)^2\right]. \qquad (A.4)$$

Subsequent applications of the Cauchy-Schwarz and triangle inequalities yield that

$$E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\Big(\sum_{i=K+1}^{M}\big[\xi_{i,1}(J)-\xi_{i,1}^{(i-K)}(J)\big]\Big)^2\right]$$

$$\le \left(E\Big[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^4\Big]\right)^{1/2}\left(E\Big[\Big(\sum_{i=K+1}^{M}\big[\xi_{i,1}(J)-\xi_{i,1}^{(i-K)}(J)\big]\Big)^4\Big]\right)^{1/2}$$

$$= \Big\|\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big\|_4^2\Big\|\sum_{i=K+1}^{M}\big[\xi_{i,1}(J)-\xi_{i,1}^{(i-K)}(J)\big]\Big\|_4^2$$

$$\le \Big(\sum_{i=N+1}^{K}\big\|\xi_{i,1}(J)\big\|_4\Big)^2\Big(\sum_{L=1}^{\infty}\Big\|\xi_{i,1}(J)-\xi_{i,1}^{(L)}(J)\Big\|_4\Big)^2$$

$$= R_1R_2.$$

By stationarity we have $R_1 = (K-N)\|\xi_{0,1}(J)\|_4$ and Assumption 3.2 yields that $\|\xi_{0,1}(J)\|_4 \to 0$ as $J \to \infty$. Assumption 3.3 implies that $R_2 < \infty$. Hence it is proved that (note that $M-K \ge 1$)

$$E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\Big(\sum_{i=K+1}^{M}\big[\xi_{i,1}(J)-\xi_{i,1}^{(i-K)}(J)\big]\Big)^2\right] \le b_1(J)(K-N)(M-K) \qquad (A.5)$$

with an appropriately defined $b_1(J) \to 0$ as $J \to \infty$. Independence leads next to

$$E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\Big(\sum_{i=K+1}^{M}\xi_{i,1}^{(i-K)}(J)\Big)^2\right] = E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\right]E\left[\Big(\sum_{i=K+1}^{M}\xi_{i,1}^{(i-K)}(J)\Big)^2\right],$$

so to prove (A.3) it suffices to show that

$$E\left[\Big(\sum_{i=N+1}^{K}\xi_{i,1}(J)\Big)^2\right] \le b_2(J)(K-N) \quad \text{and} \quad E\left[\Big(\sum_{i=K+1}^{M}\xi_{i,1}(J)\Big)^2\right] \le b_3(J)(M-K) \qquad (A.6)$$

with some $b_2(J) \to 0$ and $b_3(J) \to 0$ as $J \to \infty$. We only give the proof of the first statement in (A.6) here, since the second can be validated with similar arguments. Use stationarity to write

$$E\left[\Big(\sum_{i=N+1}^{M}\xi_{i,1}(J)\Big)^2\right] = (M-N)E\big[\xi_{0,1}^2(J)\big] + 2\sum_{i=1}^{M-N}\sum_{j=i+1}^{M-N}E[\xi_{i,1}(J)\xi_{j,1}(J)].$$

Assumption 3.2 implies that $E[\xi_{0,1}^2(J)] \to 0$ as $J \to \infty$. Utilizing stationarity once more, it follows that the second term in the last display satisfies

$$\Big|\sum_{i=1}^{M-N}\sum_{j=i+1}^{M-N} E[\xi_{i,1}(J)\xi_{j,1}(J)]\Big| \le \sum_{i=1}^{M-N}\sum_{j=i+1}^{M-N}\big|E[\xi_{i,1}(J)\xi_{j,1}(J)]\big|$$
$$\le (M-N)\sum_{L=1}^{\infty}\big|E[\xi_{0,1}(J)\xi_{L,1}(J)]\big| \le (M-N)\sum_{L=1}^{\infty}\big|E\big[\xi_{0,1}(J)[\xi_{L,1}(J)-\xi_{L,1}^{(L)}(J)]\big]\big|,$$

since $E[\xi_{0,1}(J)\xi_{L,1}^{(L)}(J)] = E[\xi_{0,1}(J)E\xi_{L,1}^{(L)}] = 0$. Using the Cauchy-Schwarz inequality in combination with Assumption 3.3, we conclude that

