UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2014/15

Cambridge Working Paper in Economics: 1464

ESTIMATING THE SPOT COVARIATION OF ASSET PRICES – STATISTICAL THEORY AND EMPIRICAL EVIDENCE

Markus Bibinger (Humboldt-Universität)
Nikolaus Hautsch (University of Vienna and Center for Financial Studies)
Peter Malec (University of Cambridge)
Markus Reiss (Humboldt-Universität)

ABSTRACT

We propose a new estimator for the spot covariance matrix of a multi-dimensional continuous semi-martingale log asset price process which is subject to noise and non-synchronous observations. The estimator is constructed based on a local average of block-wise parametric spectral covariance estimates. The latter originate from a local method of moments (LMM) which recently has been introduced by Bibinger et al. (2014). We extend the LMM estimator to allow for autocorrelated noise and propose a method to adaptively infer the autocorrelations from the data. We prove the consistency and asymptotic normality of the proposed spot covariance estimator. Based on extensive simulations we provide empirical guidance on the optimal implementation of the estimator and apply it to high-frequency data of a cross-section of NASDAQ blue chip stocks. Employing the estimator to estimate spot covariances, correlations and betas in normal but also extreme-event periods yields novel insights into intraday covariance and correlation dynamics. We show that intraday (co-)variations (i) follow underlying periodicity patterns, (ii) reveal substantial intraday variability associated with (co-)variation risk, (iii) are strongly serially correlated, and (iv) can increase strongly and nearly instantaneously if new information arrives.

Estimating the Spot Covariation of Asset Prices – Statistical Theory and Empirical Evidence*

Markus Bibinger† Nikolaus Hautsch‡ Peter Malec§ Markus Reiss¶

October 2014

Abstract

We propose a new estimator for the spot covariance matrix of a multi-dimensional continuous semi-martingale log asset price process which is subject to noise and non-synchronous observations. The estimator is constructed based on a local average of block-wise parametric spectral covariance estimates. The latter originate from a local method of moments (LMM) which recently has been introduced by Bibinger et al. (2014). We extend the LMM estimator to allow for autocorrelated noise and propose a method to adaptively infer the autocorrelations from the data. We prove the consistency and asymptotic normality of the

*For helpful comments and discussions we thank the participants of the Financial Statistics Conference, Stevanovich Center, University of Chicago, September 2014, the 7th Annual Meeting of the Society for Financial Econometrics, University of Toronto, 2014, the Workshop "Measuring and Modeling Financial Risk with High Frequency Data", European University Institute, Florence, June 2013, the conference "Statistics for Stochastic Processes and Analysis of High Frequency Data", University Pierre and Marie Curie, Paris, December, 2013, the seminars at the University of Pennsylvania, Philadelphia, Institute for Advanced Studies, Vienna, University of St. Gallen and Leibniz University Hannover. Financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Economic Risk" and FOR 1735 "Structural Inference in Statistics: Adaptation and Efficiency" is gratefully acknowledged. Hautsch also acknowledges financial support from the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF). Malec thanks the Cambridge INET for financial support.

†Department of Mathematics, Humboldt-Universität zu Berlin. Email: bibinger@math.hu-berlin.de. Address: Unter den Linden 6, D-10099 Berlin, Germany.

‡Faculty of Business, Economics and Statistics, University of Vienna and Center for Financial Studies, Frankfurt. Email: nikolaus.hautsch@univie.ac.at. Address: Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria.

§Faculty of Economics, University of Cambridge. Email: pm563@cam.ac.uk. Address: Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom.

¶Department of Mathematics, Humboldt-Universität zu Berlin. Email: mreiss@math.hu-berlin.de. Address: Unter den Linden 6, D-10099 Berlin, Germany.

proposed spot covariance estimator. Based on extensive simulations we provide empirical guidance on the optimal implementation of the estimator and apply it to high-frequency data of a cross-section of NASDAQ blue chip stocks. Employing the estimator to estimate spot covariances, correlations and betas in normal but also extreme-event periods yields novel insights into intraday covariance and correlation dynamics. We show that intraday (co-)variations (i) follow underlying periodicity patterns, (ii) reveal substantial intraday variability associated with (co-)variation risk, (iii) are strongly serially correlated, and (iv) can increase strongly and nearly instantaneously if new information arrives.

Keywords: local method of moments, spot covariance, smoothing, intraday (co-)variation risk

JEL classification: C58, C14, C32

1 Introduction

Recent literature in financial econometrics and empirical finance reports strong empirical evidence for distinct time variations in daily and long-term correlations between asset prices. Clearly rejecting the constancy of covariances over time underlines the importance of suitable econometric models to capture covariance dynamics and challenges risk management, portfolio management and asset pricing. Surprisingly little, however, is known about *intraday* variations of asset return covariances. While the literature proposes several approaches to estimate spot *variances*[1], there is a lack of empirical approaches and corresponding statistical theory to estimate spot *covariances* using high-frequency data.

In this paper, we aim at filling this gap in the literature and propose an estimator for the spot covariance matrix of a multi-dimensional continuous semi-martingale log asset price process which is observed at non-synchronous times under noise. The estimator is constructed based on local averages of block-wise parametric spectral covariance estimates. The latter are estimated employing the local method of moments (LMM) estimator proposed by Bibinger et al. (2014), which is shown to be a rate-optimal and asymptotically efficient estimator for the integrated covariation. As the LMM estimator builds on locally constant approximations of the underlying covariance process and estimates them block-wise, it provides a natural setting to construct a spot covariance estimator.

Our methodological contribution is as follows: First, we extend the LMM method to allow for autocorrelated market microstructure noise and propose consistent estimators of the autocorrelations. Second, we derive a stable central limit theorem, showing the consistency and asymptotic normality of the resulting spot covariance estimator. Apart from being (to our

[1] See, e.g., Kristensen (2010), Mykland and Zhang (2008), Mancini et al. (2012), Bos et al. (2012) or Zu and Boswijk (2014).

best knowledge) the first estimator for a spot *covariance matrix*, an important result is that the rate-optimality of the underlying LMM estimator carries over to the spot estimator. It is shown that it converges considerably faster than existing spot estimators. We provide simulation-based evidence on an optimal implementation of the estimator depending on the choice of underlying smoothing parameters. Finally, an important objective of this paper is to provide first empirical evidence on the intraday behavior of spot covariances, correlations and asset price betas.

The quantification of spot covariance estimators is useful for intraday risk management, market microstructure research, but also for market surveillance and monitoring. For instance, with the availability of estimates of (co-)volatilities in the market, intraday traders can assess intraday correlation risks. Empirical studies on the role of high-frequency trading, the impact of market fragmentation and the usefulness of volatility circuit breakers might heavily benefit from the availability of high-frequency covariance estimators which are applicable in higher dimensions. Moreover, analyzing the behavior of spot (co-)volatilities on days with distinct information arrival or (flash) crashes provides important insights into high-frequency dependence structures and their reaction to common shocks. For example, it is an empirically well-documented fact that correlations are higher and diversification opportunities are smaller during bear markets than during bull markets at a lower frequency level (see, e.g., De Santis and Gerard, 1997; Longin and Solnik, 2001). This raises the question if such changing correlation structures might be observed also on an intraday level, e.g., during "flash crash" periods. Finally, spot covariance estimates are a necessary building block for co-jump tests (see Bibinger and Winkelmann, 2013).

Our study is mainly related to two fields of literature. First, there is a vast body of papers on the estimation of integrated covariance matrices, while accounting for market microstructure noise and the asyncronicity of observations. Starting from the seminal realized covariance estimator by Barndorff-Nielsen and Shephard (2004) which neglects both types of frictions, Hayashi and Yoshida (2011) propose a consistent and efficient estimator under asynchronicity, but in the absence of microstructure noise. Estimators accounting for both types of frictions are, among others, the quasi maximum likelihood estimator by Ait-Sahalia et al. (2010), the multivariate realized kernel estimator by Barndorff-Nielsen et al. (2011), the multivariate pre-averaging estimator by Christensen et al. (2013), the two-scale estimator by Zhang (2011), and the LMM estimator by Bibinger et al. (2014). Second, there is considerable literature on spot volatility estimation. A nonparametric (kernel-type) estimator in the absence of microstructure noise is put forward by Foster and Nelson (1996), Fan and Wang (2008) and Kristensen (2010). To account for noise, the predominant approach is to compute a difference quotient based on a noise-robust integrated volatility estimator, e.g., the (univariate) realized kernel, the pre-averaging estimator or the two-scale estimator. Here, examples include Mykland and Zhang

(2008), Mancini et al. (2012), Bos et al. (2012) and Zu and Boswijk (2014). Finally, estimators that are robust to jumps, but neglect microstructure noise are put forward, e.g., by Ait-Sahalia and Jacod (2009), Andersen et al. (2009) and Bandi and Reno (2009).

Interestingly, the problem of estimating the spot *covariance matrix* in the presence of microstructure noise and asynchronicity effects has not yet been addressed in a study on its own. Our paper thus bridges the gap between the two fields of literature outlined above. Compared to integrated (co-)variance estimators, spot estimators inherently feature slower convergence rates due to the additional smoothing involved. Hence, to maximize precision, as many observations as possible should be used. The latter can be achieved by employing data sampled at the highest frequency possible, i.e., at tick-by-tick level. Hansen and Lunde (2006) and, more recently, Ait-Sahalia et al. (2011), however, show that at this frequency, microstructure noise appears to violate the traditional i.i.d. assumption, exhibiting more complex dependence structures. Extending the LMM estimator to allow for serially dependent noise and proposing a data-driven procedure to select the order of serial dependence is therefore a crucial step to make the estimator applicable to spot covariance estimation and to benefit from its optimality properties. In particular, we show that it satisfies a central limit theorem at almost optimal rate.

The approach presented here does not account for jumps in the log-price process. From a methodological point of view, an extension to disentangle jumps and continuous components utilizing a truncation technique as in Bibinger and Winkelmann (2013) appears feasible. In the given framework, however, due to additional tuning parameters involved this would require a comprehensive extension, which would dilute the main new estimation ideas. Consequently, our proposed spot covariance estimator does not separate between a diffusive and jump component. For our empirical results and corresponding conclusions, this is not a limitation since in any case, potential jumps are consistently captured by the spot estimator. Moreover, Christensen et al. (2014) show that, when considering data sampled at the tick-by-tick level, jumps are detected far less frequently than based on a coarser sampling grid.

Based on extensive simulation studies, we investigate the impact of different choices of the relevant input parameters on the estimator's finite sample precision and provide guidance of how to choose these parameters in applications. Moreover, we demonstrate that the proposed procedure for estimating the order of serial dependence in the microstructure noise process performs well under realistic conditions.

Applying the spot covariance estimator to four years of trade and quote data for 30 of the most liquid constituents of the NASDAQ100 and an ETF tracking the latter, provide novel empirical evidence on the intraday behavior of covariances and correlations. First, there is a distinct intraday seasonality pattern with covariances declining and correlations increasing throughout the day, while betas remain relatively stable. Second, spot (co-)variation reveals substantial

intraday variability and thus reflect (co-)variation risk. Third, spot (co-)variation is strongly serially correlated within a day and across days. Finally, spot covariances and correlations substantially change during flash crashes or the arrival of fundamental information. We show that in these extreme scenarios, spot covariances and variances of the 30 blue chips strongly and nearly instantaneously shoot up, causing the resulting correlations either to substantially increase or decrease. These novel insights show that dependence structures between assets can be very variable and sensitive to news. It turns out that our estimator is able to capture extreme (co-)variance movements on a high time resolution.

The remainder of the paper is structured as follows. Section 2 introduces the proposed spot estimator, gives its asymptotic properties and shows how to estimate underlying noise autocovariances. In Section 3, we present a simulation study analyzing the estimator's sensitivity to the choice of input parameters and examining the performance of the procedure for estimating the autocovariance structure of the noise process. Section 4 provides empirical evidence on spot (co-)variances, correlations and betas based on NASDAQ data. Finally, Section 5 concludes.

2 Estimation of Spot Covariances

2.1 Theoretical Setup and Assumptions

Let $(X_t)_{t\geq 0}$ denote the d-dimensional efficient log-price process. In line with the literature and motivated by well-known no-arbitrage arguments, we assume that X_t follows a continuous Itô semi-martingale

$$X_t = X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s \, dB_s, \ t \in [0,1], \tag{1}$$

defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, \mathbb{P}\right)$ with drift b_s, d-dimensional standard Brownian motion B_s and instantaneous volatility matrix σ_s. The latter yields the $(d \times d)$-dimensional spot covariance matrix $\Sigma_s = \sigma_s \sigma_s^\top$, which is our object of interest. We consider a setting in which discrete and non-synchronous observations of the process (1) are diluted by market microstructure noise, i.e.,

$$Y_i^{(p)} = X_{t_i^{(p)}}^{(p)} + \epsilon_i^{(p)}, \ i = 0, \ldots, n_p, \ p = 1, \ldots, d, \tag{2}$$

with observation times $t_i^{(p)}$, and observation errors $\epsilon_i^{(p)}$. Observed returns for component $p \in \{1, \dots, d\}$ are given by

$$\begin{aligned} \Delta_i Y^{(p)} &= Y_i^{(p)} - Y_{i-1}^{(p)} = \Delta_i X^{(p)} + \Delta_i \epsilon^{(p)} \\ &= X_{t_i^{(p)}}^{(p)} - X_{t_{i-1}^{(p)}}^{(p)} + \epsilon_i^{(p)} - \epsilon_{i-1}^{(p)}, \quad i = 1, \dots, n_p. \end{aligned} \tag{3}$$

Let $n = \min_{1 \le p \le d} n_p$ denote the number of observations of the "slowest" asset. In Section 2.3, we consider high-frequency asymptotics with $n/n_p \to \nu_p$ for constants $0 < \nu_p \le 1$, such that the asymptotic variance-covariance matrices for estimators of Σ_s are regular. Feasible inference on Σ based on the considered methods, however, is tractable even in a broader framework with different speeds in increasing sample sizes. For this technical generalization, we refer to Bibinger et al. (2014).

Below we summarize the assumptions on the instantaneous volatility matrix and drift process, noise properties and observation times. Convergence rates of spot estimators crucially depend on the smoothness of the underlying functions. We therefore consider balls in Hölder spaces of order $\alpha \in (0, 1]$ and with radius $R > 0$:

$$C^{\alpha,R}([0,1], \mathcal{E}) = \{f : [0,1] \to \mathcal{E} | \|f\|_\alpha \le R\}, \ \|f\|_\alpha := \|f\|_\infty + \sup_{x \ne y} \frac{\|f(x) - f(y)\|}{|x - y|^\alpha},$$

where $\|\cdot\|$ denotes the usual spectral norm and $\|f\|_\infty := \sup_{t \in [0,1]} \|f(t)\|$ for functions on $[0, 1]$. In our setup, we have $\mathcal{E} = \mathbb{R}^{d \times d'}$ for matrix-valued functions, $\mathcal{E} = \mathbb{R}^d$ for vectors or $\mathcal{E} = [0, 1]$ for distribution functions.

First, for the drift process in (1), we only assume a very mild regularity:

Assumption 1. *$(b_s)_{s \in [0,1]}$ is an $(\mathcal{F}_s)$-adapted process with $b_s \in C^{\nu,R}([0,1], \mathbb{R}^d)$ for some $R < \infty$ and some $\nu > 0$.*

Furthermore, the assumptions regarding the instantaneous volatility matrix process in (1) can be summarized as:

Assumption 2. *(i) $(\sigma_s)_{s \in [0,1]}$ follows an $(\mathcal{F}_s)$-adapted process satisfying $\Sigma_s = \sigma_s \sigma_s^\top \ge \underline{\Sigma}$ uniformly for some strictly positive definite matrix $\underline{\Sigma}$.*

(ii) $(\sigma_s)_{s \in [0,1]}$ satisfies $\sigma_s = f\big(\sigma_s^{(1)}, \sigma_s^{(2)}\big)$ with $f : \mathbb{R}^{2d \times 2d'} \to \mathbb{R}^{d \times d'}$ being a continuously differentiable function in all coordinates, where

- *For $\alpha \in (0, 1/2]$, $\sigma_s^{(1)}$ is a continuous Itô semi-martingale of the form (1) with càdlàg adapted volatility of volatility $\tilde{\sigma}_s$ satisfying $\tilde{\Sigma}_s = \tilde{\sigma}_s \tilde{\sigma}_s^\top \ge \underline{\tilde{\Sigma}}$ uniformly for some strictly positive definite matrix $\underline{\tilde{\Sigma}}$. The drift of $\sigma^{(1)}$ is an adapted càdlàg process.*
  For $\alpha \in (1/2, 1]$, $\sigma^{(1)}$ vanishes.

- $\sigma_s^{(2)} \in C^{\alpha,R}\big([0,1], \mathbb{R}^{d\times d'}\big)$ *with some* $R < \infty$.

Hence, σ_s is a function of a continuous Itô semi-martingale $\sigma_s^{(1)}$ and an additional component $\sigma_s^{(2)}$. The latter can capture intraday periodicity effects (see, e.g., Andersen and Bollerslev, 1997). The larger α, the more restrictive becomes Assumption 2. If $\alpha > 1/2$, the semi-martingale component vanishes and σ_s is exclusively driven by the component $\sigma_s^{(2)}$. Hence, the more interesting case is $\alpha \leq 1/2$. Importantly, the above assumptions also allow for leverage effects, i.e., a non-zero correlation between σ_s and the Brownian motion B_s in (1). It is natural to develop results under this general smoothness assumption as it is commonly known that in nonparametric estimation problems, the underlying regularity determines the size of smoothing windows and a fortiori the resulting (optimal) convergence rates.

Our assumptions on the microstructure noise process in (2) are stated in observation time, which is in line with, e.g., Hansen and Lunde (2006) and Barndorff-Nielsen et al. (2011):

Assumption 3. *(i)* $\epsilon = \{\epsilon_i^{(p)}, i = 0, \ldots, n_p, p = 1, \ldots, d\}$ *is independent of* X *and has independent components, i.e.,* $\epsilon_i^{(p)}$ *is independent of* $\epsilon_j^{(q)}$ *for all* i, j *and* $p \neq q$.

(ii) At least the first eight moments of $\epsilon_i^{(p)}, i = 0, \ldots, n_p$, *exist for each* $p = 1, \ldots, d$.

(iii) $\epsilon_i^{(p)}, i = 0, \ldots, n_p$, *follows an* R*-dependent process for some* $R < \infty$, *implying that* $\mathbb{C}\mathrm{ov}\big(\epsilon_i^{(p)}, \epsilon_{i+u}^{(p)}\big) = 0$ *for* $u > R$ *and each* $p = 1, \ldots, d$. *Define by*

$$\eta_p = \eta_0^{(p)} + 2\sum_{u=1}^{R} \eta_u^{(p)}, \quad \textit{with} \quad \eta_u^{(p)} := \mathbb{C}\mathrm{ov}\big(\epsilon_i^{(p)}, \epsilon_{i+u}^{(p)}\big), u \leq R, \tag{4}$$

the component-wise long-run noise variances, where the $\eta_u^{(p)}, 0 \leq u \leq R$, *are constant for all* $0 \leq i \leq n - u$. *We impose that* $\eta_p > 0$ *for all* p.

The independence between noise and the efficient price, as stated in part (i) of Assumption 3, is standard in the literature (see, e.g., Zhang et al., 2005).[2] Considering serially dependent noise is non-standard and motivated by empirical results, e.g., in Hansen and Lunde (2006). The moving-average-type dependence structure in the noise process in part (iii) of Assumption 3 follows, e.g., Hautsch and Podolskij (2013), implying the long-run variance (4).

Finally, we assume that the timing of observations in (2) is driven by c.d.f.'s F_p governing the transformations of observation times to equidistant sampling schemes by means of suitable quantile transformations:

Assumption 4. *There exist differentiable cumulative distribution functions* $F_p, p = 1, \ldots, d$, *such that the observation regimes satisfy* $t_i^{(p)} = F_p^{-1}(i/n_p)\,, 0 \leq i \leq n_p, p \in \{1, \ldots, d\}$, *where*

[2]For the case of endogenous noise, see, e.g., Kalnina and Linton (2008).

$F_p' \in C^{\alpha,R}\big([0,1],[0,1]\big), p = 1,\dots,d$, *with* α *being the smoothness exponent in Assumption 2 for some* $R < \infty$.

Assumption 4 implies that observation times are either non-stochastic or random, but independent from the log-price process. A treatment of endogenous times in the given theoretical framework is beyond the scope of this paper. See Koike (2014) for a recent study of endogenous times and Li et al. (2014) for a study in a setting neglecting microstructure noise.