$$\sum_{L=1}^{\infty}\big|E\big[\xi_{0,1}(J)[\xi_{L,1}(J)-\xi_{L,1}^{(L)}(J)]\big]\big| \le \sqrt{E[\xi_{0,1}^2(J)]}\sum_{L=1}^{\infty}\big\|\xi_{L,1}(J)-\xi_{L,1}^{(L)}(J)\big\|_2,$$

where $E[\xi_{0,1}^2(J)] \to 0$ as $J \to \infty$. Moreover, it follows from Assumption 3.3 that the sum on the right-hand side of the last inequality is finite. Hence (A.6) is established and (A.3) follows from combining the results in (A.4)–(A.6). To complete the proof of the lemma, observe that Theorem 12.1 of Billingsley (1968, p. 89) and (A.3) imply that, for all $x > 0$,

$$P\left(\max_{1\le k\le mT}\Big|\sum_{i=1}^{k}\xi_{i,1}(J)\Big| \ge \sqrt{m}x\right) \le \frac{C(b(J)m)^2}{(\sqrt{m}x)^4} \to 0 \qquad (m, J \to \infty)$$

with some $C > 0$. Since a similar statement is also true for the partial sums of $\xi_{i,2}$, we obtain

$$\frac{1}{\sqrt{m}}\max_{1\le k\le mT}\Big|\sum_{i=1}^{k}\xi_{i,1}\Big| \stackrel{P}{\to} 0 \qquad\text{and}\qquad \frac{1}{\sqrt{m}}\max_{1\le k\le mT}\Big|\sum_{i=1}^{k}\xi_{i,2}\Big| \stackrel{P}{\to} 0 \qquad (m\to\infty),$$

which implies the assertion of Lemma A.1. □

Lemma A.2. *If Assumptions 3.1–3.4 are satisfied, then* $\boldsymbol{A}_{m,1,2} \stackrel{P}{\to} 0$ *and, for all* $T > 0$,

$$\sup_{0\le t\le T}|\boldsymbol{B}_{m,1,2}(t)| \stackrel{P}{\to} 0 \qquad (m\to\infty). \tag{A.7}$$

Proof. The first statement of Lemma A.2 is immediately implied by Lemma A.1. Since, by stationarity,

$$\big(\boldsymbol{B}_{m,1,2}(t)\colon t\in[0,T]\big) \stackrel{\mathcal{D}}{=} \Big(\frac{1}{\sqrt{m}}\sum_{i=1}^{\lfloor mT\rfloor}\boldsymbol{\xi}_i(J)\colon t\in[0,T]\Big),$$

the second statement (A.7) follows also from Lemma A.1. □

Lemma A.3. *If Assumptions 3.1–3.4 are satisfied, then, for all* $T > 0$,

$$\max_{1\le k\le mT}\frac{1}{m^{3/2}}\frac{1}{J}\Big|\sum_{j=1}^{J}\sum_{i=1}^{k}\rho_i(s_j)\sum_{n=1}^{k}\boldsymbol{\varepsilon}_n(s_j)\Big| \stackrel{P}{\to} 0 \qquad (m\to\infty).$$

Proof. As in the proof of Lemma A.1, for the sake of notational simplicity, we denote by ε_n a generic coordinate of the d-dimensional $\boldsymbol{\varepsilon}_n$. It then follows that

$$\frac{1}{J}\sum_{j=1}^{J}\sum_{i=1}^{k}\rho_i(s_j)\sum_{n=1}^{k}\varepsilon_n(s_j)=\sum_{i=1}^{k}\sum_{n=1}^{k}\gamma_{i,n},\qquad \gamma_{i,n}=\frac{1}{J}\sum_{j=1}^{J}\rho_i(s_j)\varepsilon_n(s_j).$$

By independence and stationarity we obtain, for all i,n, that

$$E[\gamma_{i,n}^2]=\frac{1}{J^2}\sum_{j=1}^{J}\sum_{j'=1}^{J}E\big[\rho_i(s_j)\rho_i(s_{j'})\varepsilon_n(s_j)\varepsilon_n(s_{j'})\big]\le C \tag{A.8}$$

with some constant C. Similarly, for all $g,h>0$,

$$\begin{aligned}E[\gamma_{i,n}\gamma_{i+g,n+h}]&=\frac{1}{J^2}\sum_{j=1}^{J}\sum_{j'=1}^{J}E\big[\rho_i(s_j)\rho_{i+g}(s_{j'})\varepsilon_n(s_j)\varepsilon_{n+h}(s_{j'})\big]\\&=\frac{1}{J^2}\sum_{j=1}^{J}\sum_{j'=1}^{J}E\big[\rho_0(s_j)\rho_g(s_{j'})\big]E\big[\varepsilon_0(s_j)\varepsilon_h(s_{j'})\big].\end{aligned}$$