Combining time-invariant (long-run) noise variances η_p and locally different observation frequencies from Assumptions 3 and 4 implies locally varying noise levels:

Definition 1. *In the asymptotic framework with* $n/n_p \to \nu_p$, *where* $0 < \nu_p < \infty, p = 1,\dots,d$, *for* $n \to \infty$, *we define the continuous-time noise level matrix*

$$H_s = \operatorname{diag}\left((\eta_p \nu_p (F_p^{-1})'(s))^{1/2}\right)_{1\le p\le d}. \tag{5}$$

Note that for equally-spaced observations, we have $F_p(s) = s$, such that $(F_p^{-1})'(s) = 1$. Then, the p-specific (asymptotic) noise level is $(\eta_p\nu_p)^{1/2}$ with ν_p expressing the inverse of the sample size of the p-th process relative to the "slowest" process. Hence, having less frequent observations on a sub-interval is equivalent to having higher noise dilution by microstructure effects on this sub-interval. This interplay between noise and liquidity has been discussed by Bibinger et al. (2014).

2.2 Local Method of Moments Estimation of the Spot Covariance Matrix

Our approach for estimating the instantaneous covariance matrix rests upon the concept of the local method of moments (LMM) introduced in Bibinger et al. (2014). We partition the interval $[0,1]$ into equidistant blocks $[kh_n,(k+1)h_n], k = 0,\dots,h_n^{-1}-1$, with the block length h_n asymptotically shrinking to zero, $h_n \to 0$ as $n \to \infty$. The key idea is to approximate the underlying process (1) in model (2) by a process with block-wise constant covariance matrices and noise levels. In the (more simplified) setting of Bibinger et al. (2014), it is shown that such a locally constant approximation induces an estimation error for the integrated covariation, which, however, can be asymptotically neglected for sufficient smoothness of Σ_t and F_p if the block sizes h_n shrink sufficiently fast with increasing n. This opens the path to construct an asymptotically efficient estimator of the integrated covariation matrix based on optimal block-wise estimates.

In the present setting, we build on the idea of block-wise constant approximations of the underlying covariance and noise process and show that it allows constructing a consistent spot covariance estimator, which can attain an optimal rate. A major building block is the

construction of an unbiased estimator of the block-wise covariance matrix $\Sigma_{kh_n} = \sigma_{kh_n}\sigma_{kh_n}^\top$ based on the local spectral statistics

$$S_{jk} = \pi j h_n^{-1} \left(\sum_{i=1}^{n_p} \left(Y_i^{(p)} - Y_{i-1}^{(p)} \right) \Phi_{jk} \left(\frac{t_{i-1}^{(p)} + t_i^{(p)}}{2} \right) \right)_{1 \le p \le d}, \tag{6}$$

where Φ_{jk} denote orthogonal sine functions with (spectral) frequency j, whose derivatives Φ'_{jk} form another orthogonal system corresponding to the eigenfunctions of the covariance operator of a Brownian motion, and are given by

$$\Phi_{jk}(t) = \frac{\sqrt{2h_n}}{j\pi} \sin\left(j\pi h_n^{-1} \left(t - kh_n \right) \right) \mathbb{1}_{[kh_n,(k+1)h_n)}(t), j \ge 1\,. \tag{7}$$

The statistics (6) de-correlate the noisy observations (3) and can be thought of as representing their block-wise principal components.[3] They bear some resemblance to the pre-averaged returns as employed in Jacod et al. (2009). While pre-averaging estimators, however, utilize rolling (local) windows around each observation, our approach relies on fixed blocks and optimal combinations in the spectral frequency domain. It can be shown that

$$\mathbb{C}\mathrm{ov}(S_{jk}) = (\Sigma_{kh_n} + \pi^2 j^2 h_n^{-2} \mathbf{H}_k^n)(1 + o(1)), \tag{8}$$

where $\mathbf{H}_k^n$ denotes the block-wise constant diagonal noise level matrix with entries

$$\left(\mathbf{H}_k^n \right)^{(pp)} = n_p^{-1} \eta_p (F_p^{-1})'(kh_n)\,, \tag{9}$$

and $(F_p^{-1})'(kh_n)$ denotes the inverse of the block-wise constantly approximated observation frequency evaluated at kh_n. The relation (8) suggests estimating Σ_{kh_n} based on the empirical covariance $S_{jk}S_{jk}^\top$, which is bias-corrected by the noise-induced term $\pi^2 j^2 h_n^{-2} \mathbf{H}_k^n$.

An initial (pre-) estimator of the spot covariance matrix at time $s \in [0,1]$, Σ_s, is then constructed based on bias-corrected block-wise empirical covariances $S_{jk}S_{jk}^\top$, which are averaged across spectral frequencies $j = 1, \ldots, J_n^p$, and a set of adjacent blocks,

$$\mathrm{vec}\left(\hat{\Sigma}_{kh_n}^{pre} \right) = (U_{s,n} - L_{s,n} + 1)^{-1} \sum_{k=L_{s,n}}^{U_{s,n}} (J_n^p)^{-1} \sum_{j=1}^{J_n^p} \mathrm{vec}\left(S_{jk}S_{jk}^\top - \pi^2 j^2 h_n^{-2} \hat{\mathbf{H}}_k^n \right), \tag{10}$$

with the vec operator stacking the elements of a $(d \times d)$-matrix into a $\left(d^2 \times 1\right)$-vector column-wise, while $L_{s,n} = \max\{\lfloor s h_n^{-1} \rfloor - K_n, 0\}$ and $U_{s,n} = \min\{\lfloor s h_n^{-1} \rfloor + K_n, \lceil h_n^{-1} \rceil - 1\}$, such that

[3] Somewhat related approaches for a univariate framework can be found in Hansen et al. (2008) and Curci and Corsi (2012).

the length of the smoothing window obeys $U_{s,n} - L_{s,n} + 1 \leq 2K_n + 1$. $\hat{\mathbf{H}}_k^n$ is a $\sqrt{n}$-consistent estimator of $\mathbf{H}_k^n$ with p-th diagonal element

$$\big(\hat{\mathbf{H}}_k^n\big)^{(pp)} = \frac{\hat{\eta}_p}{h_n} \sum_{kh_n \leq t_i^{(p)} \leq (k+1)h_n} \big(t_i^{(p)} - t_{i-1}^{(p)}\big)^2. \tag{11}$$

Details on the construction of the estimator of the component-wise long-run noise variances, $\hat{\eta}_p$, are provided in Section 2.5 below.

For each spectral frequency j, the statistic $S_{jk}S_{jk}^\top - \pi^2 j^2 h_n^{-2}\hat{\mathbf{H}}_k^n$ is an (asymptotically) unbiased though inefficient estimator of Σ_{kh_n}. Averaging across different frequencies therefore increases the estimator's efficiency. Equally weighting as in (10), however, is not necessarily optimal. A more efficient estimator can be devised by considering (10) as the pre-estimated spot covariance matrix and then, derive estimated optimal weight matrices $\hat{W}_j$, yielding the final LMM spot covariance matrix estimator as

$$\begin{aligned}\operatorname{vec}\big(\hat{\Sigma}_s\big) = (U_{s,n} - L_{s,n} + 1)^{-1} \sum_{k=L_{s,n}}^{U_{s,n}} \sum_{j=1}^{J_n} \hat{W}_j\big(\hat{\mathbf{H}}_k^n, \hat{\Sigma}_{kh_n}^{pre}\big) \\ \times \operatorname{vec}\Big(S_{jk}S_{jk}^\top - \pi^2 j^2 h_n^{-2}\hat{\mathbf{H}}_k^n\Big).\end{aligned} \tag{12}$$

As outlined in detail in Bibinger et al. (2014), the true optimal weights are given proportionally to the local Fisher information matrices according to

$$\begin{aligned}W_j\big(\mathbf{H}_k^n, \Sigma_{kh_n}\big) &= \Big(\sum_{u=1}^{J_n} \big(\Sigma_{kh_n} + \pi^2 u^2 h_n^{-2}\mathbf{H}_k^n\big)^{-\otimes 2}\Big)^{-1} \big(\Sigma_{kh_n} + \pi^2 j^2 h_n^{-2}\mathbf{H}_k^n\big)^{-\otimes 2} \\ &= I_k^{-1} I_{jk},\end{aligned} \tag{13}$$

with I_{jk} being the Fisher information matrix associated with block k and spectral frequency j, given by

$$I_{jk} = \big(\Sigma_{kh_n} + \pi^2 j^2 h_n^{-2}\mathbf{H}_k^n\big)^{-\otimes 2}, \tag{14}$$

and $I_k = \sum_{j=1}^{J_n} I_{jk}$ denoting the k-specific Fisher information (exploiting the independence across frequencies j). Here, $A^{\otimes 2} = A \otimes A$ denotes the Kronecker product of a matrix with itself and $A^{-\otimes 2} = A^{-1} \otimes A^{-1} = (A \otimes A)^{-1}$. We show in Section 2.3 that the estimator (12), which builds on the idealized model considered in Bibinger et al. (2014), is consistent and satisfies a stable CLT under the more realistic and general assumptions of Section 2.1.

For practical purposes, note that, while both the pilot estimator (10) and the LMM estimator (12) are symmetric, none is guaranteed to yield positive semi-definite estimates. Confidence is based on estimated Fisher information matrices, which are by construction positive-definite. For the estimates themselves, we can set negative eigenvalues equal to zero, which is tantamount to a projection on the space of positive semi-definite matrices. This adjustment does not affect the asymptotic properties of the estimator as outlined below, but only improves the finite sample performance.

2.3 Asymptotic Properties

As a prerequisite for the discussion of the central limit theorem for the estimator (12), some considerations regarding K_n, which determines the length of the smoothing window, are needed. For this purpose, suppose that a certain smoothness $\alpha \in (0, 1]$ of the instantaneous volatility matrix is granted according to Assumption 2. Then, a simple computation yields $\|\mathbb{C}\mathrm{OV}(\hat{\Sigma}_s)\| = \mathcal{O}(K_n^{-1})$, implying a bias-variance trade-off in the mean square error $\mathrm{MSE}(\hat{\Sigma}_s) := \mathbb{E}\left[(\hat{\Sigma}_s - \Sigma_s)^2\right]$. More precisely, for a specific $\alpha > 0$, we have

$$\mathrm{MSE}(\hat{\Sigma}_s) = \mathcal{O}_P(K_n^{-1}) + \mathcal{O}_P(K_n^{2\alpha} h_n^{2\alpha}), \tag{15}$$

where the first term originates from the variance and the second term is induced by the squared bias. Consequently, for given $h_n \sim \log(n) n^{-1/2}$, which optimally balances noise and discretization error as derived in Bibinger et al. (2014), choosing $K_n \sim n^{\alpha/(2\alpha+1)}$ minimizes the MSE and facilitates an estimator with $\sqrt{K_n}$ convergence rate. Finally, the desired central limit theorem for the estimator (12) requires a slight undersmoothing, resulting in a smaller choice of K_n:

Theorem 1. *We assume a setup with observations of the type (2), a signal (1) and the validity of Assumptions 1-4. Then, for $h_n = \kappa_1 \log(n) n^{-1/2}$, $K_n = \kappa_2 n^{\beta} (\log(n))^{-1}$ with constants κ_1, κ_2 and $0 < \beta < \alpha(2\alpha+1)^{-1}$, for $J_n \to \infty$ and $n/n_p \to \nu_p$ with $0 < \nu_p < \infty, p = 1, \ldots, d$, as $n \to \infty$, the spot covariance matrix estimator (12) satisfies the pointwise stable central limit theorem:*

$$n^{\beta/2} \operatorname{vec}(\hat{\Sigma}_s - \Sigma_s) \xrightarrow{d-(st)} \mathbf{N}\Big(0, 2\big(\Sigma \otimes \Sigma_H^{1/2} + \Sigma_H^{1/2} \otimes \Sigma\big)_s \mathcal{Z}\Big), \quad s \in [0, 1], \tag{16}$$

where $\Sigma_H = H\big(H^{-1}\Sigma H^{-1}\big)^{1/2} H$, with noise level H from (5) and $\mathcal{Z} = \mathbb{C}\mathrm{OV}(\mathrm{vec}(ZZ^{\top}))$ for $Z \sim \mathbf{N}(0, E_d)$ being a standard normally distributed random vector.

Theorem 1 is proved in Appendix A below. The convergence in (16) is stable, which is equivalent to joint weak convergence with any measurable bounded random variable defined on

the same probability space as X. This allows for a feasible version of the limit theorem if we re-scale the estimator by the implicitly obtained estimated variance:

Corollary 1. *Under the assumptions of Theorem 1, the spot covariance matrix estimator (12) satisfies the feasible central limit theorem given by*

$$(U_{s,n} - L_{s,n} + 1)^{1/2} \big(\hat{\mathbb{V}}_s^n\big)^{-1/2} \operatorname{vec}\left(\hat{\Sigma}_s - \Sigma_s\right) \xrightarrow{d} \mathbf{N}\Big(0, \mathcal{Z}\Big), \quad s \in [0,1]\,, \tag{17a}$$

$$\textit{where} \qquad \hat{\mathbb{V}}_s^n = (U_{s,n} - L_{s,n} + 1)^{-1} \sum_{k=L_{s,n}}^{U_{s,n}} \left(\sum_{j=1}^{J_n} \hat{I}_{jk} \right)^{-1}, \tag{17b}$$

with $U_{s,n}$ and $L_{s,n}$ defined as in (10) and (12). $\hat{I}_{jk}$ is defined according to (14) with $\mathbf{H}_k^n$ and Σ_{kh_n}, $k = 0, \ldots, h_n^{-1} - 1$, replaced by the estimators (10) and (11), respectively.

Unlike in (16), in which we obtain a mixed normal limiting distribution, the matrix $\mathcal{Z}$ is completely known. It is given by twice the "symmetrizer matrix" introduced by Abadir and Magnus (2005, ch. 11) and corresponds to the covariance structure of the empirical covariance of a d-dimensional (standard) Gaussian vector.

The asymptotic variance-covariance matrix in (16) is the same instantaneous process that appears integrated over $[0,1]$ as variance-covariance matrix of the integrated covariance matrix estimator in Bibinger et al. (2014). Accordingly, Theorem 1 is in line with the results on classical realized volatility in the absence of noise for $d = 1$ and the nonparametric Nadaraya-Watson-type kernel estimator by Kristensen (2010) with asymptotic variance $2\sigma_s^4 \int_{\mathbb{R}} k^2(z)\, dz$, where k denotes the used kernel. In our case, the estimator is of histogram-type and the rectangle kernel does not appear in the asymptotic variance. Let us point out that estimator (12), building on optimal combinations over spectral frequencies, is more advanced than a usual histogram-estimator. When comparing our nonparametric estimator (12), e.g., to the aforementioned one by Kristensen (2010), in our case, the actual bandwidth is $(2K_n + 1)h_n$ (or smaller), since we smooth over (up to) $(2K_n + 1)$ adjacent blocks of length h_n. In this context, one can as well think of employing h_n^{-1} de-noised block statistics as underlying observations.

Regarding the convergence rate in (16), we may focus on the case $\alpha \geq 1/2$, which is tantamount to the spot volatility matrix process $(\sigma_s)_{s\in[0,1]}$ being at least as smooth as a continuous semi-martingale. This assumption yields the rate $n^{1/8-\varepsilon}$, for any $\varepsilon > 0$, such that we almost attain the optimal rate $n^{1/8}$, which is obviously lower than the corresponding rate for *integrated* (co-)variance estimators in the setting with noise, $n^{1/4}$. Notably, our spot covariance matrix estimator (12) converges considerably faster than existing noise-robust spot volatility estimators based on the difference quotient of, e.g., pre-averaging estimates (Zu and Boswijk, 2014).

Though the undersmoothed estimator satisfying (16) has a slightly slower convergence rate than the optimal one, the two-step approach (12) with combinations over different frequencies strongly reduces the estimator's variance (compared to simpler methods). This is well confirmed in our finite-sample simulations in Section 3.

Theorem 1 and Corollary 1 hold for estimation points $s \in [0,1]$, both in the interior and in the boundary region of the unit interval. This result is a consequence of the estimators (10) and (12) being of histogram-type, implying that smoothing is conducted by averaging over a set of adjacent blocks, which merely needs to contain time t, and does not have to be centered around the point of estimation. The above property is a major difference compared to kernel-based spot volatility estimators, as e.g., the one proposed by Kristensen (2010), which require suitable correction methods to eliminate the so-called boundary bias, thus making implementation more involved.

Finally, Theorem 1 may be employed to deduce asymptotic results for the estimators of spot correlations and spot betas. These can be considered as the instantaneous counterparts to the integrated quantities studied, e.g., in Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004). In this context, focus on those elements of the spot covariance matrix $\Sigma_t, t \in [0,1]$, involving only the indices $p, q \in \{1, \ldots, d\}$. Further, denote the spot correlation and beta estimators based on (12) by $\hat{\rho}_s^{(pq)} = \hat{\Sigma}_s^{(pq)} / \sqrt{\hat{\Sigma}_s^{(pp)} \hat{\Sigma}_s^{(qq)}}$ and $\hat{\beta}_s^{(pq)} = \hat{\Sigma}_s^{(pq)} / \hat{\Sigma}_s^{(pp)}$. Then, Theorem 1 implies by simple application of the Delta-method that

$$n^{\beta/2} \operatorname{vec}\left(\hat{\rho}_s^{(pq)} - \rho_s^{(pq)}\right) \xrightarrow{d-(st)} \mathbf{N}\Big(0, \mathbb{AV}_{\rho,s}\Big), \quad s \in [0,1], \tag{18a}$$

$$n^{\beta/2} \operatorname{vec}\left(\hat{\beta}_s^{(pq)} - \beta_s^{(pq)}\right) \xrightarrow{d-(st)} \mathbf{N}\Big(0, \mathbb{AV}_{\beta,s}\Big), \quad s \in [0,1], \tag{18b}$$

with

$$\begin{aligned}
\mathbb{AV}_{\rho,s} = {} & \Sigma_s^{(pp)} \Sigma_s^{(qq)} \mathbb{AV}_s^{(p-1)d+q,(p-1)d+q} + \frac{\left(\Sigma_s^{(pq)}\right)^2}{4\left(\Sigma_s^{(pp)}\right)^3 \Sigma_s^{(qq)}} \mathbb{AV}_s^{(p-1)d+p,(p-1)d+p} \\
& + \frac{\left(\Sigma_s^{(pq)}\right)^2}{4\left(\Sigma_s^{(qq)}\right)^3 \Sigma_s^{(pp)}} \mathbb{AV}_s^{(q-1)d+q,(q-1)d+q} - \frac{\Sigma_s^{(pq)}}{\Sigma_s^{(pp)}} \mathbb{AV}_s^{(p-1)d+q,(p-1)d+p} \\
& - \frac{\Sigma_s^{(pq)}}{\Sigma_s^{(qq)}} \mathbb{AV}_s^{(p-1)d+q,(q-1)d+q} + \frac{\left(\Sigma_s^{(pq)}\right)^2}{2\left(\Sigma_s^{(pp)} \Sigma_s^{(qq)}\right)^2} \mathbb{AV}_s^{(p-1)d+p,(q-1)d+q},
\end{aligned} \tag{18c}$$

$$\begin{aligned}
\mathbb{AV}_{\beta,s} = {} & \left(\Sigma_s^{(pp)}\right)^{-2} \mathbb{AV}_s^{(p-1)d+q,(p-1)d+q} + \left(\Sigma_s^{(pq)}\right)^2 \left(\Sigma_s^{(pp)}\right)^{-4} \mathbb{AV}_s^{(p-1)d+p,(p-1)d+p} \\
& - 2\Sigma_s^{(p,q)} \left(\Sigma_s^{(p,p)}\right)^{-3} \mathbb{AV}_s^{(p-1)d+q,(p-1)d+p},
\end{aligned} \tag{18d}$$

where $\mathbb{AV}_s$ denotes the asymptotic variance-covariance matrix in (16). Feasible versions of the central limit theorems (18a) and (18b) can be readily obtained analogously to Corollary 1.

2.4 Choice of Inputs

The proposed spot covariance matrix estimator (12) depends on four input parameters to be chosen: (i) the block length h_n, (ii) the maximum spectral frequency J_n, (iii) the maximum frequency for the pre-estimator (10), J_n^p, as well as (iv) the length of the smoothing window, K_n.