Following the arguments in the proof of Lemma A.1, the Cauchy-Schwarz inequality yields

$$\big|E[\rho_0(s_j)\rho_g(s_{j'})]\big|=\big|E\big[\rho_0(s_j)[\rho_g(s_{j'})-\rho_g^{(g)}(s_{j'})]\big]\big|\le C\sup_{0\le s\le 1}\big\|r_{M,0}-r_{M,0}^{(g)}\big\|_2 \tag{A.9}$$

and similarly,

$$\big|E[\varepsilon_0(s_j)\varepsilon_h(s_{j'})]\big|\le C\sup_{0\le s\le 1}\big\|\varepsilon_0(s)-\varepsilon_0^{(h)}(s)\big\|_2 \tag{A.10}$$

with some constant C. Putting together the estimates in (A.8)–(A.10) with Assumption 3.3, we conclude that there is a constant C such that for all $K<N$ and $M<L$

$$E\Big[\sum_{i=K+1}^{N}\sum_{n=M+1}^{L}\gamma_{i,n}\Big]^2\le C(N-K)(L-M).$$

Now, using Menshov's inequality (see Móricz, 1977) we obtain that

$$\max_{1\le k\le mT}\Big|\sum_{i=1}^{k}\sum_{n=1}^{k}\gamma_{i,n}\Big|=O_P\left(m(\log m)^2\right)\qquad(m\to\infty),$$

which completes the proof of the lemma. □

Lemma A.4. *If Assumptions 3.1–3.4 are satisfied, then* $\boldsymbol{A}_{m,2}\xrightarrow{P}0$ *and, for any* $T>0$,

$$\sup_{0\le t\le T}\big|\boldsymbol{B}_{m,2}(t)\big|\xrightarrow{P}0\qquad(m\to\infty).$$

Proof. The result follows from Lemma A.3 by stationarity. □

According to Lemmas A.2 and A.4 we need to consider only the weak convergence of the process $(\boldsymbol{Z}_m(t)\colon t \in [0,T])$ given by

$$\boldsymbol{Z}_m(t) = \frac{1}{\sqrt{m}} \sum_{i=1}^{\lfloor mt \rfloor} \boldsymbol{z}_i,$$

where $\boldsymbol{z}_i$ is defined in (3.1). We do so in the next lemma.

Lemma A.5. *If Assumptions 3.1–3.4 are satisfied, then*

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

where $(\boldsymbol{W}_{\boldsymbol{D}}(t)\colon t \in [0,T])$ *is a Gaussian process with mean vector* $E[\boldsymbol{W}_{\boldsymbol{D}}(t)] = \mathbf{0}$ *and covariance function* $E[\boldsymbol{W}_{\boldsymbol{D}}(t)\boldsymbol{W}'_{\boldsymbol{D}}(s)] = \boldsymbol{D}\min(t,s)$, *and* $\overset{\mathcal{D}^d[0,T]}{\longrightarrow}$ *denotes weak convergence in the Skorohod space* $\mathcal{D}^d[0,T]$.

Proof. The proof is based on Theorem 21.1 of Billingsley (1968, p. 184). Let

$$\boldsymbol{z}_i^{(L)} = \big(z_{i,1}^{(L)}, \ldots, z_{i,d}^{(L)}\big)' = \int_0^1 \rho_i^{(L)}(s)\boldsymbol{\varepsilon}_i(s)ds.$$

Using the Cauchy-Schwartz inequality we get, for any $1 \le \ell \le d$,

$$\begin{aligned}
\big\| z_{i,\ell} - z_{i,\ell}^{(L)} \big\|_2 &= \Big(E\Big[\int_0^1 \big[\rho_i(s)\varepsilon_{i,\ell}(s) - \rho_i^{(L)}(s)\varepsilon_{i,\ell}^{(L)}(s)\big]ds\Big]^2\Big)^{1/2} \\
&\le \Big(E\Big[\int_0^1 \varepsilon_{i,\ell}(s)\big[\rho_i(s) - \rho_i^{(L)}(s)\big)ds\Big]^2\Big)^{1/2} \\
&\quad + \Big(E\Big[\int_0^1 \rho_i^{(L)}(s)\big[\varepsilon_{i,\ell}(s) - \varepsilon_{i,\ell}^{(L)}(s)\big]ds\Big]^2\Big)^{1/2} \\
&\le \Big(E\Big[\int_0^1 \varepsilon_{i,\ell}^2(s)ds \int_0^1 \big[\rho_i(s) - \rho_i^{(L)}(s)\big]^2 ds\Big]\Big)^{1/2} \\
&\quad + \Big(E\Big[\int_0^1 \big[\rho_i^{(L)}(s)\big]^2 ds \int_0^1 \big[\varepsilon_{i,\ell}(s) - \varepsilon_{i,\ell}^{(L)}(s)\big]^2 ds\Big]\Big)^{1/2} \\
&= \Big(\int_0^1 E[\varepsilon_{i,\ell}^2(s)]ds\Big)^{1/2}\Big(\int_0^1 E\big[\big(\rho_i(s) - \rho_i^{(L)}(s)\big)^2\big]ds\Big)^{1/2} \\
&\quad + \Big(\int_0^1 E\big[\big(\rho_i^{(L)}(s)\big)^2\big]ds\Big)^{1/2}\Big(\int_0^1 E\big[\big(\varepsilon_{i,\ell}(s) - \varepsilon_{i,\ell}^{(L)}(s)\big)^2\big]ds\Big)^{1/2}.
\end{aligned}$$

Now Assumption 3.3 gives that, for all $1 \le \ell \le d$,

$$\sum_{L=1}^{\infty} \big\| z_{i,\ell} - z_{i,\ell}^{(L)} \big\|_2 < \infty.$$

Applying Theorem 21.1 of Billingsley (1968, p. 184) we conclude that all coordinates of the processes $(\boldsymbol{Z}_m(t)\colon t \in [0,T])$ are tight and that therefore $(\boldsymbol{Z}_m(t)\colon t \in [0,T])$ is also tight. For any set of constants $(c_1, \ldots, c_d)$ we obtain that

$$\sum_{L=1}^{\infty} \Big\| \sum_{\ell=1}^{d} c_\ell z_{i,\ell} - \sum_{\ell=1}^{d} c_\ell z_{i,\ell}^{(L)} \Big\|_2 < \infty,$$

so the normality of the finite-dimensional distributions is again implied by Billingsley (1968, p. 184). The proof is thus complete. □

Theorem A.1. *If Assumptions 3.1–3.4 are satisfied, then, for all $T > 0$,*

$$\big(\boldsymbol{A}_m, \boldsymbol{B}_m(t)\colon t \in [0,T]\big) \overset{\mathcal{D}^{2d}[0,T]}{\longrightarrow} \Big(\boldsymbol{W}_{\boldsymbol{D}}^{(1)}(1), \boldsymbol{W}_{\boldsymbol{D}}^{(2)}(t)\colon t \in [0,T]\Big) \qquad (m \to \infty),$$

where $(\boldsymbol{W}_{\boldsymbol{D}}^{(1)}(t)\colon t \geq 0)$ and $(\boldsymbol{W}_{\boldsymbol{D}}^{(2)}(t)\colon t \geq 0)$ are independent copies of $(\boldsymbol{W}_{\boldsymbol{D}}(t)\colon t \geq 0)$ defined in Lemma A.5.

Proof. The result follows immediately from Lemmas A.2, A.4 and A.5, and from the fact that $\boldsymbol{W}_{\boldsymbol{D}}(1)$ and $(\boldsymbol{W}_{\boldsymbol{D}}(t+1) - \boldsymbol{W}_{\boldsymbol{D}}(1)\colon t \geq 0)$ are independent. Moreover,

$$\big(\boldsymbol{W}_{\boldsymbol{D}}(t+1) - \boldsymbol{W}_{\boldsymbol{D}}(1)\colon t \geq 0\big) \overset{\mathcal{D}}{=} (\boldsymbol{W}_{\boldsymbol{D}}(t)\colon t \geq 0).$$

thus completing the proof. □

Lemma A.6. *If Assumptions 3.1–3.4 are satisfied, then*

$$\frac{1}{mJ}\sum_{i=1}^{m}\sum_{j=1}^{J}\big[r_{M,i}(s_j) - \hat{r}_{M,m}(s_j)\big]^2 \to \int_0^1 \mathrm{var}(r_{M,0}(s))ds \qquad (m \to \infty).$$