For (i) , Theorem 1 requires that $h_n = \mathcal{O}\big(\log(n)n^{-1/2}\big)$. (ii) is given by $\lfloor \min_p n_p h_n \rfloor$, but a spectral cut-off $J_n = \mathcal{O}\big(\log(n)\big)$ can be chosen, since the weights decay fast with increasing frequency j, making higher frequencies asymptotically negligible. The effect of quickly diminishing weights implies that (iii) should be fixed at a value not "too large", e.g., $J_n^p = 5$. The reason is that the cut-off directly determines the (uniform) weights in the pre-estimator (10). For (iv), we generally set $K_n = \mathcal{O}\big(n^{\alpha/(2\alpha+1)}\big)$. The latter choice implies undersmoothing, thereby forfeiting rate-optimality of the estimator, but provides us a central limit theorem. Under the "continuous semi-martingale or smoother" assumption ($\alpha \geq 1/2$) for the spot volatility matrix process, which seems admissible in most financial applications, we set $K_n = \mathcal{O}\big(n^{1/4-\varepsilon}\big)$ for any $\varepsilon > 0$.

In practice, we introduce proportionality parameters for (i), (ii) and (iv), i.e. $h_n = \theta_h \log(n)n^{-1/2}$, $J_n = \lfloor \theta_J \log(n) \rfloor$ and $K_n = \lceil \theta_K n^{1/4-\delta} \rceil$, where $\theta_h, \theta_J, \theta_K > 0$ and δ denotes a small positive number. This approach is in line with the selection of the window length in the context of pre-averaging estimators (Jacod et al., 2009). We discuss the specific choice of the above input parameters in more detail in Sections 3 and 4.2.

2.5 Estimating Noise Autocovariances

According to part (i) of Assumption 3, the estimation of the long-run noise variance $\eta_p, p = 1, \ldots, d$, defined in (4), only requires estimates of component-wise auto-covariances, but no covariances *across* processes. Therefore, for ease of exposition, we restrict the analysis to $d = 1$, focusing on a one-dimensional model with $n+1$ observations of $Y_i = X_{t_i} + \epsilon_i, i = 0, \ldots, n$. Further, we set $\eta_u = \eta_u^{(1)} = \mathbb{C}\mathrm{ov}(\epsilon_i, \epsilon_{i+u})$ for the u-th order autocovariance, while the long-run variance is now simply denoted by η. Following part (iii) of Assumption 3, for some $R \geq 0$, we may neglect all dependencies $\eta_u, u > R$. Hence, $\epsilon_i, i = 0, \ldots, n$, is an R-dependent process and the returns $\Delta_i Y$ have a MA(R)-structure.

Fix $R \geq 0$ as the order of serial dependence. We shall discuss below how to choose R from the data in practice. We successively estimate the autocovariances by

$$\hat{\eta}_R = (2n)^{-1} \sum_{i=1}^{n} \left(\Delta_i Y\right)^2 + n^{-1} \sum_{r=1}^{R} \sum_{i=1}^{n-r} \Delta_i Y \Delta_{i+r} Y \,, \tag{19a}$$

$$\hat{\eta}_r - \hat{\eta}_{r+1} = (2n)^{-1} \sum_{i=1}^{n} \left(\Delta_i Y\right)^2 + n^{-1} \sum_{u=1}^{r} \sum_{i=1}^{n-u} \Delta_i Y \Delta_{i+u} Y \,,\ 0 \leq r \leq R-1 \,. \tag{19b}$$

In particular, this includes

$$\hat{\eta}_0 - \hat{\eta}_1 = (2n)^{-1} \sum_{i=1}^{n} \left(\Delta_i Y\right)^2 , \tag{19c}$$

which is the classical estimator of η_0 in an i.i.d. setting as in Zhang et al. (2005). The estimators are $\sqrt{n}$-consistent and satisfy central limit theorems. To construct an estimator for the variance of $\hat{\eta}_r$, denote for $q, r, r' \in \{0, \ldots, R\}$

$$\begin{aligned} \tilde{\Gamma}_q^{rr'} = n^{-1} \sum_{i=1}^{n-(r \vee (q+r'))} & \Delta_i Y \Delta_{i+r} Y, \Delta_{i+q} Y \Delta_{i+q+r'} Y - [\hat{\eta}_r - \hat{\eta}_{r+1} - (\hat{\eta}_{r-1} - \hat{\eta}_r)] \\ & \times [\hat{\eta}_{r'} - \hat{\eta}_{r'+1} - (\hat{\eta}_{r'-1} - \hat{\eta}_{r'})] \,, \end{aligned} \tag{20}$$

where $\hat{\eta}_r - \hat{\eta}_{r+1}$, $\hat{\eta}_{r-1} - \hat{\eta}_r$, $\hat{\eta}_{r'} - \hat{\eta}_{r'+1}$ and $\hat{\eta}_{r'-1} - \hat{\eta}_{r'}$ are computed according to (19b). Then, the variance of $\hat{\eta}_r, 0 \leq r \leq R$, is consistently estimated by

$$\widehat{\mathbb{V}\mathrm{ar}}(\hat{\eta}_r) = n^{-1} \left(V_{r+1}^n + V_r^n + 2C_{r,r+1}^n\right) , \tag{21a}$$

$$\text{with } C_{r,r+1}^n = \left(\frac{\hat{\Gamma}_0^{00}}{4} + \frac{1}{2} \sum_{u=1}^{r} \hat{\Gamma}_u^{00} + \sum_{u=0}^{r} \sum_{u'=1}^{r+1} \left(\hat{\Gamma}_0^{uu'} + 2 \sum_{q=1}^{R} \hat{\Gamma}_q^{uu'} \right) \right) , \tag{21b}$$

and $V_r^n = C_{r,r}^n$. Particularly, for $r = R$, we have $\widehat{\mathbb{V}\mathrm{ar}}(\hat{\eta}_R) = n^{-1} V_R^n$. Below, we give a feasible central limit theorem, which entails an asymptotic distribution-free test of the hypotheses $\mathbb{H}_0^Q : \eta_u = 0$ for all $u \geq Q$, $Q = R + 1$.

Theorem 2. *Under Assumption 3, the following central limit theorem applies to the estimators defined by* (19a) *and* (19b)*:*

$$\sqrt{n} (V_r^n + V_{r+1}^n + 2C_{r,r+1}^n)^{-1/2} \left(\hat{\eta}_r - \eta_r\right) \xrightarrow{d} \mathbf{N}(0,1) \,. \tag{22}$$

Consequently, under $\mathbb{H}_0^Q$*:*

$$T_Q^n(Y) = \sqrt{n/V_Q^n}\,\hat{\eta}_Q \xrightarrow{d} \mathbf{N}(0,1)\,. \tag{23}$$

Theorem 2 is proved in Appendix A.

The statistic $T_Q^n(Y)$ serves as a test statistic for the significance of non-zero autocovariances for certain lags. An accurate strategy to select the order of serial dependence R thus requires computing the test statistics $T_Q^n(Y)$ for $Q \leq \tilde{Q} = \tilde{R} + 1$ large enough and incorporating all autocovariances until the first hypothesis of a zero autocovariance cannot be rejected for a given significance level. Then, denoting the determined order by $\hat{R}$, an estimate of the long-run noise variance, $\hat{\eta}$, is obtained according to (4) based on the individual estimates $\hat{\eta}_0, \ldots, \hat{\eta}_{\hat{R}}$.

3 Simulation Study

We conduct a simulation study to examine two issues. First, we analyze the impact of different choices of the input parameters θ_h, θ_J and θ_K introduced in Section 2.4 on the estimator's finite sample performance. Second, we study the precision of the procedure for estimating the long-run noise variance η as outlined in Section 2.5. As these issues are not related to the dimensionality of the process, we set $d = 1$ for reasons of parsimony.

We assume that the efficient log-price process incorporates both a stochastic and a non-stochastic seasonal volatility component, which is modeled by a Flexible Fourier Form as introduced by Gallant (1981). Correspondingly, we assume the underlying process as given by

$$dX_t = \mu dt + \phi_t \tilde{\sigma}_t dB_t, \tag{24a}$$

$$\ln\left(\phi_t^2\right) = \alpha_\phi t + \beta_\phi t^2 + \sum_{q=1}^{Q} \left[\gamma_{\phi,q}\cos(2\pi q t) + \delta_{\phi,q}\sin(2\pi q t)\right], \tag{24b}$$

where B_t is a standard Brownian motion and $\tilde{\sigma}_t$ the stochastic volatility component. The drift term is set to $\mu = 0.03$. The seasonality component is normalized such that $\int_0^1 \phi_t^2 dt = 1$ with parameter values given by median estimates for mid-quotes of 30 highly liquid constituents of the NASDAQ100 in 2010-14. See Section 4.1 for a summary of the dataset. In this context, we employ the estimation procedure by Andersen and Bollerslev (1997).[4] The stochastic volatility component $\tilde{\sigma}_t$ is assumed to follow one- and two-factor models according to Huang and Tauchen (2005). Details are provided in Appendix B. In all settings, we simulate the process in (24a) and

[4]We estimate the daily volatility component based on sub-sampled five-minute realized variances instead of a parametric GARCH approach. The number of sinusoids Q is chosen based on the BIC.

(24b) by a Euler discretization scheme based on a one-second grid assuming $23,400$ seconds per trading day, while setting $n = 23,400$.

We dilute the observations of the efficient log-price process by serially dependent microstructure noise with $R = 1$, i.e., $\epsilon_i = \theta_\epsilon \epsilon_{i-1} + u_i$, $u_i \sim N\big(0, \eta/\left(1+\theta_\epsilon\right)^2\big)$, $i = 1, \ldots, n$. To ensure that the absolute noise level is in line with the variation in the volatility process, we determine η by choosing the noise-to-signal ratio per trade $\xi^2 := n\eta/\sqrt{\int_0^1 \phi_t^4 \tilde{\sigma}_t^4 dt}$ (see Oomen, 2006). We consider low-noise and high-noise scenarios, setting $\xi = 0.659$ and $\xi = 1.941$, respectively. These numbers correspond to the first and third quartile of respective estimates based on the NASDAQ data employed in our empirical applications. Here, η is estimated following the procedure from Section 2.5, while the integrated quarticity is approximated by the squared sub-sampled five-minute realized variance. Finally, we choose $\theta_\epsilon = 0.441$, yielding a first-order autocorrelation of $\eta_1 = 0.6$, which is the median estimate for the underlying NASDAQ data.

To investigate the impact of chosen input parameters, we compute the LMM estimator (12) over a grid of values for θ_h, θ_J and θ_K. For each combination, we evaluate a normalized root mean integrated squared error, RMISE $:= M^{-1} \sum_{m=1}^{M} \int_0^1 \left[\hat{\sigma}_{t,m}^2 / \left(\tilde{\sigma}_{t,m}^2 \phi_t^2\right) - 1\right]^2 dt$, where $\hat{\sigma}_{t,m}^2$ is the spot variance estimate for time t in replication m, $\tilde{\sigma}_{t,m}^2$ is the corresponding true stochastic volatility component and M is the number of replications.

Table 1 reports the values of the input parameters yielding minimal RMISEs for $M = 3000$. For the setting including the one-factor stochastic volatility model ("1F") and a low noise level, the RMISE-optimal values of the input parameters yield a configuration with $\lceil h_n^{-1} \rceil = 77$ blocks spanning about 5 minutes each, a spectral cut-off $J_n = 120$ and a smoothing window of $K_n = 4$ blocks. A higher noise level implies a decrease in the cut-off to $J_n = 80$ and a lengthening of the smoothing window to $K_n = 5$ blocks. In the two-factor specification ("2F"), the rougher volatility paths result in a smaller optimal value of θ_h, translating into $\lceil h_n^{-1} \rceil = 305$ blocks with length of only about 1 minute and 15 seconds each. For the high noise level, the cut-off reduces to $J_n = 15$, while the smoothing window lengthens to $K_n = 15$. In general, we see that the two-step method (12) clearly outperforms a simple histogram estimator which relies only on the first frequency $J_n = 1$.

The effect of a deviation from the RMISE-optimal values of θ_h, θ_J and θ_K is illustrated in Figure 1 for the one-factor model assuming a high noise level. The plots depict RMISEs for a grid of two of the three input parameters, while the third one is fixed at its optimal value. Evidently, the performance of the estimator is not overly sensitive to the choice of θ_h. Unless the latter is set to an extremely low value or θ_K is extremely small, resulting in excessive undersmoothing, the RMISE is fairly stable. Regarding θ_J, only very small choices imply a considerable reduction in the estimator's precision, as the spectral frequencies are cut off too early. Finally, although θ_K appears to be the most influential parameter, deviations from its

Table 1: RMISE-optimal values of θ_h, θ_J and θ_K based on grid search. Normalized root mean integrated squared errors $M^{-1}\sum_{m=1}^{M}\int_0^1\left[\hat{\sigma}^2_{t,m}/\left(\tilde{\sigma}^2_{t,m}\phi^2_t\right)-1\right]^2 dt$ computed for $M = 3000$ Monte Carlo replications. "1F" and "2F" represent a one- and two-factor model for the stochastic volatility component of the efficient log-price, respectively. ξ denotes the square root of the noise-to-signal ratio per trade, i.e., $\xi^2 := n\eta/\sqrt{\int_0^1 \phi_t^4\tilde{\sigma}_t^4 dt}$. Normalized RMISEs are reported in percentage points.

ξ	Vol. Spec.	θ_h^*	θ_J^*	θ_K^*	RMISE*
0.659	1F	0.20	12.0	0.3	3.11
	2F	0.05	2.0	1.2	4.80
1.941	1F	0.20	8.0	0.4	4.37
	2F	0.05	1.5	1.2	5.84

"optimal" value imply rather mild increases in the RMISE in both directions. We can conclude that the performance of the spot covariance matrix estimator (12) is quite robust for a range of sensible input choices.

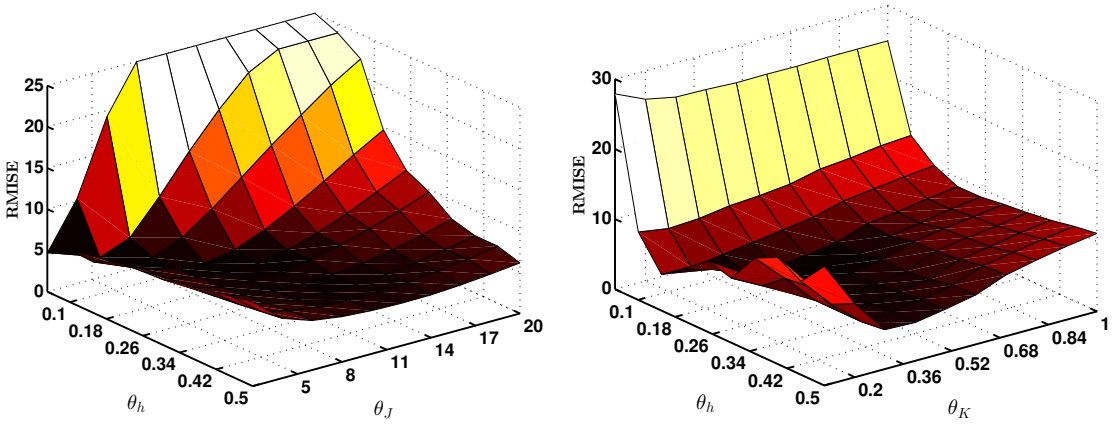
To evaluate the proposed estimator of the long-run noise variance, we consider microstructure noise processes based on different orders of serial dependence R. Hence, we assume $\epsilon_i = \Theta_R(L)\,u_i$, $\Theta_R(L) := \sum_{r=0}^{R}\theta_{\epsilon,r}L^r$, $\theta_{\epsilon,0} = 1$, $u_i \sim N\big(0, \eta/\Theta_R(1)^2\big)$, $i = 1,\ldots,n$. We employ the following settings: (i) $R = 0$, (ii) $R = 1$ with $\theta_{\epsilon,1} = 0.5$, and (iii) $R = 2$ with $\theta_{\epsilon,1} = 0.5$ and $\theta_{\epsilon,2} = 0.3$. Moreover, we select a high noise level by setting $\xi = 3$, which can be considered as a "stress test" for the proposed procedure.

For $M = 5000$ replications, Table 2 shows means and standard deviations for the estimates of R, as well as biases and standard deviations for the estimates of the long-run noise variance η based on $\alpha = 0.05$ and $\tilde{Q} = 15$. We observe that the procedure proposed in Section 2.5 slightly over-estimates R, resulting in more conservative estimates of the order of serial dependence in the noise process. Moreover, when comparing the results for the one- and two-factor stochastic volatility model, the precision of the estimate of R is similar. The two-factor model, however, consistently implies a higher volatility in the estimates of η. Generally, we can conclude that the proposed approach provides a satisfactory precision in a realistic scenario.

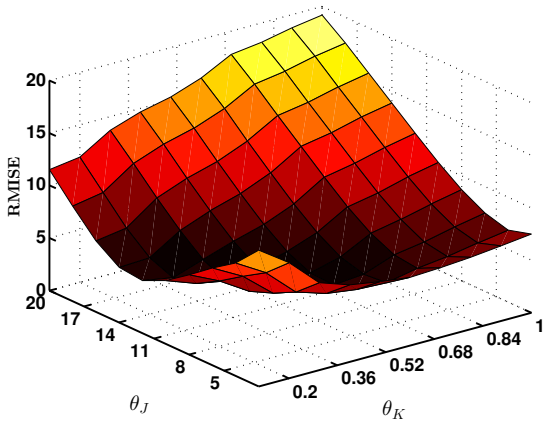
4 Empirical Study

4.1 Data

We consider mid-quote and trade data at the highest possible frequency for 30 of the most liquid constituents of the NASDAQ100 index as well as the PowerShares QQQ Trust, an ETF tracking the NASDAQ100. The sample period is from May 2010 to April 2014. Mid-quotes



(a) $\theta_K = 0.4$ (b) $\theta_J = 8$



(c) $\theta_h = 0.2$

Figure 1: RMISE depending on input parameters (one-factor model and high noise level). Normalized root mean integrated squared errors $M^{-1}\sum_{m=1}^{M}\int_0^1\left[\hat{\sigma}_{t,m}^2/\left(\tilde{\sigma}_{t,m}^2\phi_t^2\right)-1\right]^2 dt$ computed for $M = 3000$ Monte Carlo replications and reported in percentage points. In each subplot, the remaining input parameter is fixed at its optimal value according to Table 1.

Table 2: Descriptive statistics of the estimated order of serial dependence R and estimated long-run variance η of the noise process. R denotes the true order of dependence of the noise process. The settings are: (i) $R = 0$, (ii) $R = 1$ with $\theta_{\epsilon,1} = 0.5$, and (iii) $R = 2$ with $\theta_{\epsilon,1} = 0.5$ and $\theta_{\epsilon,2} = 0.3$. "1F" and "2F" represent a one- and two-factor model for the stochastic volatility component of the efficient log-price, respectively. $\hat{R}$ and $\hat{\eta}$ are computed following Section 2.5 using $\alpha = 0.05$ and $\tilde{Q} = 15$. BIAS$(\hat{\eta})$ and STD$(\hat{\eta})$ are re-scaled by 10^3. Results based on $M = 5000$ Monte Carlo replications.

R	Vol. Spec.	$\bar{\hat{R}}$	STD$(\hat{R})$	BIAS$(\hat{\eta})$	STD$(\hat{\eta})$
0	1F	0.00	0.05	0.002	1.985
	2F	0.00	0.05	0.000	3.467
1	1F	1.13	0.71	-0.031	2.305
	2F	1.11	0.61	-0.026	3.186
2	1F	2.35	1.28	-0.092	2.046
	2F	2.37	1.34	-0.073	2.678

are computed from first-level limit order book data provided by the LOBSTER database.[5] The latter reconstructs the order book from a message stream, which is part of NASDAQ's historical TotalView-ITCH data and contains all limit order submissions, cancellations and executions on each trading day (see Huang and Polak, 2011). Accordingly, the corresponding transaction data can be read out from the above message files directly. A crucial advantage of the resulting datasets, e.g., compared to similar ones sampled from the Trade and Quote (TAQ) database, is the fact that all recorded events are time stamped with at least millisecond precision, which allows for an econometric analysis at the highest resolution possible. Despite the "clean" nature of the data, resulting from the latter being directly taken from NASDAQ's message stream, we handle remaining errors in the trade and mid-quote samples using the cleaning procedures proposed by Barndorff-Nielsen et al. (2009).

Table 3 gives summary statistics of ask and bid quotes recorded whenever the underlying limit order book changes, i.e., induced by a submission of a limit or market order or a cancellation of an existing limit order.[6] The number of limit order book updates is enormous, amounting to several hundred thousands for some stocks. It turns out, however, that a considerable amount of best ask/bid quote revisions equals zero. This is most remarkable for Microsoft, where this number amounts to more than 99%. Since this sheer amount of data makes the computation of the estimators challenging and cumbersome, we construct the estimators employing quote *revisions*. Table 3 also reports the long-run noise variance estimates computed according to Section 2.5 and the corresponding noise-to-signal ratios per observation. It is shown that there is a considerable variation of the noise-to-signal ratio across the different assets. Finally, we find

[5] See `https://lobster.wiwi.hu-berlin.de/`.