For any $x > 0, y > 0$ there are $N_0 = N_0(x,y)$ and $J_0 = J_0(x,y)$ such that

$$P\left(\sup_{N \leq k < \infty}\Big|\frac{1}{kJ}\sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\big[r_{M,i}(s_j) - \tilde{r}_{M,k}(s_j)\big]^2 - \int_0^1 \mathrm{var}\big(r_{M,0}(s)\big)ds\Big| \geq x\right) \leq y$$

for all $N \geq N_0$ and $J \geq J_0$.

Proof. We only provide the proof of the second statement because it implies the first. Elementary arguments give

$$\Big|\frac{1}{J}\sum_{j=1}^{J}\rho_i^2(s_j) - \int_0^1 \rho_i^2(s)ds\Big| \leq \sup_{0 \leq s \leq 1-h}\sup_{0 \leq u \leq h}\big|\rho_i^2(s+u) - \rho_i^2(s)\big|$$

for all $h \geq 1/J$. The ergodic theorem yields that, with probability one and for all $h \geq 1/J$,

$$\begin{aligned}
&\limsup_{k\to\infty}\frac{1}{k}\Big(\sum_{i=1}^{k}\Big|\frac{1}{J}\sum_{j=1}^{J}\rho_i^2(s_j) - \int_0^1 \rho_i^2(s)ds\Big|\Big) \\
&\leq \limsup_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\sup_{0 \leq s \leq 1-h}\sup_{0 \leq u \leq h}\big|\rho_i^2(s+u) - \rho_i^2(s)\big| \\
&= E\Big[\sup_{0 \leq s \leq 1-h}\sup_{0 \leq u \leq h}\big|\rho_0^2(s+u) - \rho_0^2(s)\big|\Big],
\end{aligned}$$

since $(\sup_{s\le 1-h}\sup_{u\le h}|\rho_i^2(s+u)-\rho_i^2(s)|\colon i\in\mathbb{Z})$ is a stationary and ergodic sequence with finite mean. By Assumption 3.2 we have that

$$\lim_{h\to 0} E\Big[\sup_{0\le s\le 1-h}\sup_{0\le u\le h}\big|\rho_0^2(s+u)-\rho_0^2(s)\big|\Big]=0.$$

Thus we obtain

$$\Big|\frac{1}{k}\sum_{i=1}^{k}\frac{1}{J}\sum_{j=1}^{J}\rho_i^2(s_i)-\frac{1}{k}\sum_{i=1}^{k}\int_0^1\rho_i^2(s)ds\Big|\to 0 \qquad \text{a.s.}$$

as $k\to\infty$ and $J\to\infty$. A second application of the ergodic theorem yields, as $k\to\infty$,

$$\frac{1}{k}\sum_{i=1}^{k}\int_0^1\rho_i^2(s)\to E\Big[\int_0^1\rho_0^2(s)ds\Big]=\int_0^1\mathrm{var}\big(r_{M,0}(s)\big)ds \qquad \text{a.s.}$$

Similar arguments give

$$\frac{1}{J}\sum_{j=1}^{J}\Big(\frac{1}{k}\sum_{i=1}^{k}\big[r_{M,i}(s_j)-E[r_{M,i}(s_j)]\big]\Big)^2\to 0 \qquad \text{a.s.}$$

Hence we conclude

$$\sup_{N\le k<\infty}\Big|\frac{1}{kJ}\sum_{i=1}^{k}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\hat r_{M,k}(s_j)\big]^2-\int_0^1\mathrm{var}\big(r_{M,0}(s)\big)ds\Big|\to 0 \qquad \text{a.s.} \tag{A.11}$$

Now the second assertion of the lemma follows from (A.11) via stationarity. □

Proof of Theorem 3.1. We show that

$$\frac{1}{\sqrt{m}}\tilde U_{\lfloor mt\rfloor}\big(\tilde{\boldsymbol\beta}_{\lfloor mt\rfloor}-\hat{\boldsymbol\beta}_m\big)=\frac{1}{\sqrt{m}}\boldsymbol{B}_m(t)-\frac{\tilde U_{\lfloor mt\rfloor}}{\hat U_m}\boldsymbol{A}_m,$$

where $\tilde U_k$ and $\hat U_m$ are as defined in (2.3) and (2.5), respectively. An application of Lemma A.6 leads to

$$\sup_{0\le t\le T}\Big|\frac{\tilde U_{\lfloor mt\rfloor}}{\hat U_m}-t\Big|\xrightarrow{P}0$$

as $m\to\infty$ and $J\to\infty$. Theorem A.1 implies that

$$\Big(\frac{1}{\sqrt{m}}\tilde U_{\lfloor mt\rfloor}\big(\tilde{\boldsymbol\beta}_{\lfloor mt\rfloor}-\hat{\boldsymbol\beta}_m\big)\colon t\in[0,T]\Big)\xrightarrow{\mathcal{D}^d[0,T]}\Big(\boldsymbol{W}_{\boldsymbol D}^{(2)}(t)-t\boldsymbol{W}_{\boldsymbol D}^{(1)}(1)\colon t\in[0,1]\Big).$$

Comparing covariance functions, we obtain that

$$\Big(\big[\boldsymbol{W}_{\boldsymbol D}^{(2)}(t)-t\boldsymbol{W}_{\boldsymbol D}^{(1)}(1)\big]'\boldsymbol{D}^{-1}\big[\boldsymbol{W}_{\boldsymbol D}^{(2)}(t)-t\boldsymbol{W}_{\boldsymbol D}^{(1)}(1)\big]\colon t\in[0,T]\Big)\stackrel{\mathcal{D}}{=}\left(\Gamma_d(t)\colon t\in[0,T]\right),$$

resulting in

$$\Big(\frac{1}{m}\tilde U_{\lfloor mt\rfloor}^2\big(\tilde{\boldsymbol\beta}_{\lfloor mt\rfloor}-\hat{\boldsymbol\beta}_m\big)'\boldsymbol{D}^{-1}(\tilde{\boldsymbol\beta}_{\lfloor mt\rfloor}-\hat{\boldsymbol\beta}_m)\colon t\in[0,T]\Big)\xrightarrow{\mathcal{D}[0,T]}(\Gamma_d(t)\colon t\in[0,T]). \tag{A.12}$$

According to Slutsky's lemma and Assumption 3.5, (A.12) remains true if $\boldsymbol{D}^{-1}$ is replaced by $\hat{\boldsymbol{D}}_m^{-1}$. Thus, we obtain from the continuous mapping theorem that

$$\sup_{0\le t\le T}\frac{1}{w(t)}\frac{1}{m}\tilde{U}^2_{\lfloor mt\rfloor}\big(\tilde{\boldsymbol{\beta}}_{\lfloor mt\rfloor}-\hat{\boldsymbol{\beta}}_m\big)'\hat{\boldsymbol{D}}_m^{-1}\big(\tilde{\boldsymbol{\beta}}_{\lfloor mt\rfloor}-\hat{\boldsymbol{\beta}}_m\big)\ \xrightarrow{\mathcal{D}}\ \sup_{0\le t\le T}\frac{\Gamma_d(t)}{w(t)},$$

completing the proof. □

Proof of Theorem 3.2. In the proof, we allow even the parameter γ to change to γ^* at lag k^*. For any $k>k^*$, we have

$$\begin{aligned}
&\sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\tilde{r}_{M,k}(s_j)\big]\big[\boldsymbol{r}_i(s_j)-\tilde{\boldsymbol{r}}_k(s_j)\big]\\
&\quad=(\mathbf{1}-\boldsymbol{\beta})\gamma\sum_{i=m+1}^{m+k^*}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\tilde{r}_{M,k}(s_j)\big]+(\mathbf{1}-\boldsymbol{\beta}^*)\gamma^*\sum_{i=m+k^*+1}^{m+k}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\tilde{r}_{M,k}(s_j)\big]\\
&\qquad+\boldsymbol{\beta}\sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\tilde{r}_{M,k}(s_j)\big]^2+(\boldsymbol{\beta}^*-\boldsymbol{\beta})\sum_{i=m+k^*+1}^{m+k}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\tilde{r}_{M,k}(s_j)\big]r_{M,i}(s_j)\\
&\qquad+\sum_{i=m+1}^{m+k}\sum_{j=1}^{J}\big[r_{M,i}(s_j)-\tilde{r}_{M,k}(s_j)\big]\boldsymbol{\varepsilon}_i(s_j)\\
&\quad=\tilde{L}_1+\tilde{L}_2+\tilde{L}_3+(\boldsymbol{\beta}^*-\boldsymbol{\beta})\tilde{L}_4+\tilde{L}_5.
\end{aligned}$$