[6] All tables and figures underlying the empirical study are given in the Appendix C.

strong empirical support for significant autocorrelations in the noise process, clearly violating the traditional i.i.d. assumption.

We use the mid-quote revisions for the NASDAQ100 constituents to estimate 30×30 spot covariance matrices according to (12), yielding pair-wise spot covariances and correlations, as well as individual volatilities. Further, we also include the QQQ ETF and estimate 31×31 spot covariance matrices to obtain estimates of spot betas with QQQ as market proxy. In both cases, we select the relevant inputs as discussed in Section 2.4. The corresponding proportionality parameters are set to the values found to be "optimal" in the simulation study of Section 3. Taking a more conservative stance, we rely on the "1F"-setting assuming a high noise level.[7] Hence, we set $\theta_h = 0.2$, $\theta_J = 8$, $\theta_K = 0.4$ and $J_n^p = 5$.

Table 4 reports summary statistics for the number of blocks, the spectral cut-off and the length of the smoothing window as induced by the underlying data for both the entire sample and each year. On average, we use approximately 23 blocks per day, resulting in an average block length of 17 minutes. Spectral frequencies are cut off at nearly 54, while the average length of the smoothing window is about 2 blocks, translating into roughly half an hour. Regarding the evolution over time, the average number of blocks increases from the first to second year, subsequently drops, only to increase again in the last year. Each time, changes in the block number are accompanied by moves of the spectral cut-off and the length of the smoothing window in the same direction. Notably, the variation in all three inputs is considerably higher in the first two years than in the last two years of the sample.

4.2 Intraday Behavior of Spot (Co-)Variances

Figure 2 shows the cross-sectional deciles of across-day averages of spot covariances and correlations for each asset pair. Figure 3 depicts the underlying volatilities and betas with respect to the ETF tracking the NASDAQ100 index. We observe distinct intraday seasonality patterns. Covariances clearly decline at the beginning of the trading day, stabilize around noon on a widely constant level and slightly increase before market closure. Interestingly, the resulting correlations show a reverse pattern and significantly increase during the first trading hour. The latter is caused by spot volatilities, that decay faster than the corresponding covariances at the beginning of the trading day. Hence, the (co-)variability between assets is highest after start of trading which might be caused by the processing of common information. The assets' idiosyncratic risk, however, as reflected by spot volatilities, is even higher, overcompensating the effect of high covariances and leading to lower correlations at the beginning of the trading day. Interestingly, spot volatilities drop significantly faster than underlying covariances during

[7] We alternatively compute spot covariances using the configuration that is optimal in the"2F"-scenario and employing transaction prices and find qualitatively similar results. For sake of brevity, we do not report them here.

the first trading hour. Shortly after opening, spot volatilities are approximately twice as high as the (average) *daily* (based on the open-to-close integrated variance estimate) volatility, but strongly decline thereafter. This makes correlations sharply increasing between 10:00 and 11:00 am. Accordingly, we observe that median spot correlations range between approximately 0.2 and 0.4 across a day.[8] This is in contrast to a *daily* correlation (computed from the open-to-close integrated covariance estimate) of approximately 0.32 and shows that even on average, intraday variability of correlations and covariances is substantial. Moreover, daily covariances tend to vary widely in locksteps with the volatility of the underlying ETF, making the corresponding betas close to be constant across the day. Hence, systematic risk (with respect to the NASDAQ100) varies much less than idiosyncratic risk.

In Figure 4, we compute, for each asset pair and each point during the day, the standard deviation of spot covariances and correlations *across* days. We observe that the across-day variability in covariances is highest after market opening and shortly before closure. A similar picture is also observed for spot volatilities in Figure 5. We associate these patterns with effects arising from (overnight) information processing in the morning and increased trading activities in the afternoon, where traders tend to re-balance or close positions before the end of trading. Hence, idiosyncratic effects seem to become stronger during these periods, increasing the variability of (co-)variances. Interestingly, the across-day standard deviations in correlations show a reverse pattern. Thus, across-day variability in intraday correlations is lowest at the beginning of trading, increases until mid-day and is widely constant during the afternoon hours. Here, increased across-day covariance and volatility risk seem to compensate each other. Likewise, the daily variability of spot betas is widely constant through different intra-day time points (except for the highest decile).

To evaluate the intraday variability of all spot quantities, we compute a proxy for the total intraday variation normalized by the L_1-norm, i.e.,

$$\tilde{V}_f^{\text{norm}} = \sum_{i=1}^{n_g} |f(t_i) - f(t_{i-1})| \left[\sum_{i=1}^{n_g} |f(t_i)| \, \Delta t_i \right]^{-1},$$

where $f(\cdot)$ stands for the respective spot quantity of interest and n_g denotes the number of underlying grid points per day. Figures 6 and 7 show the time series of cross-sectional medians of the corresponding intraday variation measures. The measures' normalization makes them comparable across days and quantities, while providing insights into intraday (co-)variation risks. We observe that the latter are strongly time-varying and are clustered over time. Hence, intraday (co-)variation risks seem to be relatively persistent, following some long-term movements. On

[8] Note that we obtain very similar pictures for *absolute* correlations and covariances. Hence, during the analyzed period, correlations between NASDAQ100 assets are widely positive.

individual days, we observe, nevertheless, exceptionally high intraday fluctuations of these quantities. These are likely days, where the market faces fundamental news arrival or unusual activities, e.g., induced by flash crashes. Indeed, in Section 4.3, we will focus on three selected days where markets are confronted with extreme events. These three days, corresponding to the dates 05/06/10, 12/27/12 and 04/23/13 (more details will follow in Section 4.3), are indicated by the vertical lines and indeed reflect salient intraday (co-)variation risks. Interestingly, normalized variations are highest for covariances and correlations, which might be also partly induced by higher estimation errors for covariances compared to variances. Accordingly, the variability of spot volatilities is generally lower, but can still be substantial on individual days. Confirming the findings above, the variability in betas is generally lowest. However, even betas – reflecting systematic risk – are far from being constant on selective days.

Figure 8 reports the (averaged) autocorrelation functions (ACFs) of all four quantities with the figures being constructed such that one lag corresponds to approximately five minutes. We observe that all (co-)variability measures are strongly serially correlated across short time intervals with first-order autocorrelations being around 0.95. Nevertheless, the ACFs decay relatively fast *within* a day. This is most extreme for spot volatilities, where the ACF declines from 0.95 at the 5 minute lag to below 0.1 after approximately 3.5 hours. The noticeable seasonality pattern in the ACFs for covariances and volatilities underlines distinct *daily* autocorrelations, which considerably exceed the *intradaily* autocorrelations at slightly smaller lags. For correlations and betas, this pattern is less pronounced. Here, long-term autocorrelations stabilize around 0.25 and decay very slowly.

4.3 Event Studies

The previous section shows that spot correlations and covariances can substantially vary during a day, even if these patterns are averaged across time and assets. Here, we aim at analyzing the behavior of spot (co-)variability in extreme market periods. The first study analyzes the flash crash on 05/06/10. Figure 9 shows the intraday movements of the QQQ ETF over the entire trading day and a time window starting at 1:30 pm. We associate individual time points with the following events:[9] (1) At $\approx$ 2:00 pm, protests in Athens related to the Euro crisis trigger a sharp down movement of the Euro especially vs. the Yen. In the U.S., fund managers start large-scale short-selling of futures contracts on the S&P500 ("E-Mini"), leading to a trading volume which is six times higher than usual. (2) At $\approx$ 2:35 pm, the E-mini market makers cut back trading. (3) At 2:37 pm, NASDAQ stops routing orders to ARCA, the electronic trading platform of NYSE, due to huge lags in order acknowledgement. (4) At $\approx$ 2:45 pm, rumors

[9]See `http://online.wsj.com/news/articles/SB10001424052748704545004575353443450790402`. For a detailed overview, we refer to CFTC and SEC (2010).

spread suggesting that the decline occurred due to a "fat-finger" error of a Citigroup trader, and not because of an adverse news shock. This helps stabilizing markets and liquidity in futures trading rebounds. (5) At 3:01 pm, NASDAQ resumes routing to ARCA. Until market closure, trading remains erratic.

Figure 10 shows the cross-sectional deciles of resulting spot covariances and correlations on this day. We observe that covariances are virtually constant during the morning, but nearly instantaneously increase shortly after 2:00 pm, i.e., after large-scale short selling of E-Mini futures started. The deciles show that the cross-sectional distribution of covariances across all asset pairs is very skewed, revealing huge upward shifts in some covariances, but only very moderate reactions in others. Figure 11, however, shows that the corresponding reactions in spot volatilities have been much stronger. This even leads to *declining* correlations with median correlations dropping from approx. 0.5 before 2:00 pm to approx. 0.3 around the peak of the crash. Hence, this event is characterized by an explosion of idiosyncratic risk, which overcompensates increases in covariances and ultimately *decreases* correlations. Likewise, systematic risk, as reflected by the corresponding spot betas significantly declines during this period. This is due to the fact that the spot volatility of the NASDAQ100 increases much stronger than the underlying covariances. Hence, we can summarize that the May 2010 flash crash particularly moved spot volatilities, but not too much spot covariances, even reducing systematic risk.

The next event study illustrates an extreme situation of different type. Here, we analyze intraday risk on 12/27/12, when at approximately 10:00 am the U.S. senate majority leader stated that a resolution of the U.S. "fiscal cliff" (i.e., budgetary deficits reaching the legal upper bound) before January 1, 2013, was unlikely due to lack of cooperation by Republicans.[10] As shown in Figure 12, this caused prices to fall. Around 2:20 pm, a news release reported that the House of Representatives would convene on the following Sunday in an attempt to end the "fiscal cliff" crisis. This, in turn pushed the market significantly upwards. Figures 13 and 14 show that immediately after the announcement of this positive news, covariances rise significantly and more than triple (on average). Similarly, volatilities increase, as well, but on average only moderately. This is in sharp contrast to the (co-)variability patterns found for the May 2010 flash crash, where an explosion in idiosyncratic risk dominates. Consequently, the positive "fiscal cliff news" lead to a significant *increase* in spot correlations as co-movements dominate. Systematic risk as reflected by spot betas remains widely unchanged. Thus, covariances increase to the same extent as the volatility of the (NASDAQ100) market.

[10] See `http://money.msn.com/now/post.aspx?post=73878f29-4fdb-45c5-baf6-0f83f40-b821c&_p=986b65a2-3eea-479c-a4a0-ba9d3988b0e0`.

Finally, we study a third type of event, which is characterized by completely non-anticipated (and ultimately wrong!) news. On 04/23/13 at around 1:07 pm, a fake tweet from the account of the Associated Press (AP) reported "breaking" news on two explosions in the White House, where the U.S. president (supposedly) got injured.[11] At 1:10 pm, AP officially denied this message and suspended its twitter account at 1:14 pm. Figure 15 shows the underlying price process and the timing of the corresponding events. Our results in Figures 16 and 17 show that (co-)variances and correlations strongly increase even before 1:00 pm. While the latter finding suggests that rumors on the market might have been present even before the AP announcement, one has to take into account that our spot estimator actually computes the rolling mean over, on average, 5 block-wise estimates with each block being almost 17 minutes long. The increase in covariances is stronger than for volatilities, whereby correlations (on average) increase from approximately 0.2 to 0.8(!). As in the previous scenario, betas remain widely unaffected. Our estimates show that this effect has been present for approximately one hour.[12] This finding is remarkable given that the flash crash itself lasted only a couple of minutes and is similar to the effects observed during the May 2010 flash crash. Hence, effects of (flash) crashes on covariances tend to remain in the market for a considerable time period.

In summary, we conclude on the following findings: First, extreme (news) events cause abrupt upward shifts in (co-)variances. After the market has processed this news, (co-)variances move back, which again occurs relatively sharply. Our spot estimator seems to capture these effects quite well, since the observed reactions in the spot quantities are very well aligned with the timing of the underlying event. This indicates that the estimators are suitable to capture changes in dependence structures on a high time resolution. Second, correlations do change abruptly, as well, but the direction of the movements is ambiguous. Obviously, depending on the type of news, we observe that the magnitudes of changes in covariances can be much higher or smaller than those in volatilities. In the first case, the market tends to strongly co-move and seems to be driven by news, which affects all assets very similarly. In the second case, idiosyncratic effects dominate, leading to a *decline* in correlations. Third, betas seem to be least affected by these extreme scenarios. In a situation of increasing correlations, we observe that covariances tend to change by the same magnitude as the volatility of the market (NASDAQ100) index, leaving betas widely constant. Conversely, in a scenario with decreasing correlations, the dominance of idiosyncratic effects tends to lower betas and thus reduces systematic risk components.

[11] See `http://buzz.money.cnn.com/2013/04/23/ap-tweet-fake-white-house/?iid=EL`.

[12] We double-checked that this is *not* induced by the choice of the smoothing window.

5 Conclusion

In this paper, we introduce an estimator for spot covariance matrices, which is constructed based on local averages of block-wise estimates of locally constant covariances. The proposed estimator builds on the local method of moments approach introduced by Bibinger et al. (2014). We show how to extend the LMM approach to the case of autocorrelations in market microstructure noise and provide a suitable procedure for choosing the lag order in practice. For the resulting spot covariance matrix estimator, we derive a stable central limit theorem along with a feasible version that is straightforwardly applicable in empirical practice. An important result is that we are able to attain the optimal convergence rate, which is $n^{1/8}$ under the assumption of a semi-martingale volatility matrix process with the efficient log-prices being subject to noise and a non-synchronous observation scheme.

Extensive simulation exercises provide guidance on how to implement the estimator in practice and demonstrate its relative insensitivity with respect to the choice of block sizes, cut-offs and smoothing windows. Moreover, based on NASDAQ blue chip stocks, we provide detailed empirical evidence on the intraday behavior of spot covariances, correlations, volatilities and betas. We show that particularly the former three reveal distinct intraday seasonality patterns. Moreover, we analyze how spot covariances change in periods of extreme market movements and show that intraday changes of (co-)volatility structures can be quite distinct and considerable. It turns out that our estimator is able to capture these changes quite accurately, which makes it a good candidate for analyzing the (co-)variability of assets on a high-frequency time scale.

References

ABADIR, K. M. AND J. R. MAGNUS (2005): *Matrix algebra*, vol. 1 of *Econometric Exercises*, Cambridge: Cambridge University Press.

AIT-SAHALIA, Y., J. FAN, AND D. XIU (2010): "High-Frequency Estimates With Noisy and Asynchronous Financial Data," *Journal of the American Statistical Association*, 105, 1504–1516.

AIT-SAHALIA, Y. AND J. JACOD (2009): "Testing for jumps in a discretely observed process," *Annals of Statistics*, 37, 184–222.

AIT-SAHALIA, Y., L. ZHANG, AND P. A. MYKLAND (2011): "Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise," *Journal of Econometrics*, 160, 160–165.

ALTMEYER, R. AND M. BIBINGER (2014): "Functional Stable Limit Theorems for Efficient Spectral Covolatility Estimators," SFB 649 Discussion Papers SFB649DP2014-005, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

ANDERSEN, T. G. AND T. BOLLERSLEV (1997): "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

ANDERSEN, T. G., D. DOBREV, AND E. SCHAUMBURG (2009): "Duration-Based Volatility Estimation," Global COE Hi-Stat Discussion Paper Series gd08-034, Institute of Economic Research, Hitotsubashi University.

BANDI, F. M. AND R. RENO (2009): "Nonparametric Stochastic Volatility," Global COE Hi-Stat Discussion Paper Series gd08-035, Institute of Economic Research, Hitotsubashi University.

BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2009): "Realized Kernels in Practice: Trades and Quotes," *Econometrics Journal*, 12, C1–C32.

——— (2011): "Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading," *Journal of Econometrics*, 162, 149–169.

BARNDORFF-NIELSEN, O. E. AND N. SHEPHARD (2004): "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72, 885–925.

BIBINGER, M., N. HAUTSCH, P. MALEC, AND M. REISS (2014): "Estimating the Quadratic Covariation Matrix from Noisy Observations: Local Method of Moments and Efficiency," *Annals of Statistics*, 42, 1312–1346.

BIBINGER, M. AND L. WINKELMANN (2013): "Econometrics of Co-Jumps in High-Frequency Data with Noise," SFB 649 Discussion Papers SFB649DP2013-021, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

BILLINGSLEY, P. (1991): *Probability and Measure*, Springer, New York, 2 ed.

BOS, C. S., P. JANUS, AND S. J. KOOPMAN (2012): "Spot Variance Path Estimation and Its Application to High-Frequency Jump Testing," *Journal of Financial Econometrics*, 10, 354–389.

CFTC AND SEC (2010): "Findings Regarding the Market Events of May 6, 2010," Tech. rep., U.S. Commodity Futures Trading Commission and U.S. Securities & Exchange Commission.

CHERNOV, M., A. R. GALLANT, E. GHYSELS, AND G. TAUCHEN (2003): "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225 – 257.

CHRISTENSEN, K., R. OOMEN, AND M. PODOLSKIJ (2014): "Fact or Friction: Jumps at Ultra High Frequency," *Journal of Financial Economics*, forthcoming.

CHRISTENSEN, K., M. PODOLSKIJ, AND M. VETTER (2013): "On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes." *Journal of Multivariate Analysis*, 120, 59–84.

CURCI, G. AND F. CORSI (2012): "Discrete Sine Transform for Multi-Scales Realized Volatility Measures," *Quantitative Finance*, 12, 263–279.

DE SANTIS, G. AND B. GERARD (1997): "International Asset Pricing and Portfolio Diversification with Time-Varying Risk," *The Journal of Finance*, 52, 1881–1912.

FAN, J. AND Y. WANG (2008): "Spot Volatility Estimation for High-Frequency Data," *Statistics and Its Interface*, 1, 279–288.

FOSTER, D. P. AND D. B. NELSON (1996): "Continuous Record Asymptotics for Rolling Sample Variance Estimators," *Econometrica*, 64, 139–174.

GALLANT, A. R. (1981): "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form : The Fourier Flexible Form," *Journal of Econometrics*, 15, 211–245.

HANSEN, P. R., J. LARGE, AND A. LUNDE (2008): "Moving Average-Based Estimators of Integrated Variance," *Econometric Reviews*, 27, 79–111.

HANSEN, P. R. AND A. LUNDE (2006): "Realized Variance and Market Microstructure Noise," *Journal of Business & Economic Statistics*, 24, 127–161.

HAUTSCH, N. AND M. PODOLSKIJ (2013): "Preaveraging-Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence," *Journal of Business & Economic Statistics*, 31, 165–183.

HAYASHI, T. AND N. YOSHIDA (2011): "Nonsynchronous covariation process and limit theorems," *Stochastic Processes and their Applications*, 121, 2416–2454.

HUANG, R. AND T. POLAK (2011): "LOBSTER: Limit Order Book Reconstruction System," Technical report, Humboldt-Universität zu Berlin.

HUANG, X. AND G. TAUCHEN (2005): "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3, 456–499.

JACOD, J. (2012): "Statistics and high frequency data." *Proceedings of the 7th Séminaire Européen de Statistique, La Manga, 2007: Statistical methods for stochastic differential equations, edited by M. Kessler, A. Lindner and M. Sørensen*.

JACOD, J., Y. LI, P. A. MYKLAND, M. PODOLSKIJ, AND M. VETTER (2009): "Microstructure Noise in the Continous Case: the Pre-Averaging Approach," *Stochastic Processes and their Applications*, 119, 2803–2831.

KALNINA, I. AND O. LINTON (2008): "Estimating Quadratic Variation Consistently in the Presence of Endogenous and Diurnal Measurement Error," *Journal of Econometrics*, 147, 47–59.

KOIKE, Y. (2014): "Quadratic covariation estimation of an irregularly observed semimartingale with jumps and noise," *preprint, arXiv:1408.0938*.

KRISTENSEN, D. (2010): "Nonparametric filtering of the realized spot volatility: a kernel-based approach," *Econometric Theory*, 26, 60–93.

LI, Y., P. A. MYKLAND, E. RENAULT, L. ZHANG, AND X. ZHENG (2014): "Realized Volatility when Sampling Times are Possibly Endogenous," *Econometric Theory*, 30, 580–605.

LONGIN, F. AND B. SOLNIK (2001): "Extreme Correlation of International Equity Markets," *The Journal of Finance*, 56, 649–676.

MANCINI, C., V. MATTIUSSI, AND R. RENO' (2012): "Spot Volatility Estimation Using Delta Sequences," Working Papers - Mathematical Economics 2012-10, Universita' degli Studi di Firenze, Dipartimento di Scienze per l'Economia e l'Impresa.

MYKLAND, P. A. AND L. ZHANG (2008): "Inference for Volatility-Type Objects and Implications for Hedging," *Statistics and Its Interface*, 1, 255–278.

OOMEN, R. C. (2006): "Comment on Hansen, P. R., and Lunde, A. (2006), "Realized Variance and Market Microstructure Noise"," *Journal of Business & Economic Statistics*, 24, 195–202.

ZHANG, L. (2011): "Estimating Covariation: Epps Effect and Microstructure Noise," *Journal of Econometrics*, 160, 33–47.

ZHANG, L., P. A. MYKLAND, AND Y. AIT-SAHALIA (2005): "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411.

ZU, Y. AND P. H. BOSWIJK (2014): "Estimating Spot Volatility with High-Frequency Financial Data," *Journal of Econometrics*, forthcoming.

A Proofs

A.1 Preliminaries

Consider the process

$$\tilde{X}_t = \int_0^t \sigma_{\lfloor sh_n^{-1} \rfloor h_n} \, dB_s \,, \tag{25}$$

without drift and with block-wise constant volatility as a simplified approximation of X. In the following, we distinguish between the estimator of the spot covariance matrix (12) based on oracle optimal weights (13), $\hat{\Sigma}_s^{or}$, and the adaptive estimator $\hat{\Sigma}_s$. Furthermore, we write $\hat{\Sigma}_s(\tilde{X} + \epsilon)$ for the estimator built from observations in the simplified model in which $\tilde{X}$ is observed in noise and denote the associated spectral statistics by:

$$\tilde{S}_{jk} = \pi j h_n^{-1} \left(\sum_{i=1}^{n_p} \left(\tilde{X}^{(p)}_{t_i^{(p)}} + \epsilon_i^{(p)} - \tilde{X}^{(p)}_{t_{i-1}^{(p)}} - \epsilon_{i-1}^{(p)} \right) \Phi_{jk}\left(\tfrac{t_i^{(p)} - t_{i-1}^{(p)}}{2} \right) \right)_{1 \le p \le d} . \tag{26}$$

On the compact interval $[0, 1]$, it can be assumed that $\|b_s\|, \|\sigma_s\|, \|\tilde{b}_s\|, \|\tilde{\sigma}_s\|$ are uniformly bounded. This is based on Jacod (2012), Lemma 6. 6 in Section 6. 3.

For the order of the weights we have by Lemma C.1 of Bibinger et al. (2014) uniformly over all k that

$$\|W_j\big(\mathbf{H}_k^n, \Sigma_{kh_n}\big)\| \lesssim (\log(n))^{-1}\big(1 + j^2(nh_n^2)^{-1}\big)^{-2}. \tag{27}$$

We introduce the short notation $\bar{t}_i^{(p)} = (1/2)\big(t_i^{(p)} + t_{i-1}^{(p)}\big)$. Recall the summation by parts identity from Altmeyer and Bibinger (2014) given by

$$\begin{aligned} S_{jk}^{(p)} &\approx -\sum_{v=1}^{n_p-1} Y_v^{(p)}\Big(\Phi_{jk}\big(\bar{t}_{v+1}^{(p)}\big) - \Phi_{jk}\big(\bar{t}_v^{(p)}\big)\Big) \\ &\approx -\sum_{v=1}^{n_p-1} Y_v^{(p)}\varphi_{jk}(t_v^{(p)})\frac{t_{v+1}^{(p)} - t_{v-1}^{(p)}}{2}, \end{aligned} \tag{28}$$

with $\varphi_{jk}(t) = \Phi_{jk}'(t) = \sqrt{2}h_n^{-1/2}\cos\big(j\pi h_n^{-1}(t - kh_n)\big)\mathbb{1}_{[kh_n,(k+1)h_n]}(t)$. The first remainder, which is only due to end-effects when $t_0^{(p)} \neq 0$ or $t_{n_p}^{(p)} \neq 1$, and the second remainder by application of the mean value theorem and passing to arguments $t_v^{(p)}$ are asymptotically negligible. The following orthogonality approximations are satisfied by Φ_{jk}, φ_{jk}:

$$\begin{aligned} \sum_{i=1}^{n_p}\Phi_{jk}\big(\bar{t}_i^{(p)}\big)\Phi_{qk}\big(\bar{t}_i^{(p)}\big)\big(t_i^{(p)} - t_{i-1}^{(p)}\big) &= (\delta_{jq} + \mathcal{O}(1))\int_0^1 \Phi_{jk}^2(t)\,dt \\ &= (\delta_{jq} + \mathcal{O}(1))h_n^2\pi^{-2}j^{-2}, \end{aligned} \tag{29a}$$

with $\delta_{jq} = \mathbb{1}_{\{j=q\}}$ being Kronecker's delta. Likewise,

$$\begin{aligned} \sum_{i=1}^{n_p-1}\left(\varphi_{jk}\big(t_i^{(p)}\big)\varphi_{qk}\big(t_i^{(p)}\big)\frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\right) &= (\delta_{jq} + \mathcal{O}(1))\int_0^1 \varphi_{jk}^2(t)\,dt \\ &= (\delta_{jq} + \mathcal{O}(1)). \end{aligned} \tag{29b}$$

Moreover, we have the following approximations:

$$\begin{aligned} \sum_{i=1}^{n_p-1}\varphi_{jk}^2\big(t_i^{(l)}\big)\Big(\tfrac{t_{i+1}^{(l)} - t_{i-1}^{(l)}}{2}\Big)^2 &\approx \sum_{i=1}^{n_l-1}\varphi_{jk}^2\big(t_i^{(l)}\big)\tfrac{t_{i+1}^{(l)} - t_{i-1}^{(l)}}{2}\tfrac{(F_l^{-1})'(kh_n)}{n_l} \\ &\approx \Big(\int_0^1 \varphi_{jk}^2(t)\,dt\Big)\tfrac{(F_l^{-1})'(kh_n)}{n_l}, \end{aligned} \tag{30}$$

$$\sum_{kh_n \le t_i^{(l)} \le (k+1)h_n} \big(t_i^{(l)} - t_{i-1}^{(l)}\big)^2 \approx \sum_{kh_n \le t_i^{(l)} \le (k+1)h_n} (F_l^{-1})'(kh_n) n_l^{-1} \big(t_i^{(l)} - t_{i-1}^{(l)}\big) \tag{31}$$
$$= (F_l^{-1})'(kh_n) n_l^{-1} h_n \,,$$

where the remainders are asymptotically negligible. The last quantity reflects the local variation of observation times similar to the (global) quadratic variation of time by Zhang et al. (2005).

A.2 Proof of Theorem 1

For the sake of notational brevity, we present the proof of Theorem 1 for time points that lie in the interior of the unit interval. Therefore, in (12), we have $L_{s,n} = \lfloor sh_n^{-1} \rfloor - K_n$, $U_{s,n} = \lfloor sh_n^{-1} \rfloor + K_n$ and $U_{s,n} - L_{s,n} + 1 = 2K_n + 1$. For points in the boundary region, the proof proceeds completely analogous and requires only the obvious changes in the limits of summation, as well as in the resulting length of the smoothing window.

Recall the notation from (4), where $\eta_p, p = 1, \dots, d$, refers to the long-run noise variance, namely the sum of all autocovariances, in component p.

Lemma 1. *By Assumption 3, we have for $p, q \in \{1, \dots, d\}, p \neq q$, that*

$$\mathbb{E}\left[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \tfrac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2} \Big)^2 \right] = \nu_p \big(F_p^{-1}\big)' \eta_p n^{-1} + \mathcal{O}\big(n^{-1}\big) \,, \tag{32a}$$

$$\mathbb{E}\left[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \tfrac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2} \Big) \Big(\sum_{i=1}^{n_q} \epsilon_i^{(q)} \varphi_{jk}\big(t_i^{(q)}\big) \tfrac{t_{i+1}^{(q)} - t_{i-1}^{(q)}}{2} \Big) \right] = 0 \,, \tag{32b}$$

$$\mathbb{E}\left[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \tfrac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2} \Big)^4 \right] = \nu_p^2 \big(\big(F_p^{-1}\big)' \big)^2 3\, \eta_p^2 n^{-2} + \mathcal{O}\big(n^{-2}\big) \,, \tag{32c}$$

$$\mathbb{E}\left[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \tfrac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2} \Big)^2 \Big(\sum_{i=1}^{n_q} \epsilon_i^{(q)} \varphi_{jk}\big(t_i^{(q)}\big) \tfrac{t_{i+1}^{(q)} - t_{i-1}^{(q)}}{2} \Big)^2 \right]$$
$$= \nu_p \big(F_p^{-1}\big)' \eta_p \nu_q \big(F_q^{-1}\big)' \eta_q n^{-2} + \mathcal{O}\big(n^{-2}\big) \,. \tag{32d}$$

Proof. Using (31) and an analogous estimate, we infer that

$$\begin{aligned}\mathbb{E}\Big[\Big(\sum_{i=1}^{n_p}\epsilon_i^{(p)}\varphi_{jk}\big(t_i^{(p)}\big)\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)^2\Big] &= \mathbb{E}\Big[\sum_{i=1}^{n_p}\big(\epsilon_i^{(p)}\big)^2\varphi_{jk}^2\big(t_i^{(p)}\big)\Big(\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)^2\\ &\quad + 2\sum_{i=1}^{n_p}\sum_{u=1}^{R}\epsilon_i^{(p)}\epsilon_{i+u}^{(p)}\varphi_{jk}\big(t_i^{(p)}\big)\varphi_{jk}\big(t_{i+u}^{(p)}\big)\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\\ &\quad\times\tfrac{t_{i+u+1}^{(p)}-t_{i+u-1}^{(p)}}{2}\Big]\\ &= \nu_p\big(F_p^{-1}\big)'\eta_p n^{-1}+\mathcal{O}\big(n^{-1}\big).\end{aligned}$$

Here and frequently below, we consider simple approximations for $\varphi_{jk}(t)-\varphi_{jk}(s)$ as $t-s=\mathcal{O}(n^{-1})$, using $\cos(t)-\cos(s)=-2\sin((t-s)/2)\sin((t+s)/2)$.
We also introduce the shortcut $\delta_{i,v}^R=\mathbb{1}_{\{|i-v|\le R\}}$ for the following calculations. Accordingly, the fourth moments yield

$$\begin{aligned}&\mathbb{E}\Big[\Big(\sum_{i=1}^{n_p}\epsilon_i^{(p)}\varphi_{jk}\big(t_i^{(p)}\big)\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)^4\Big] =\\ &\mathbb{E}\Big[\sum_{i,v,u,r=1}^{n_p}\epsilon_i^{(p)}\epsilon_v^{(p)}\epsilon_u^{(p)}\epsilon_r^{(p)}\varphi_{jk}\big(t_i^{(p)}\big)\varphi_{jk}\big(t_v^{(p)}\big)\varphi_{jk}\big(t_u^{(p)}\big)\varphi_{jk}\big(t_r^{(p)}\big)\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\tfrac{t_{v+1}^{(p)}-t_{v-1}^{(p)}}{2}\\ &\qquad\times\tfrac{t_{u+1}^{(p)}-t_{u-1}^{(p)}}{2}\tfrac{t_{r+1}^{(p)}-t_{r-1}^{(p)}}{2}\Big]\\ &=\sum_{i,v,u,r=1}^{n_p}\mathbb{E}\big[\epsilon_i^{(p)}\epsilon_v^{(p)}\epsilon_u^{(p)}\epsilon_r^{(p)}\big]\big(\delta_{i,v}^R\delta_{u,r}^R+\delta_{i,u}^R\delta_{v,r}^R+\delta_{i,r}^R\delta_{v,u}^R\big)\varphi_{jk}\big(t_i^{(p)}\big)\varphi_{jk}\big(t_v^{(p)}\big)\\ &\qquad\times\varphi_{jk}\big(t_u^{(p)}\big)\varphi_{jk}\big(t_r^{(p)}\big)\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\tfrac{t_{v+1}^{(p)}-t_{v-1}^{(p)}}{2}\tfrac{t_{u+1}^{(p)}-t_{u-1}^{(p)}}{2}\tfrac{t_{r+1}^{(p)}-t_{r-1}^{(p)}}{2}\\ &=\nu_p^2\big(\big(F_p^{-1}\big)'\big)^2 3\,\eta_p^2 n^{-2}-R_n\,,\end{aligned}$$

with a remainder R_n, which satisfies for some constant C that

$$\begin{aligned}R_n &\lesssim \sum_{i,v,u,r=1}^{n_p} C\Big(\delta_{i,v}^R\delta_{u,r}^R\big(\delta_{i,u}^R+\delta_{v,r}^R+\delta_{i,r}^R+\delta_{v,u}^R\big)+\delta_{i,u}^R\delta_{v,r}^R\big(\delta_{i,v}^R+\delta_{u,r}^R+\delta_{i,r}^R+\delta_{v,u}^R\big)\\ &\qquad+\delta_{i,r}^R\delta_{v,u}^R\big(\delta_{i,v}^R+\delta_{u,r}^R+\delta_{i,u}^R+\delta_{v,r}^R\big)\Big)n^{-4}\\ &=\mathcal{O}\big(nR^3n^{-4}\big)=\mathcal{O}\big(n^{-3}\big)=o\big(n^{-2}\big)\,.\end{aligned}$$

Thus, R_n is of smaller order than the leading term. Roughly speaking, the expectation above vanishes if no two pairs of indices are in the range of autocorrelations. The terms where

both pairs are again correlated is asymptotically negligible. The statements for terms with $p \neq q$ readily follow from the fact that we have non-correlated noise components and that $\mathbb{E}[\epsilon_i^{(p)}] = \mathbb{E}[\epsilon_v^{(q)}] = 0$ for all i, v. □

Decompose the estimation error of the estimator (12) as follows:

$$\begin{aligned} n^{\beta/2} \operatorname{vec}\left(\hat{\Sigma}_s - \Sigma_s\right) &= n^{\beta/2} \operatorname{vec}\left(\hat{\Sigma}_s^{or}(\tilde{X} + \epsilon) - \Sigma_s\right) \\ &\quad + n^{\beta/2} \operatorname{vec}\left(\hat{\Sigma}_s^{or} - \hat{\Sigma}_s^{or}(\tilde{X} + \epsilon)\right) \\ &\quad + n^{\beta/2} \operatorname{vec}\left(\hat{\Sigma}_s - \hat{\Sigma}_s^{or}\right). \end{aligned}$$

Theorem 1 is implied by Proposition A.1, Proposition A.2 and Proposition A.3 below.

Proposition A.1. *On the assumptions of Theorem 1, for any $s \in (0,1)$ it holds true that*

$$n^{\beta/2} \operatorname{vec}\left(\hat{\Sigma}_s^{or}(\tilde{X} + \epsilon) - \Sigma_s\right) \overset{d-(st)}{\longrightarrow} \mathbf{N}\Big(0, 2\big(\Sigma \otimes \Sigma_H^{1/2} + \Sigma_H^{1/2} \otimes \Sigma\big)_s \mathcal{Z}\Big). \tag{33}$$

Proof. We consider the estimator (12) under observations of the simplified process $\tilde{X} + \epsilon$ and associated spectral statistics $\tilde{S}_{jk}$ in (26). By virtue of the identity

$$\mathbb{COV}\Big(\operatorname{vec}\left(\tilde{S}_{jk}\tilde{S}_{mk}^{\top}\right)\Big) = I_{jk}^{-1}\mathcal{Z}(\delta_{jm} + \mathcal{O}(1)), \tag{34}$$

we deduce that the variance-covariance matrix is

$$\begin{aligned} \mathbb{COV}\Big(\operatorname{vec}\left(\hat{\Sigma}_s^{or}\right)\Big) &= \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} (2K_n+1)^{-2} \sum_{j=1}^{J_n} W_{jk}\mathbb{COV}\Big(\operatorname{vec}\left(\tilde{S}_{jk}\tilde{S}_{jk}^{\top}\right)\Big)W_{jk}^{\top} \\ &\qquad + \mathcal{O}(K_n^{-1}) \\ &= \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} (2K_n+1)^{-2} I_k^{-1} \sum_{j=1}^{J_n} I_{jk} I_{jk}^{-1} \mathcal{Z} I_{jk} I_k^{-1} + \mathcal{O}(K_n^{-1}) \\ &= \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} (2K_n+1)^{-2} I_k^{-1}\mathcal{Z} + \mathcal{O}(K_n^{-1}), \end{aligned}$$

since the covariances between different blocks are asymptotically negligible, which can be seen by considering conditional expectations. Recall the definition of the symmetric matrices I_k and I_{jk} from (13). Having established the asymptotic equality (34), the asymptotic form of the variance-covariance matrix follows analogously to the proof of Corollary 4.3 of Bibinger et al. (2014). There, the asymptotic theory is pursued for continuous observations, but, once we have the illustration above for I_k^{-1}, the analysis is the same. Using block-wise transformations which

diagonalize Σ_{kh_n} and transfer the noise level (5) to the identity matrix, i.e.,

$$\Lambda_{kh_n} = O_k H_{kh_n} \Sigma_{kh_n} H_{kh_n} O_k^\top,$$

with O_k being orthogonal matrices and Λ_{kh_n} being diagonal, we can infer the asymptotic form via

$$\mathbb{COV}\Big(\operatorname{vec}\big(\hat{\Sigma}_s^{or}\big)\Big) = \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} (2K_n+1)^{-2} \big(O_k H_k^{-1}\big)^{-\otimes 2} \tilde{I}_k^{-1} \big(H_k^{-1} O_k^\top\big)^{-\otimes 2} \mathcal{Z} + \mathcal{O}(K_n^{-1}),$$

with a diagonalized version $\tilde{I}_k$ of I_k. Along the same lines as in the proof of Corollary 4.3 in Bibinger et al. (2014), we derive that

$$\mathbb{COV}\Big(\operatorname{vec}\big(\hat{\Sigma}_s^{or}\big)\Big) = (2+\mathcal{o}(1)) \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} (2K_n+1)^{-2} \times \big(\Sigma_{kh_n} \otimes \big(\Sigma_H^{kh_n}\big)^{1/2} + \big(\Sigma_H^{kh_n}\big)^{1/2} \otimes \Sigma_{kh_n}\big) \mathcal{Z},$$

with the short notation $\Sigma_H^{kh_n} = H_{kh_n}\big(H_{kh_n}^{-1}\Sigma_{kh_n}H_{kh_n}^{-1}\big)^{1/2}H_{kh_n}$. The expression in Theorem 1 now follows using the smoothness of Σ and H granted by Assumptions 2-4. Therefore, to deduce the variance-covariance structure, it remains to prove that (34) is indeed valid. The computations rely on the preliminaries above, namely summation by parts, the orthogonality relations (29a) and (29b), as well as (30) and (31). For the signal term, we apply Lemma 4.4 of Altmeyer and Bibinger (2014), which states that in our asymptotic framework we can, without loss of generality, consider the signal terms as stemming from synchronous observations. Though this Lemma directly follows from a basic approximation of $\Phi_{jk}(t) - \Phi_{jk}(s)$ as $(t-s)$ gets small, similarly as employed above, this is a main simplification of the analysis. Then, we obtain with $\mathcal{X} = \sum_{i=1}^n \Delta_i X \Phi_{jk}(\bar{t}_i)$ from some synchronous reference observation scheme $t_i, i = 0,\dots,n$, that

$$\mathbb{COV}\Big(\operatorname{vec}\big(\pi^2 j^2 h_n^{-2} \mathcal{X}\mathcal{X}^\top\big)\Big) = \frac{1}{n}\Big(\Sigma_{kh_n}^{\otimes 2}\mathcal{Z}\Big),$$

with (29a) and by Itô isometry. As the noise level (5) is diagonal, we can restrict ourselves to variances

$$\mathbb{V}\mathrm{ar}\left(\Big(\pi^2 j^2 h_n^{-2} \sum_i \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \tfrac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\Big)^2\right) = \pi^4 j^4 h_n^{-4} 2\Big(\big(\mathbf{H}_k^n\big)^{(pp)}\Big)^2,$$

with $\mathbf{H}_k^n$ as defined in (9), as well as by Lemma 1, (29b) and (30). The multiplication with $\mathcal{Z}/2$ when evaluating the variance-covariance matrix of the vectorized diagonal matrix is performed by elementary matrix calculus. Using the same preliminaries again together with the independence of signal and noise, the form of cross terms is easily proved, such that we conclude (34). Considering the statistics

$$\zeta_k^n = n^{\beta/2} \sum_{k=\lfloor sh_n^{-1} \rfloor - K_n}^{\lfloor sh_n^{-1} \rfloor + K_n} (2K_n+1)^{-1} \sum_{j=1}^{J_n} W_j\big(\mathbf{H}_k^n, \Sigma_{kh_n}\big) \tilde{Z}_{jk}\,, \tag{35a}$$

$$\tilde{Z}_{jk} = \operatorname{vec}\left(\tilde{S}_{jk}\tilde{S}_{jk}^\top - \pi j^2 h_n^{-2}\mathbf{H}_k^n - \Sigma_{kh_n}\right), \tag{35b}$$

we obtain the following convergences:

$$\mathbb{E}\left[\zeta_k^n \middle| \mathcal{F}_{kh_n}\right] \xrightarrow{p} 0\,, \tag{35c}$$

$$\mathbb{E}\left[\left(\zeta_k^n(\zeta_k^n)^\top\right)\middle|\mathcal{F}_{kh_n}\right] \xrightarrow{p} 2\big(\Sigma \otimes \Sigma_H^{1/2} + \Sigma_H^{1/2} \otimes \Sigma\big)_s \mathcal{Z}\,, \tag{35d}$$

$$\mathbb{E}\left[\left(\zeta_k^n(\zeta_k^n)^\top \zeta_k^n(\zeta_k^n)^\top\right)\middle|\mathcal{F}_{kh_n}\right] \xrightarrow{p} 0\,. \tag{35e}$$