Thus,

$$\tilde{\boldsymbol{\beta}}_k-\hat{\boldsymbol{\beta}}_m=(\boldsymbol{\beta}^*-\boldsymbol{\beta})L_4+L_1+L_2+L_3+L_5-L_6,$$

where $L_j=\tilde{L}_j/\tilde{U}_k$, $1\le j\le 5$ with $\tilde{U}_k$ from (2.3) and $L_6=\hat{\boldsymbol{\beta}}_m-\boldsymbol{\beta}$. Using the ergodic theorem as in the proof of Theorem 3.1 leads to

$$\frac{kL_4}{k-k^*}\xrightarrow{P}1$$

as $J,k,k-k^*\to\infty$. As in Theorem 3.1 one can show that all other terms L_j, $j\neq 4$, are negligible. Combining these estimates, we obtain that

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{V_k}{w(k/m)}=\Big(\int_0^1\mathrm{var}(r_{M,0}(s))ds\Big)^2\max_{1\le k\le\lfloor mT\rfloor}\frac{(k-k^*)^2}{mw(k/m)}(\boldsymbol{\beta}^*-\boldsymbol{\beta})'D^{-1}(\boldsymbol{\beta}^*-\boldsymbol{\beta})(1+o_P(1)).$$

Assumption 3.6 implies that $\inf\{w(t)\colon t\in[0,T]\}>0$ and consequently,

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{(k-k^*)^2}{mw(k/m)}(\boldsymbol{\beta}^*-\boldsymbol{\beta})'D^{-1}(\boldsymbol{\beta}^*-\boldsymbol{\beta})\ge c(T-T^*)^2m\|\boldsymbol{\beta}^*-\boldsymbol{\beta}\|^2\xrightarrow{P}\infty.$$

This completes the proof. □

References

Amsler, C.E. and Schmidt, P. (1985) A Monte Carlo investigation of the accuracy of multivariate CAPM tests. *Journal of Financial Economics* **14**, 393–375.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Wu, J. (2006) 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, Volume B, 1–40.

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

Andrews D.W.K. Monahan, J.C. (1992) An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* **60**, 953–966.

Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009) Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics* **37**, 4046–4087.

Aue, A. and Horváth, L. (2004) Delay time in sequential detection of change. *Statistics & Probability Letters* **67**, 221–231.

Aue, A., Horváth, L., Hušková, M. and Kokoszka, P. (2006a) Change-point monitoring in linear models. *Econometrics Journal* **9**, 373–403.

Aue, A., Horváth, L., Kokoszka, P. and Steinebach, J. (2006b) Detecting changes in mean: asymptotic normality of stopping times. *Test* **17**, 515–530.

Aue, A., Horváth, L. and Reimherr, M.L. (2009) Delay time of sequential procedures for multiple time series regression models. *Journal of Econometrics* **149**, 174–190.

Balduzzi, P. and Robotti, C. (2008) Mimicking portfolios, economic risk premia and tests of multi-beta models. *Journal of Business & Economic Statistics* **26**, 354–368.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2008a) Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* **76**, 1481–1536.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2008b) Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Preprint.

Barndorff-Nielsen, O.E. and Shephard, N. (2004) Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* **72**, 885–925.

Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2006) Multivariate GARCH models: A survey. *Journal of Applied Econometrics* **21**, 79–109.

Berkes, I., Gombay, E., Horváth, L. and Kokoszka, P. (2004) Sequential change-point detection in GARCH(p, q) models. *Econometric Theory* **20**, 1140–1167.

Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

Bollerslev, T. (1990) Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics* **74**, 498–505.

Bosq, D. (2000) *Linear Processes in Function Spaces: Theory and Applications.* Springer-Verlag, New York.

Černiková, A., Hušková, M., Prášková, Z. and Steinebach, J. (2011) Delay time in monitoring jump changes in linear models. *Statistics*, iFirst, 1–25.

Chu, C.-S. J., Stinchcombe, M. and White, H. (1996) Monitoring structural change. *Econometrica* **64**, 1045–1065.

Engle, R.F. (2000) The econometrics of ultra-high-frequency data. *Econometrica* **68**, 1–22.

Engle, R.F. (2002) Dynamic conditional correlation—A simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* **20**, 339–350.