The above analysis for the variance implies (35d). The bias condition (35c) is readily obtained using Lemma 1, summation by parts, Itô isometry for the signal part, as well as (29a) and (29b). The Lyapunov condition (35e) follows using analogous bounds as for the stable CLT of the integrated version in Altmeyer and Bibinger (2014).
We are left to prove that $\alpha_n = n^{\beta/2}\operatorname{vec}\left(\hat{\Sigma}_s^{or}(\tilde{X}+\epsilon) - \Sigma_s\right)$ satisfy

$$\mathbb{E}\left[Zg(\alpha_n)\right] \to \mathbb{E}\left[Zg(\alpha)\right] = \mathbb{E}[Z]\mathbb{E}\left[g(\alpha)\right]\,, \tag{36}$$

for any $\mathcal{F}$-measurable bounded random variable Z and continuous bounded function g with

$$\alpha = \big(\Sigma^{1/2} \otimes \Sigma_H^{1/4}\big)_s \mathcal{Z}U + \big(\Sigma_H^{1/4} \otimes \Sigma^{1/2}\big)_s \mathcal{Z}U' \tag{37}$$

and $U, U' \in \mathbb{R}^{d^2}$ being two independent standard normally distributed vectors independent of $\mathcal{F}$. (36) ensures that the central limit theorem (33) is stable. The limit α gives indeed the

asymptotic law of (33) as $\mathcal{Z}^2 = 2\mathcal{Z}$ and

$$\begin{aligned}&\left(\Sigma^{1/2} \otimes \Sigma_H^{1/4}\right)_s \mathcal{Z} \left(\left(\Sigma^{1/2} \otimes \Sigma_H^{1/4}\right)_s \mathcal{Z}\right)^\top \\ &\quad = 2\left(\Sigma \otimes \Sigma_H^{1/2}\right)_s \mathcal{Z}\,,\end{aligned}$$

because $\mathcal{Z}$ commutes with $\left(\Sigma^{1/2} \otimes \Sigma_H^{1/4}\right)_s$ and by the analogous transform for the second addend.

The proof of (36) relies on separating the sequence of intervals with the blocks involved in $\hat{\Sigma}_s^{or}$ from the rest of $[0,1]$ and conditioning. Thereto, set

$$\begin{aligned}A_n &= [s - (K_n+1)h_n, s + (K_n+1)h_n]\,,\\ \tilde{X}(n)_t = \int_0^t \mathbb{1}_{A_n}(s)\sigma_{\lfloor sh_n^{-1}\rfloor h_n}\, dB_s\,, \bar{X}(n)_t &= X_t - \tilde{X}(n)_t\,.\end{aligned}$$

Denote with $\mathcal{H}_n$ the σ-field generated by $\bar{X}(n)_t$ and $\mathcal{F}_0$. Then, $\left(\mathcal{H}_n\right)_{n\in\mathbb{N}}$ is an isotonic sequence and $\bigcup_{n\in\mathbb{N}} \mathcal{H}_n = \mathcal{F}_1$. Since $\mathbb{E}[Z|\mathcal{H}_n] \to Z$ in $L^1(\mathbb{P})$ as $n \to \infty$, it is enough to show that $\mathbb{E}[Zg(\alpha_n)] \to \mathbb{E}[Z]\mathbb{E}[g(\alpha)]$ for Z being $\mathcal{H}_{n_0}$-measurable for some $n_0 \in \mathbb{N}$. Observe that α_n includes only increments $\Delta_i \tilde{X}^{(p)}, p = 1,\dots,d$, of $\tilde{X}(n)_t$ and uncorrelated from those of $\bar{X}(n)_t$. For all $n \geq n_0$, we conclude $\mathbb{E}[Zg(\alpha_n)] = \mathbb{E}[Z]\mathbb{E}[g(\alpha_n)] \to \mathbb{E}[Z]\mathbb{E}[g(\alpha)]$ by a standard central limit theorem. This proves (36) and completes the proof of Proposition A.1. □

Proposition A.2. *On the assumptions of Theorem 1, for any $s \in (0,1)$ it holds true that*

$$n^{\beta/2}\,\mathrm{vec}\left(\hat{\Sigma}_s^{or} - \hat{\Sigma}_s^{or}(\tilde{X}+\epsilon)\right) \xrightarrow{p} 0\,. \tag{38}$$

Proof. The left-hand side above equals

$$\begin{aligned}n^{\beta/2}\,\mathrm{vec}\left(\hat{\Sigma}_s^{or} - \Sigma_s^{or}(\tilde{X}+\epsilon)\right) = n^{\beta/2}\left(2K_n+1\right)^{-1} &\sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} \sum_{j=1}^{J_n} W_j\left(\mathbf{H}_k^n, \Sigma_{kh_n}\right)\\ &\times \mathrm{vec}\left(S_{jk}S_{jk}^\top - \tilde{S}_{jk}\tilde{S}_{jk}^\top\right).\end{aligned}$$

Using that

$$\begin{aligned}\left\|S_{jk}S_{jk}^\top - \tilde{S}_{jk}\tilde{S}_{jk}^\top\right\| &= \left\|\tilde{S}_{jk}\left(S_{jk}^\top - \tilde{S}_{jk}^\top\right) + \left(S_{jk} - \tilde{S}_{jk}S_{jk}^\top\right\|\right.\\ &\leq \left(\|S_{jk}\| + \|\tilde{S}_{jk}\|\right)\left\|S_{jk} - \tilde{S}_{jk}\right\|,\end{aligned}$$

and the order of the weights provided by (27), a crude estimate suffices here:

$$\begin{aligned}
&\Big\| n^{\beta/2} \operatorname{vec}\big(\hat{\Sigma}_s^{or} - \Sigma_s^{or}(\tilde{X} + \epsilon)\big)\Big\| \\
&\lesssim n^{\beta/2}\big(2K_n+1\big)^{-1} \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor - K_n} \sum_{j=1}^{J_n} \big\|W_j\big(\mathbf{H}_k^n, \Sigma_{kh_n}\big)\big\|\big(\|S_{jk}\| + \|\tilde{S}_{jk}\|\big)\big\|S_{jk} - \tilde{S}_{jk}\big\| \\
&= \mathcal{O}_p\left(n^{\beta/2}\sum_{j=1}^{J_n}\big(1 + j^2(nh_n^2)^{-1}\big)^{-1}(\log(n))^{-1}\right) = \mathcal{O}_p\big(n^{\beta/2}h_n^{\alpha}\log(n)\big) = o_p(1)\,.
\end{aligned}$$

This follows with

$$\sum_{j=1}^{J_n}\big(1 \wedge j^{-2}nh_n^2\big) \lesssim \sum_{j=1}^{\sqrt{n}h_n} 1 + \sum_{j=1}^{J_n} j^{-2}nh_n^2 \lesssim \log^2(n)\,,$$

and with $\|\mathbb{COV}(S_{jk})\| = \mathcal{O}\big(\big(1 + j^2(nh_n^2)^{-1}\big)^2\big)$ and $\|\mathbb{E}[S_{jk}]\| = \mathcal{O}(1)$. □

Proposition A.3. *On the assumptions of Theorem 1, for any $s \in (0,1)$ it holds true that*

$$n^{\beta/2}\big(\operatorname{vec}\big(\hat{\Sigma}_s - \hat{\Sigma}_s^{or}\big)\big) \xrightarrow{p} 0\,. \tag{39}$$

Proof. We write the left-hand side above

$$\begin{aligned}
n^{\beta/2}\big(\operatorname{vec}\big(\hat{\Sigma}_s - \hat{\Sigma}_s^{or}\big)\big) = n^{\beta/2}\big(2K_n+1\big)^{-1} \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor - K_n} \sum_{j=1}^{J_n} \Big(\hat{W}_j\big(\hat{\Sigma}_{kh_n}\big) \\
- W_j\big(\Sigma_{kh_n}\big)\Big) Z_{jk},
\end{aligned}$$

$$Z_{jk} = \operatorname{vec}\Big(S_{jk}S_{jk}^{\top} - \pi j^2 h_n^{-2}\mathbf{H}_k^n - \Sigma_{kh_n}\Big)\,.$$

By Proposition A.2, the expectation of Z_{jk} is asymptotically negligible. To extend our asymptotic theory to an adaptive approach, we shall concentrate in the following on the error due to pre-estimating Σ_{kh_n} to determine the optimal weights in (13). The effect of an estimated noise level as in (11) is easily shown to be asymptotically negligible using Theorem 2 and (31). We

decompose the difference between adaptive and oracle estimator using

$$\begin{aligned}\hat{W}_j\big(\hat{\Sigma}_{kh_n}\big) - W_j\big(\Sigma_{kh_n}\big) = \hat{W}_j\big(\hat{\Sigma}_{m(2K_n+1)h_n}\big) - W_j\big(\Sigma_{m(2K_n+1)h_n}\big) + \\ + \hat{W}_j\big(\hat{\Sigma}_{kh_n}\big) - \hat{W}_j\big(\hat{\Sigma}_{m(2K_n+1)h_n}\big) \\ + W_j\big(\Sigma_{m(2K_n+1)h_n}\big) - W_j\big(\Sigma_{kh_n}\big)\,,\end{aligned}$$

where $\hat{\Sigma}_{m(2K_n+1)h_n}, m = 0, \dots, \lfloor (2K_n+1)^{-1}h_n^{-1}\rfloor - 1$, is a pre-estimator, which is constant on the coarse grid, such that it is constant over the smoothing window of $\hat{\Sigma}_s$ from (12), and $\Sigma_{m(2K_n+1)h_n}$ is the locally constantly approximated true covariance matrix on the same coarse grid. We apply triangular inequality and prove that all three terms tend to zero in probability. For the first term, we obtain

$$\begin{aligned}&\left\| \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor - K_n} (2K_n+1)^{-1} \sum_{j=1}^{J_n} \Big(\hat{W}_j\big(\hat{\Sigma}_{m(2K_n+1)h_n}\big) - W_j\big(\Sigma_{m(2K_n+1)h_n}\big)\Big) Z_{jk} \right\| \\ &\quad \le (2K_n+1)^{-1} \sum_{j=1}^{J_n} \left\| \hat{W}_j\big(\hat{\Sigma}_{m(2K_n+1)h_n}\big) - W_j\big(\Sigma_{m(2K_n+1)h_n}\big)\right\| \left\| \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor - K_n} Z_{jk} \right\| \\ &= \mathcal{O}_p\left(K_n^{-1/2} \sum_{j=1}^{J_n} \delta_n \log(n) \big(1 + j^2 (nh_n^2)^{-1}\big)\big(1 \vee j^{-4} n^2 h_n^4\big)\right) = o\big(n^{-\beta/2}\big)\,,\end{aligned}$$

as the weight matrices in this term do not depend on k if $\|\hat{\Sigma} - \Sigma\| = \mathcal{O}_p(\delta_n)$ is the rate of the pre-estimator on the coarse grid and by (27), as well as Lemma C.2 in Bibinger et al. (2014). The latter is a key ingredient of this proof as it gives a uniform upper bound on the norm of the matrix derivatives of $W_j(\Sigma)$ w.r.t. Σ, such that we can use the Δ-method. Actually some rate $\delta_n = n^{-\varepsilon}, \varepsilon > 0$, suffices here and we can ensure a much faster rate.

Hence, it remains to show that the two other terms are negligible, as well. Since all weight matrices satisfy $\sum_j W_j = E_{d^2\times d^2}$, i.e. their sum equals the identity matrix, we consider the

sum of the norms of the variance-covariance matrices of those terms, which is bounded by:

$$
\begin{aligned}
&(2K_n+1)^{-2}\sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor-K_n}\left\|\sum_{j=1}^{J_n}\mathbb{COV}\left(\left(W_j\big(\Sigma_{m(2K_n+1)h_n}\big)-W_j\big(\Sigma_{kh_n}\big)\right)Z_{jk}\right)\right\| \\
&\quad+(2K_n+1)^{-2}\sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor-K_n}\left\|\sum_{j=1}^{J_n}\mathbb{COV}\left(\left(\hat{W}_j\big(\hat{\Sigma}_{m(2K_n+1)h_n}\big)-\hat{W}_j\big(\hat{\Sigma}_{kh_n}\big)\right)Z_{jk}\right)\right\| \\
&\le(2K_n+1)^{-2}\sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor-K_n}\left(\sum_{j=1}^{J_n}\left(\left\|W_j\big(\Sigma_{m(2K_n+1)h_n}\big)-W_j\big(\Sigma_{kh_n}\big)\right\|^2\right.\right. \\
&\qquad\qquad\left.\left.\times\left\|\mathbb{COV}\big(Z_{jk}\big)\right\|\right)^{1/2}\right)^2 \\
&\quad+(2K_n+1)^{-2}\sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor-K_n}\left(\sum_{j=1}^{J_n}\left(\left\|\hat{W}_j\big(\hat{\Sigma}_{m(2K_n+1)h_n}\big)-\hat{W}_j\big(\hat{\Sigma}_{kh_n}\big)\right\|^2\right.\right. \\
&\qquad\qquad\left.\left.\times\left\|\mathbb{COV}\big(Z_{jk}\big)\right\|\right)^{1/2}\right)^2 \\
&=\mathcal{O}_p\left(K_n^{-1}\left(\sum_{j=1}^{J_n}\big(1\vee j^{-4}n^2h_n^4\big)\big(1\wedge j^2(nh_n^2)^{-1}\big)\right)^2\Big(\delta_n^2\vee(K_nh_n)^{-2\alpha}\Big)\right)=o_p(n^{-\beta}).
\end{aligned}
$$

This completes the proof of Theorem 1. The feasible version (17a) readily follows from the induced consistency of $\hat{I}_k^{-1}\big(\hat{\Sigma}_{kh_n}\big)$ for $I_k^{-1}\big(\Sigma_{kh_n}\big)$. □

A.3 Proof of Theorem 2

Observe that

$$
\begin{aligned}
\mathbb{E}\left[\Delta_iY\Delta_{i+r}Y\right]&=\mathbb{E}\left[\Delta_i\epsilon\Delta_{i+r}\epsilon\right]+\mathcal{o}(1)\\
&=\mathbb{E}\left[\epsilon_i\epsilon_{i+r}-\epsilon_{i-1}\epsilon_{i+r}-\epsilon_i\epsilon_{i+r-1}+\epsilon_{i-1}\epsilon_{i+r-1}\right]\\
&=2\eta_r-\eta_{r-1}-\eta_{r+1}=\gamma_r,
\end{aligned}
$$

for $r\ge1, 1\le i\le(n-r)$, and

$$
\gamma_0=2(\eta_0-\eta_1)=\mathbb{E}\left[(\Delta_iY)^2\right]+\mathcal{o}(1)\,.
$$

The remainders stem from the signal terms X, which are of smaller order. Hence, by the definition of the estimators in (19b), it readily follows that

$$\mathbb{E}\left[\hat{\eta}_R\right] = \eta_0 - \eta_1 + \sum_{u=1}^{R} \left(2\eta_u - \eta_{u-1} - \eta_{u+1}\right) + \mathcal{O}(1) = \eta_R + \mathcal{O}(1),$$

and we have consistency of all $\hat{\eta}_r, 0 \leq r \leq R$. For the analysis of the variances of the estimators (19a)-(19c), denote by

$$\Gamma^{rr'}_{|i-k|} = \mathbb{C}\mathrm{ov}\left(\Delta_i \epsilon \Delta_{i+r}\epsilon, \Delta_k \epsilon \Delta_{k+r'}\epsilon\right) .$$

The variance of $\hat{\eta}_R$ for the maximum lag $R > 1$, for which $\eta_j = 0$ for $j > R$, becomes

$$\begin{aligned}
\mathbb{V}\mathrm{ar}(\hat{\eta}_R) &= (2n)^{-2} \sum_{i=1}^{n} \mathbb{V}\mathrm{ar}\big((\Delta_i Y)^2\big) + (2n)^{-2} \sum_{i=1}^{n-R} 2 \sum_{u=1}^{R} \mathbb{C}\mathrm{ov}\big((\Delta_i Y)^2, (\Delta_{i+u} Y)^2\big) \\
&\quad + n^{-2} \sum_{i=1}^{n-R} \sum_{r=1}^{R} \sum_{r'=1}^{R} \Big(\mathbb{C}\mathrm{ov}\big(\Delta_i Y \Delta_{i+r} Y, \Delta_i Y \Delta_{i+r'} Y\big) \\
&\qquad\qquad + 2 \sum_{u=1}^{R} \mathbb{C}\mathrm{ov}\big(\Delta_i Y \Delta_{i+r} Y, \Delta_{i+u} Y \Delta_{i+u+r'} Y\big)\Big) \\
&\quad + n^{-2} \sum_{i=1}^{n-R} \sum_{r=1}^{R} \Big(\mathbb{C}\mathrm{ov}\big((\Delta_i Y)^2, \Delta_i Y \Delta_{i+r} Y\big) \\
&\qquad\qquad + 2 \sum_{u=1}^{R} \mathbb{C}\mathrm{ov}\big((\Delta_i Y)^2, \Delta_{i+u} Y \Delta_{i+u+r} Y\big)\Big) \\
&= n^{-1} \left(\tfrac{1}{4}\Gamma_0^{00} + \tfrac{1}{2} \sum_{u=1}^{R} \Gamma_u^{00} + \sum_{r=1}^{R} \sum_{r'=1}^{R} \Big(\Gamma_0^{rr'} + 2 \sum_{u=1}^{R} \Gamma_u^{rr'}\Big) \right. \\
&\qquad\qquad \left. + \sum_{r=1}^{R} \Big(\Gamma_0^{0r} + 2 \sum_{u=1}^{R} \Gamma_u^{0r}\Big)\right) + \mathcal{O}\big(n^{-1}\big) \\
&= n^{-1} \left(\tfrac{1}{4}\Big(\Gamma_0^{00} + 2 \sum_{u=1}^{R} \Gamma_u^{00}\Big) + \sum_{r=0}^{R} \sum_{r'=1}^{R} \Big(\Gamma_0^{rr'} + 2 \sum_{u=1}^{R} \Gamma_u^{rr'}\Big)\right) + \mathcal{O}\big(n^{-1}\big) .
\end{aligned}$$

Write the above variance as $n^{-1}\mathcal{V}_R^n$. By construction of $\eta_r, 0 \leq r \leq R-1$, we derive

$$n\mathbb{V}\mathrm{ar}(\hat{\eta}_r) + \mathcal{O}(1) = \mathcal{V}_{r+1}^n + \mathcal{V}_r^n + 2\mathcal{C}_{r,r+1}^n,$$

where $\mathcal{V}_r^n$ are defined analogously to $\mathcal{V}_R^n$ except replacing R by $r < R$ and

$$\mathcal{C}_{r,r+1}^n = \left(\frac{\Gamma_0^{00}}{4} + \frac{1}{2} \sum_{u=1}^{R} \Gamma_u^{00} + \sum_{u=0}^{r} \sum_{u'=1}^{r+1} \left(\Gamma_0^{uu'} + 2 \sum_{q=1}^{R} \Gamma_q^{uu'} \right) \right).$$

Inserting the observed returns $\Delta_i Y$ as estimators of the noise increments $\Delta_i \epsilon$ gives consistent estimators of the variances. Sufficient conditions for a central limit theorem can easily be shown here by applying, for example, Theorem 27.4 from Billingsley (1991).

B Stochastic Volatility Specifications for Simulation Study

Following Huang and Tauchen (2005), we consider one- and two-factor stochastic volatility models for the efficient log-price process in (24a). The one-factor specification reads

$$\tilde{\sigma}_t = \exp(\beta_0 + \beta_1 v_t), \quad dv_t = \alpha v_t dt + dW_t, \tag{40}$$

where W_t is a standard Brownian motion with $\mathbb{C}\text{orr}(dW_t, dB_t) = \rho$. The parameter values are chosen as $\beta_0 = 0$, $\beta_1 = 0.125$, $\alpha = -0.025$ and $\rho = -0.62$.

The two-factor model introduced by Chernov et al. (2003) allows for more pronounced movements in the instantaneous volatility by a feedback mechanism. The corresponding parameterization is

$$\begin{aligned} \tilde{\sigma}_t &= \text{s–exp}(\beta_0 + \beta_1 v_{1,t} + \beta_2 v_{2,t}), \\ dv_{1,t} = \alpha_1 v_{1,t} dt + dW_{1,t}, &\quad dv_{2,t} = \alpha_2 v_{2,t} dt + (1 + \beta_v v_{2,t})\, dW_{2,t}, \\ \text{s–exp}(u) &= \begin{cases} \exp(u) & \text{if} \quad u \le u_0 \\ \exp(u_0)\sqrt{1 - u_0 + u^2/u_0} & \text{else,} \end{cases} \end{aligned} \tag{41}$$

where $W_{1,t}$ and $W_{2,t}$ are standard Brownian motions with $\mathbb{C}\text{orr}(dW_{1,t}, dB_t) = \rho_1$ and $\mathbb{C}\text{orr}(dW_{2,t}, dB_t) = \rho_2$. We consider the configuration $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.137e^{-2}$, $\alpha_2 = -1.386$, $\beta_v = 0.25$, $\rho_1 = \rho_2 = -0.3$ and $u_0 = \ln(1.5)$.

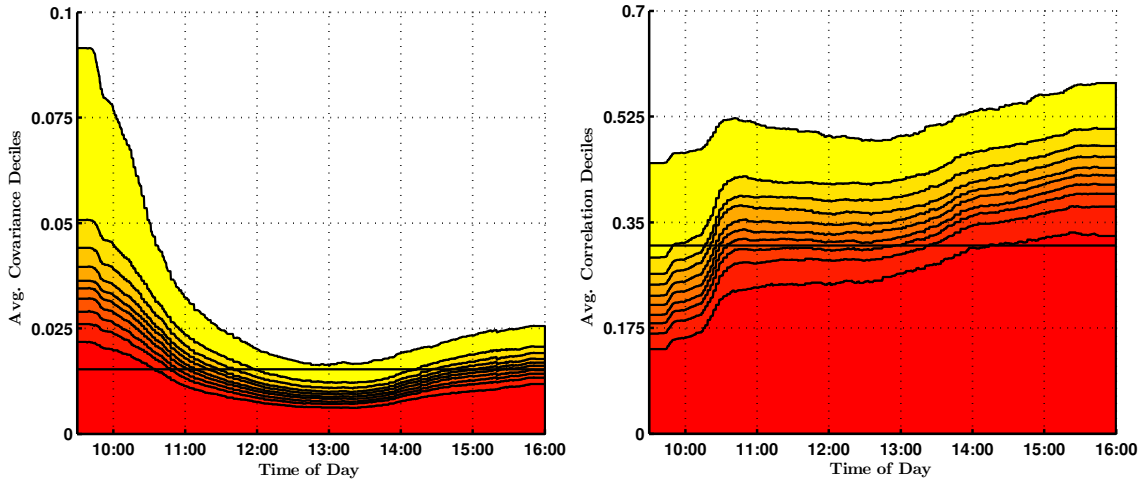
C Tables and Figures

Table 3: Summary statistics for NASDAQ quote data. $\overline{n}$: avg. # of observations. $\overline{\Delta t}$: avg. duration in seconds between observations. $\%_{(|\Delta Y|>0)}$: % of observations associated with price changes. $\overline{\Delta t_{(|\Delta Y|>0)}}$: avg. duration in seconds between price changes. $\overline{\sqrt{\hat{\eta}^*}}$: $(10^6\times)$ avg. of square root of long-run noise variance estimate for quote *revisions* based on $\tilde{Q} = 50$. $\overline{\hat{\xi}^*}$: avg. noise-to-signal ratio per observation, where $\hat{\xi}^{2*} = n\hat{\eta}^*/\text{RV}^{ss}_{5m}$ with RV^{ss}_{5m} denoting the sub-sampled five-minute realized variance. $\overline{\hat{R}^*}$: avg. estimate of order of serial dependence in the noise process.

Symbol	$\overline{n}$	$\overline{\Delta t}$	$\%_{(\lvert\Delta Y\rvert>0)}$	$\overline{\Delta t_{(\lvert\Delta Y\rvert>0)}}$	$\overline{\sqrt{\hat{\eta}^*}}$	$\overline{\hat{\xi}^*}$	$\overline{\hat{R}^*}$
AAPL	176608.121	0.184	49.720	0.366	0.005	0.344	12.083
MSFT	440650.911	0.063	0.759	9.227	0.098	0.748	2.734
GOOG	59790.343	0.466	50.918	0.937	0.022	0.656	8.877
CSCO	293862.329	0.092	0.628	16.566	0.055	0.427	1.507
ORCL	320960.857	0.089	1.390	6.971	0.203	1.151	3.713
INTC	350144.419	0.079	0.681	12.739	0.086	0.526	1.976
QCOM	321429.205	0.087	4.064	2.533	0.209	1.950	7.604
AMGN	134118.345	0.245	13.509	2.389	0.215	2.038	6.461
TEVA	107736.535	0.312	7.557	4.558	0.216	1.701	4.507
GILD	180836.653	0.155	6.696	3.263	0.272	1.950	6.729
AMZN	97905.278	0.333	47.455	0.692	0.033	0.591	10.475
EBAY	271235.321	0.101	3.336	3.731	0.362	1.789	6.286
NWSA	198976.385	0.158	1.176	14.179	0.212	0.798	1.919
DTV	156096.381	0.201	6.187	3.904	0.209	1.739	4.930
CELG	76910.909	0.377	24.930	1.715	0.202	1.719	7.219
INFY	59641.755	0.594	22.721	2.828	0.165	1.966	3.088
YHOO	250724.521	0.133	1.061	15.243	0.200	0.671	2.455
COST	72083.358	0.412	21.294	2.172	0.106	1.963	5.998
SPLS	137575.532	0.201	1.184	19.160	0.122	0.486	1.373
SBUX	169759.285	0.167	6.190	3.527	0.227	1.578	5.748
ADP	108426.889	0.291	7.240	4.707	0.204	2.258	4.586
ESRX	128535.116	0.251	11.346	2.418	0.229	1.915	6.195
ADBE	161843.232	0.174	4.353	5.087	0.238	1.391	3.779
FSLR	73666.222	0.428	31.716	1.593	0.817	1.743	7.862
BIIB	51072.535	0.579	37.997	1.664	0.229	2.241	5.594
SYMC	152508.502	0.182	1.212	16.638	0.127	0.598	1.462
JNPR	160994.910	0.205	4.310	6.507	0.545	1.246	3.436
BRCM	239822.793	0.119	3.563	4.193	0.473	1.965	6.293
AMAT	155696.834	0.174	0.775	24.924	0.093	0.397	0.968
CMCSA	266752.353	0.106	1.558	7.801	0.144	1.002	3.194
QQQ	1063732.448	0.025	0.784	3.948	0.059	1.097	4.596

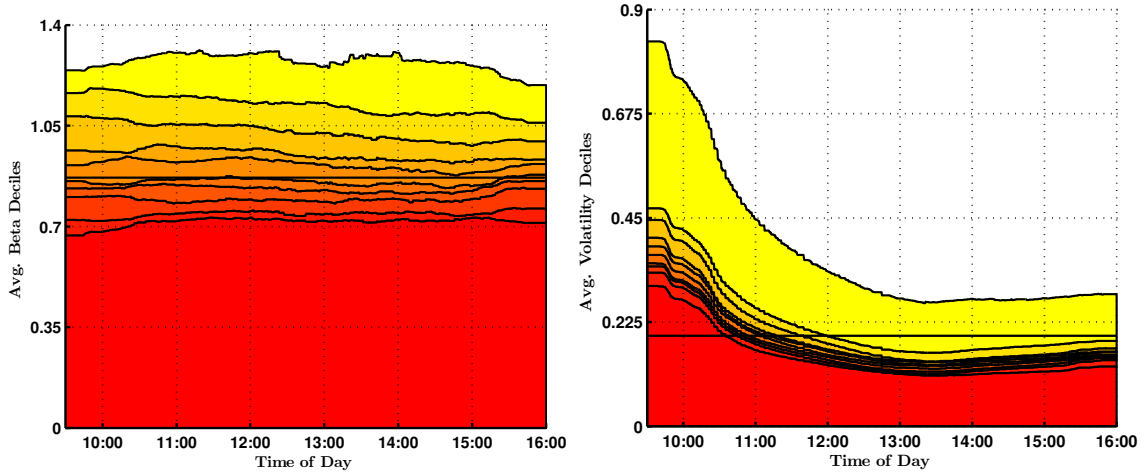
Table 4: Summary statistics of number of blocks, spectral cut-off and length of smoothing window for LMM estimator. No. of blocks $\lceil h_n^{-1} \rceil$, spectral cut-off J_n and length of smoothing window K_n (in blocks) are chosen as described in Section 2.4, using the input parameters that are optimal in the "1F"-setting given a high noise level in the simulation study of Section 3: $\theta_h = 0.2$, $\theta_J = 8$ and $\theta_K = 0.4$.

Sample	Input	$q_{0.05}$	Mean	$q_{0.95}$	Std.
05/10	$\lceil h_n^{-1} \rceil$	18.000	22.516	29.000	3.922
–	J_n	48.000	53.532	60.000	3.672
04/14	K_n	2.000	2.435	3.000	0.300
05/10	$\lceil h_n^{-1} \rceil$	18.000	22.744	30.000	4.362
–	J_n	48.000	53.711	60.000	3.928
04/11	K_n	2.000	2.446	3.000	0.323
05/11	$\lceil h_n^{-1} \rceil$	18.650	23.761	32.000	4.663
–	J_n	49.000	54.708	61.350	4.072
04/12	K_n	2.000	2.516	3.000	0.330
05/12	$\lceil h_n^{-1} \rceil$	17.000	20.495	26.000	2.587
–	J_n	47.000	51.450	57.000	2.883
04/13	K_n	2.000	2.280	2.500	0.262
05/13	$\lceil h_n^{-1} \rceil$	19.000	22.902	27.000	2.532
–	J_n	50.000	54.127	58.000	2.529
04/14	K_n	2.000	2.488	3.000	0.194



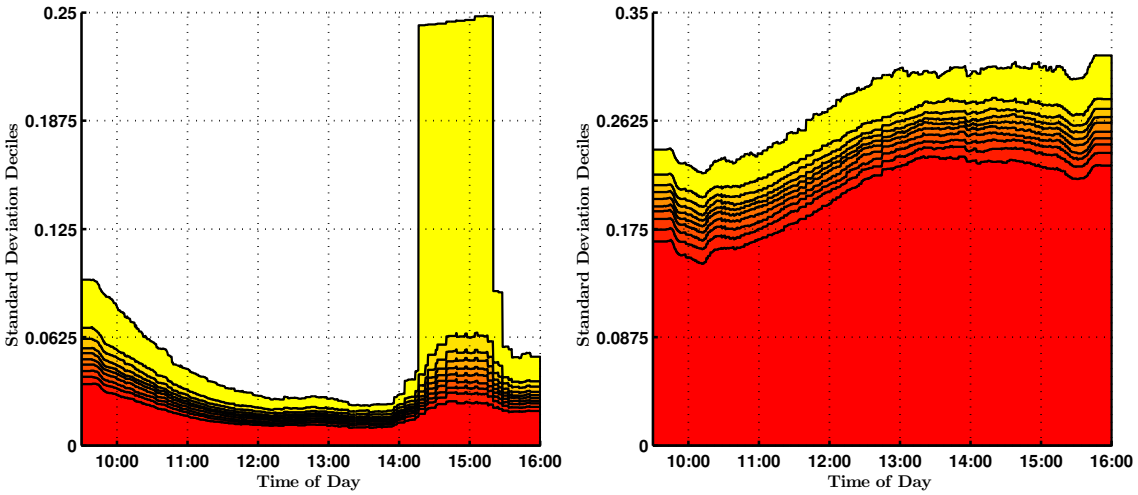
(a) Covariances (b) Correlations

Figure 2: Cross-sectional deciles of across-day averages of spot covariances and correlations. Spot estimates are first averaged across days for each asset pair. Subsequently, cross-sectional sample deciles of the across-day averages are computed. Solid horizontal line corresponds to the cross-sectional median of the across-day averages of *integrated* covariance and correlation estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Covariances are annualized.



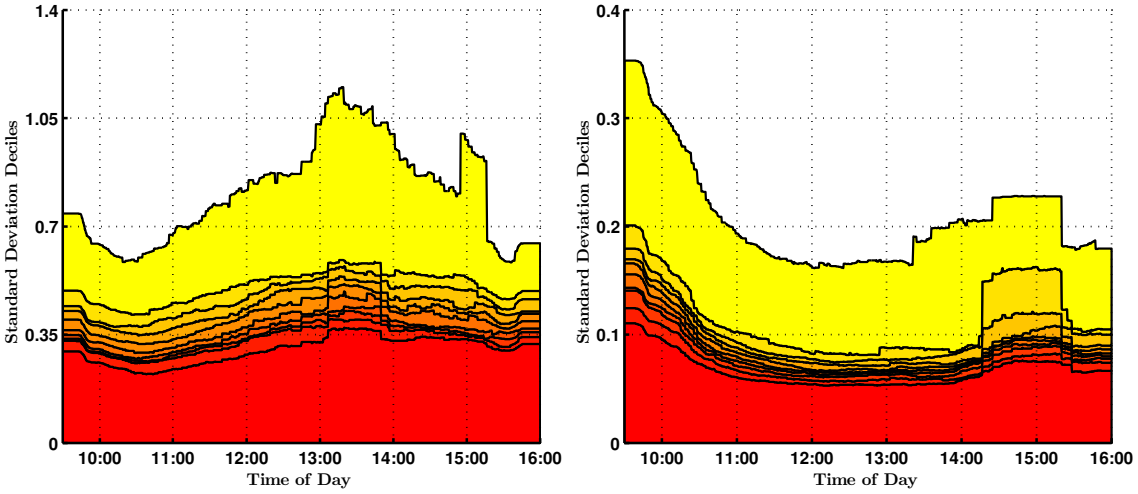
(a) Betas (b) Volatilities

Figure 3: Cross-sectional deciles of across-day averages of spot betas and volatilities. Spot estimates are first averaged across days for each asset. Subsequently, cross-sectional sample deciles of the across-day averages are computed. Solid horizontal line corresponds to the cross-sectional median of the across-day averages of *integrated* beta and volatility estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Volatilities are annualized.



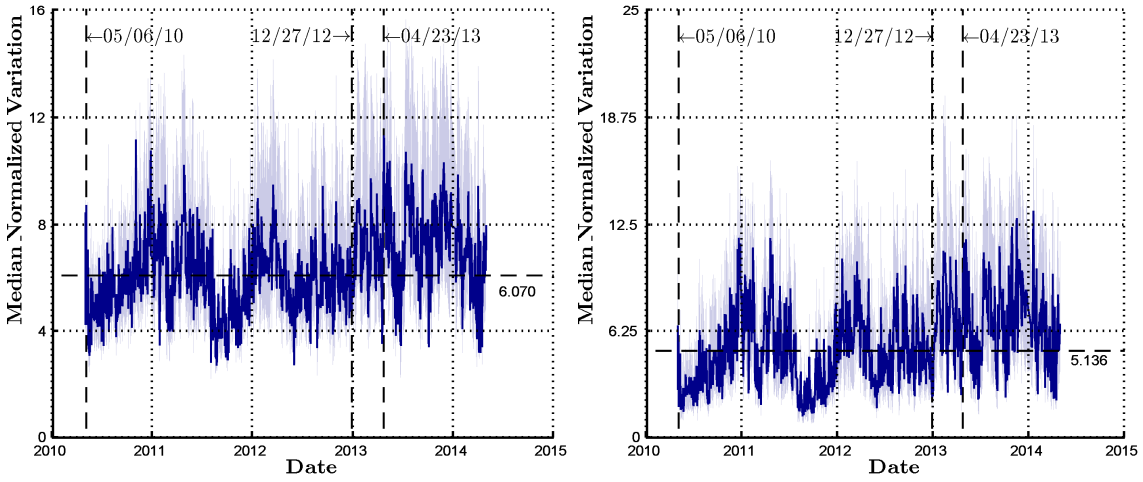
(a) Covariances (b) Correlations

Figure 4: Cross-sectional deciles of across-day standard deviations of spot covariances and correlations. First, sample standard deviations of spot estimates are computed across days for each asset pair. Subsequently, cross-sectional sample deciles of the across-day standard deviations are computed. Covariances are annualized.



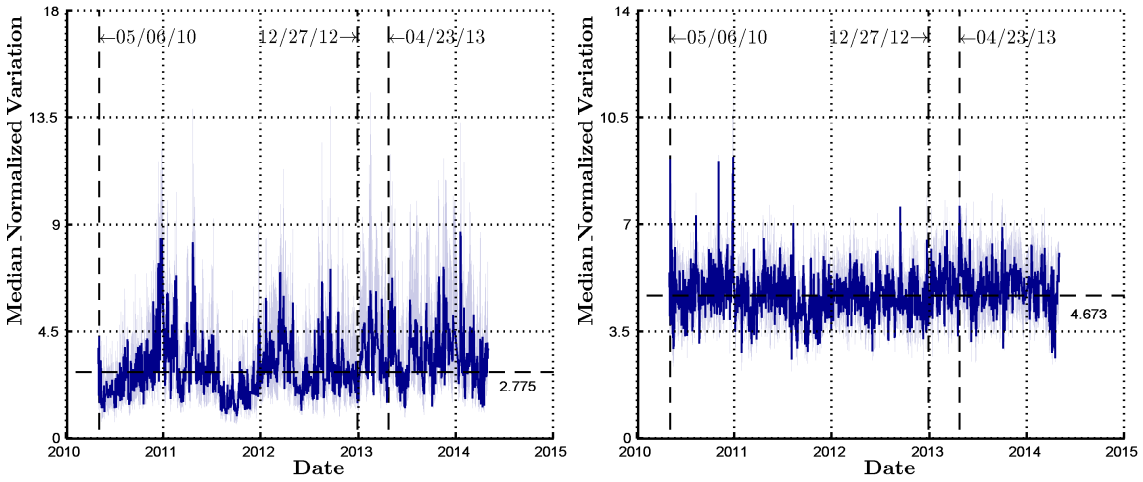
(a) Betas (b) Volatilities

Figure 5: Cross-sectional deciles of across-day standard deviations of spot betas and volatilities. First, sample standard deviations of spot estimates are computed across days for each asset. Subsequently, cross-sectional sample deciles of the across-day standard deviations are computed. Volatilities are annualized.