Engle, R.F. and Kroner, K.F. (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**, 122–150.

Engle, R.F., Ng, V., and Rothschild, M. (1990) Asset pricing with a factor-ARCH structure: Empirical estimates for Treasury bills. *Journal of Econometrics* **45**, 213–237.

Garcia, R. and Ghysels, E. (1998) Structural change and asset pricing in emerging markets. *Journal of International Money and Finance* **17**, 455–473.

Ghysels, E. (1998) On stable factor structures in the pricing of risk: do time-varying betas help or hurt? *Journal of Finance* **53**, 549–573.

Goodhart, C.A.E. and O'Hara, M. (1997) High frequency data in financial markets: issues and applications. *Journal of Empirical Finance* **4**, 73–114.

Hafner, C.M. and Preminger, A. (2009) Asymptotic theory for a factor GARCH model. *Econometric Theory* **25**, 336–363.

Hafner, C.M. and Preminger, A. (2009+) On asymptotic theory for multivariate GARCH models. *Journal of Multivariate Analysis*, forthcoming.

Harrison, J.M., Pitbladdo, R. and Schaefer, S.M. (1984) Continuous prices in frictionless markets have infinite variation. *Journal of Business* **57**, 353–365.

Harvey, C.R. (1991) The world price of covariance risk. *Journal of Finance* **46**, 111–157.

Hörmann, S., Horváth, L. and Reeder, R. (2010+) Functional volatility sequences. Preprint, University of Utah, Salt Lake City, Utah, USA.

Huang, H.C. and Cheng, W.H. (2005) Tests of the CAPM under structural changes. *International Economic Journal* **19**, 523–541.

Hušková, M., Prašková, Z. and Steinebach, J. (2007) On the detection of changes in autoregressive time series, I. Asymptotics. *Journal of Statistical Planning and Inference* **137**, 1243–1259.

Ibragimov, I.A. (1962) Some limit theorems for stationary processes. *Theory of Probability and Its Applications* **7**, 349–382.

Jacod, J., Li, Y., Mykland, P.A., Podolskij, M. and Vetter, M. (2009) Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications* **119**, 2249–2276.

Jansson, M. (2002) Consistent covariance matrix estimation for linear processes. *Econometric Theory* **18**, 1449–1459.

Jeantheau, T. (1998) Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* **14**, 70–86.

Lintner, J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* **47**, 13–37.

Liu, W. and Wu, W.B. (2010) Asymptotics of spectral density estimates. *Econometric Theory* **26**, 1218–1245.

MacKinley, A.C. (1987) On multivariate tests of the CAPM. *Journal of Financial Economics* **18** 341–371.

Markowitz, H.M. (1999) The early history of portfolio theory: 1600–1960. *Financial Analysts Journal* **55**, 5–16.

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

Móricz, F. (1977) Moment inequalities for the maximum of partial sums of random fields. *Acta Scientiarum Mathematicarum (Szeged)* **39**, 353–366.

Newey, W.K. and West K.D. (1987) A simple, positive-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**, 703–708.

Philipps, P.C.B. (2005) HAC estimation by automated regression. *Econometric Theory* **21**, 116–142.

Santos, A.O. (1998) Capital asset pricing model and changes in volatility. International Center for Financial Asset Management and Engineering, Research Paper No. 4.

Shao, X. and Wu, W.B. (2004) Limit theorems for iterated random functions. *Journal of Applied Probability* **41**, 425–436.

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

Smith, R.J. (2005) Automatic positive semidefinite HAC covariance matrix and GMM estimation. *Econometric Theory* **21**, 158–170.

Tsay, R.S. (2002) *Analysis of Financial Time Series.* Wiley, New York.

Wang, J., Meric, G., Liu, Z. and Meric, I. (2009) Stock market crashes, firm characteristics, and stock returns. *Journal of Banking & Finance* **33**, 1563–1574.

Wu, W.B. (2005) Nonlinear systems theory: Another look at dependence. *Proceedings of the National Academy of Sciences U.S.A.* **102**, 14150–14154.

Wu, W.B. (2007) Strong invariance principles for dependent random variables. *The Annals of Probability* **35**, 2294–2320.

Zeileis, A. (2004) Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software* **11**, 1–17.

Zhang, L., Mykland, P.A. and Aït-Sahalia, Y. (2005) A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* **100**, 1394–1411.