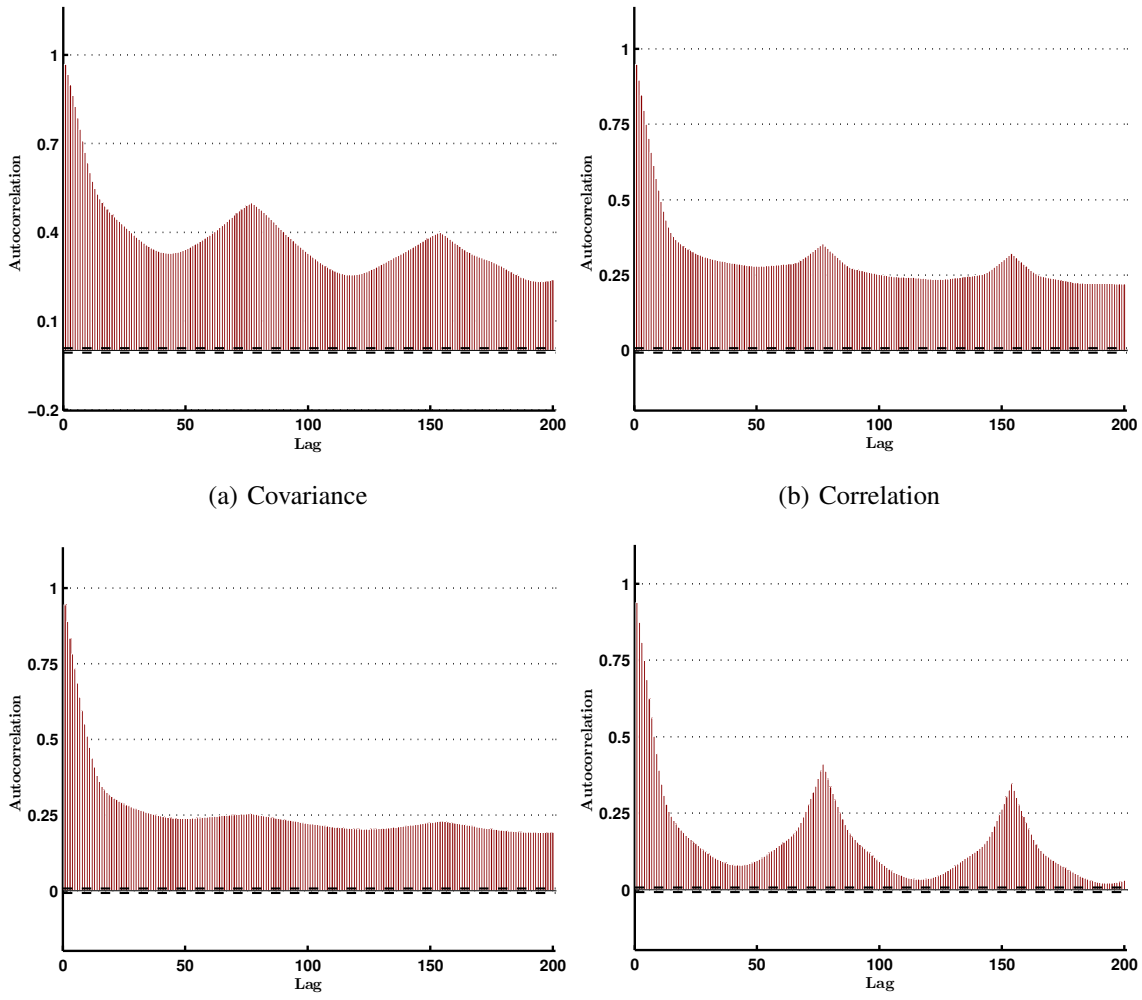
(a) Covariances (b) Correlations

Figure 6: Cross-sectional medians of normalized intraday variation proxy for spot covariances and correlations. (Total) intraday variation normalized by the L_1-norm is proxied by $\tilde{V}_f^{\text{norm}} = \sum_{i=1}^{n_g} |f(t_i) - f(t_{i-1})| \left[\sum_{i=1}^{n_g} |f(t_i)| \Delta t_i\right]^{-1}$, where n_g is the number of grid points used. Lower and upper boundary of shaded area correspond to cross-sectional 10% and 90% percentiles, respectively. The horizontal dashed line corresponds to the across-day median. The vertical broken lines indicate three selected "event days".



(a) Betas (b) Volatilities

Figure 7: Cross-sectional medians of normalized intraday variation proxy for spot betas and volatilities. (Total) intraday variation normalized by the L_1-norm is proxied by $\tilde{V}_f^{\text{norm}} = \sum_{i=1}^{n_g} |f(t_i) - f(t_{i-1})| \left[\sum_{i=1}^{n_g} |f(t_i)| \Delta t_i\right]^{-1}$, where n_g is the number of grid points used. Lower and upper boundary of shaded area correspond to cross-sectional 10% and 90% percentiles, respectively. The horizontal dashed line corresponds to the across-day median. The vertical broken lines indicate three selected "event days".



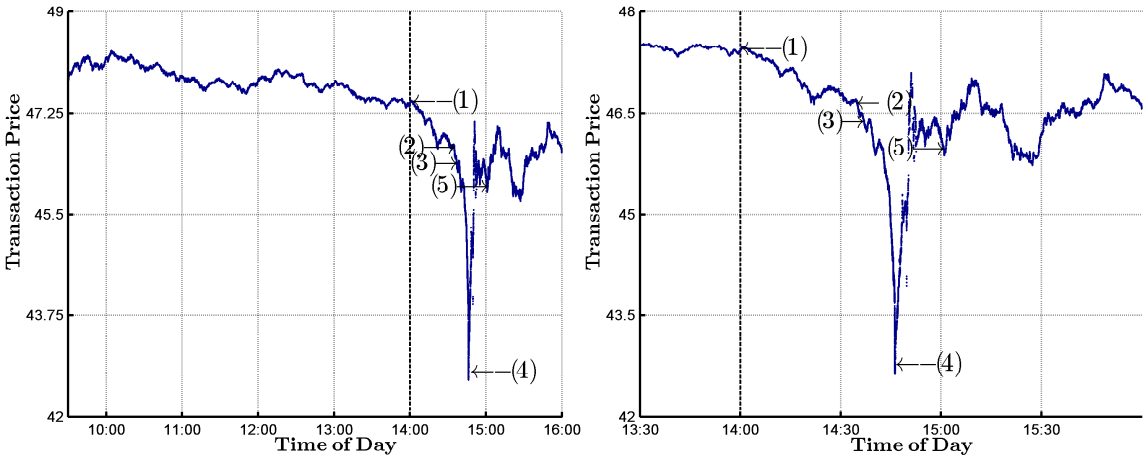
(a) Covariance

(b) Correlation

(c) Beta

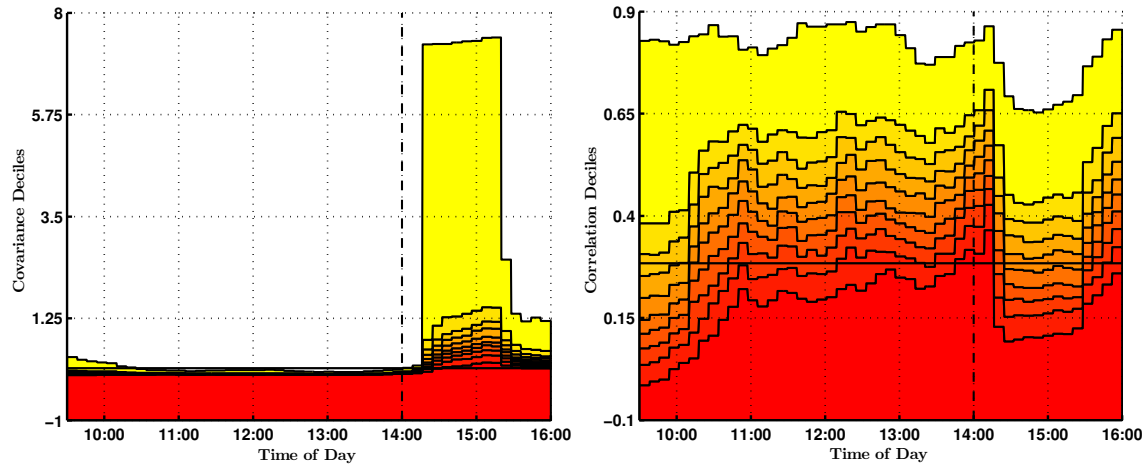
(d) Volatility

Figure 8: Avg. ACFs of spot covariance, correlation, beta and volatility estimates with one lag representing approximately five minutes. ACFs with corresponding confidence intervals are first computed for each asset or asset pair and subsequently averaged across all assets or pairs. Dashed lines correspond to cross-sectional averages of point-wise 95% confidence intervals ($\pm 1.96/\sqrt{n}$).



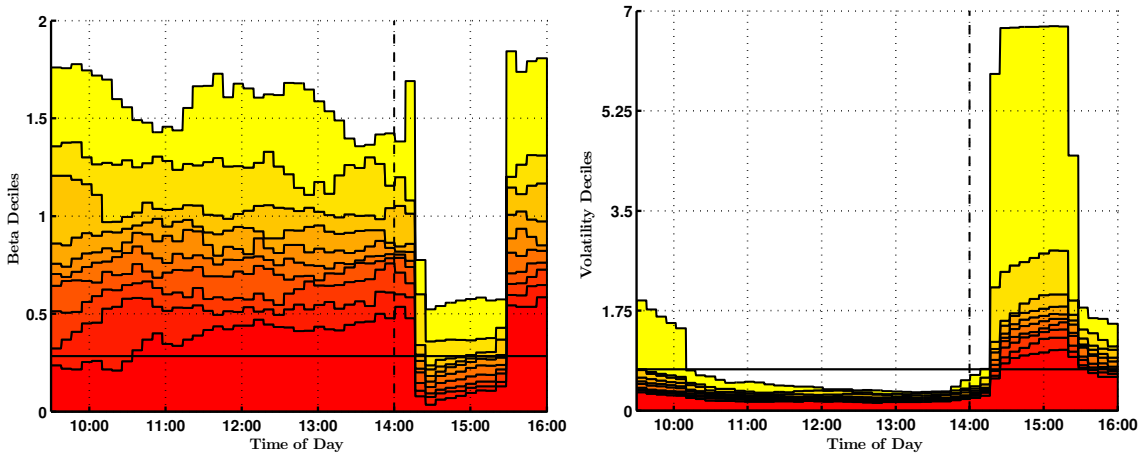
(a) Entire trading day (b) 1:30 pm – 4:00 pm

Figure 9: QQQ transaction prices (05/06/10). (1): Protests in Athens trigger Euro down movement vs. Yen; U.S. fund managers short-sell E-Mini contracts in vast amounts. (2): E-Mini market makers cut back trading. (3): NASDAQ stops order routing to ARCA. (4): Rumors suggesting that decline occurred due to "fat-finger" error, and not bad news. (5): NASDAQ resumes routing to ARCA.



(a) Covariances (b) Correlations

Figure 10: Cross-sectional deciles of spot covariances and correlations (05/06/10). Solid horizontal line corresponds to the cross-sectional median of *integrated* covariance and correlation estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Covariances are annualized.



(a) Betas (b) Volatilities

Figure 11: Cross-sectional deciles of spot betas and volatilities (05/06/10). Solid horizontal line corresponds to the cross-sectional median of *integrated* beta and volatility estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Volatilities are annualized.

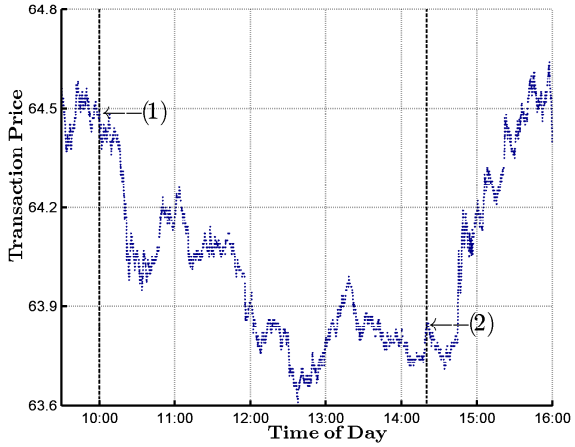
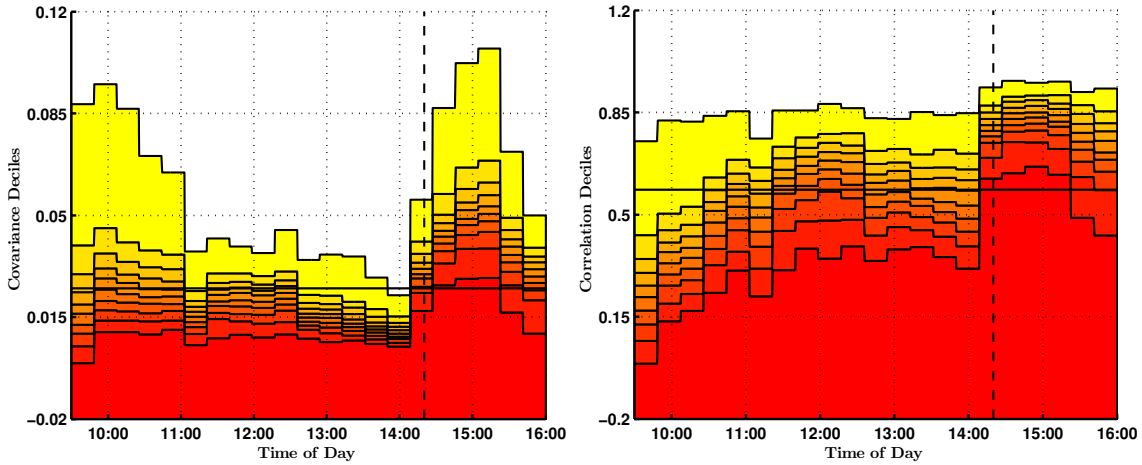
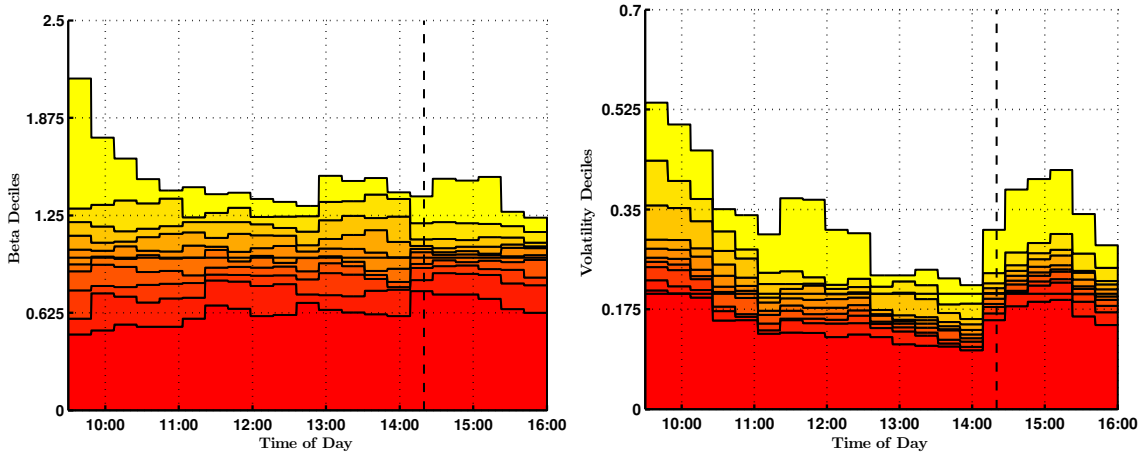


Figure 12: QQQ transaction prices (12/27/12). (1): Senate Majority Leader states that resolution to "fiscal cliff" crisis before January 1, 2013, unlikely. (2): News that the House of Representatives will convene on the following Sunday in an attempt to end the "fiscal cliff" crisis.



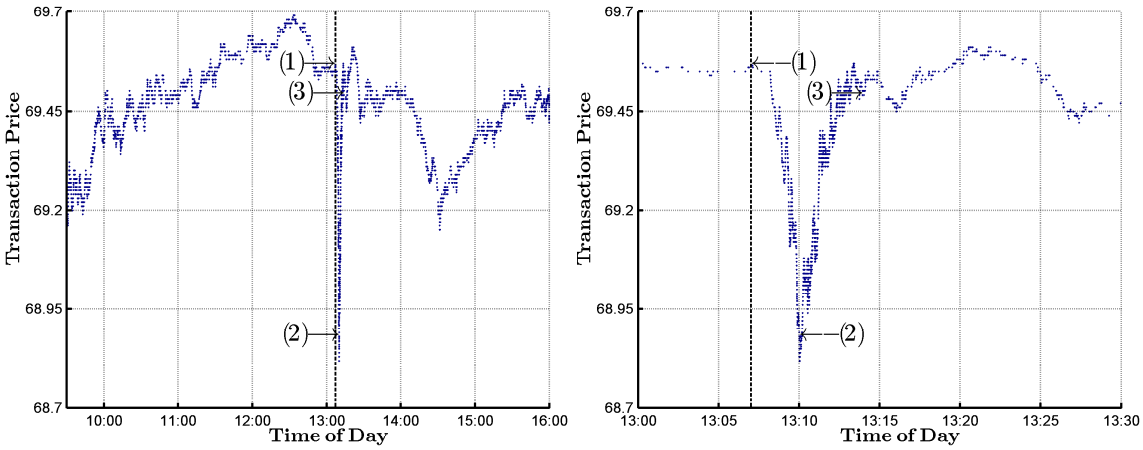
(a) Covariances (b) Correlations

Figure 13: Cross-sectional deciles of spot covariances and correlations (12/27/12). Solid horizontal line corresponds to the cross-sectional median of *integrated* covariance and correlation estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Covariances are annualized.



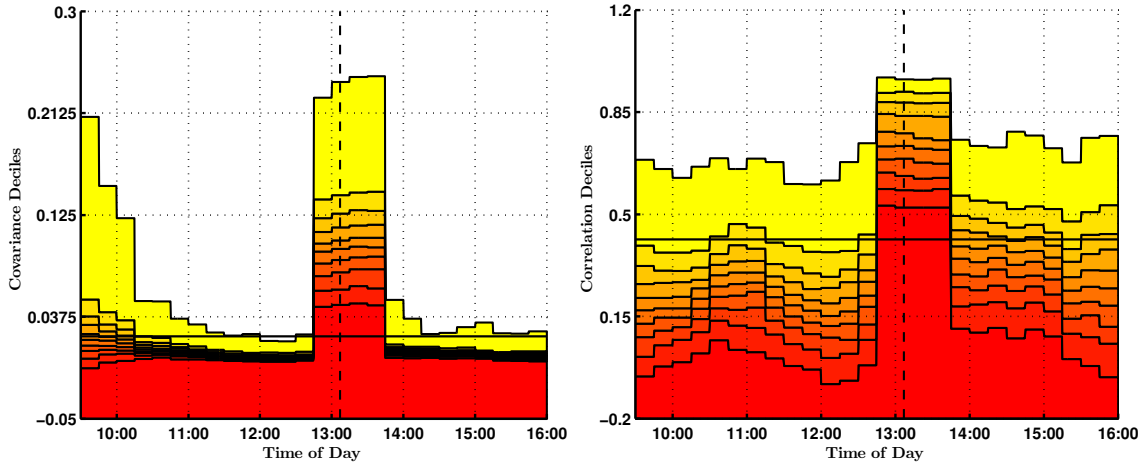
(a) Betas (b) Volatilities

Figure 14: Cross-sectional deciles of spot betas and volatilities (12/27/12). Solid horizontal line corresponds to the cross-sectional median of *integrated* beta and volatility estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Volatilities are annualized.



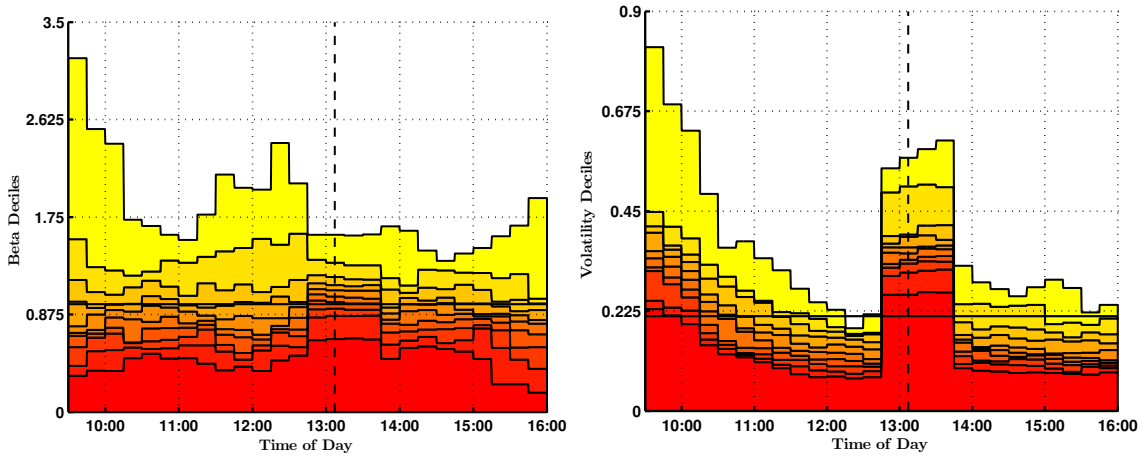
(a) Entire trading day (b) 1:00 pm – 1:30 pm

Figure 15: QQQ transaction prices (04/23/13). (1): Fake tweet from the account of AP stating "Breaking: Two Explosions in the White House and Barack Obama is injured". (2): Official denial by AP. (3): AP's twitter account suspended.



(a) Covariances (b) Correlations

Figure 16: Cross-sectional deciles of spot covariances and correlations (04/23/13). Solid horizontal line corresponds to the cross-sectional median of *integrated* covariance and correlation estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Covariances are annualized.



(a) Betas (b) Volatilities

Figure 17: Cross-sectional deciles of spot betas and volatilities (04/23/13). Solid horizontal line corresponds to the cross-sectional median of *integrated* beta and volatility estimates. These are based on the LMM estimator of the integrated (open-to-close) covariance matrix by Bibinger et al. (2014) accounting for serially dependent noise and using the same input parameter configuration as the spot estimators. Volatilities are annualized.