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

Web Appendix

Markus Bibinger Nikolaus Hautsch Peter Malec Markus Reiss

June 2017

Contents

1 Theoretical Supplement

1.1 Preliminaries

Consider the process

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

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 $\hat{\Sigma}_s$.[1] 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:

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

On the compact interval $[0,1]$, it can be assumed that $\|b_s\|$, $\|\sigma_s\|\|$ 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

$$\|W_j\big(\mathbf{H}_k^n, \Sigma_{kh_n}\big)\| \lesssim (\log(n))^{-1}\big(1 + j^2 (n h_n^2)^{-1}\big)^{-2} . \tag{3}$$

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 Equation (41c) and its generalization on p. 34 in Altmeyer and Bibinger (2015), given by

$$\begin{aligned} S_{jk}^{(p)} &\approx - \sum_{v=1}^{n_p - 1} Y_v^{(p)} \Big(\Phi_{jk}\big(\bar{t}_{v+1}^{(p)}\big) - \Phi_{jk}\big(\bar{t}_v^{(p)}\big) \Big) \\ &\approx - \sum_{v=1}^{n_p - 1} Y_v^{(p)} \varphi_{jk}(t_v^{(p)}) \frac{t_{v+1}^{(p)} - t_{v-1}^{(p)}}{2} \,, \end{aligned} \tag{4}$$

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.

[1]Numbers marked by an asterisk refer to equations in the paper.

The following orthogonality approximations are satisfied by Φ_{jk}, φ_{jk}:

$$\sum_{i=1}^{n_p} \Phi_{jk}\big(\bar{t}_i^{(p)}\big)\Phi_{qk}\big(\bar{t}_i^{(p)}\big)\big(t_i^{(p)} - t_{i-1}^{(p)}\big) = (\delta_{jq} + \mathcal{O}(1))\int_0^1 \Phi_{jk}^2(t)\,dt \qquad (5a)$$
$$= (\delta_{jq} + \mathcal{O}(1))h_n^2\pi^{-2}j^{-2},$$

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

$$\sum_{i=1}^{n_p-1}\left(\varphi_{jk}\big(t_i^{(p)}\big)\varphi_{qk}\big(t_i^{(p)}\big)\frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\right) = (\delta_{jq} + \mathcal{O}(1))\int_0^1 \varphi_{jk}^2(t)\,dt \qquad (5b)$$
$$= (\delta_{jq} + \mathcal{O}(1))\,.$$

Moreover, we have the following approximations:

$$\sum_{i=1}^{n_p-1} \varphi_{jk}^2\big(t_i^{(l)}\big)\Big(\tfrac{t_{i+1}^{(l)}-t_{i-1}^{(l)}}{2}\Big)^2 \approx \sum_{i=1}^{n_l-1} \varphi_{jk}^2\big(t_i^{(l)}\big)\tfrac{t_{i+1}^{(l)}-t_{i-1}^{(l)}}{2}\tfrac{(F_l^{-1})'(kh_n)}{n_l} \qquad (6)$$
$$\approx \Big(\int_0^1 \varphi_{jk}^2(t)\,dt\Big)\tfrac{(F_l^{-1})'(kh_n)}{n_l}\,,$$

$$\sum_{kh_n \le t_i^{(l)} \le (k+1)h_n} \big(t_i^{(l)} - t_{i-1}^{(l)}\big)^2 \approx \sum_{kh_n \le t_i^{(l)} \le (k+1)h_n} (F_l^{-1})'(kh_n)n_l^{-1}\big(t_i^{(l)} - t_{i-1}^{(l)}\big) \qquad (7)$$
$$= (F_l^{-1})'(kh_n)n_l^{-1}h_n\,,$$

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).

1.2 Proof of Theorem 1 from the Paper

For the sake of notational brevity, we present the proof of Theorem 1 for the two-sided version of the estimator and 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 the one-sided estimator and/or 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 $\eta_p, p = 1, \ldots, d$, refers to the long-run noise variance, namely the sum of all autocovariances, in component p.

Lemma 1. *By Assumption 3,*[2] *we have for* $p, q \in \{1, \ldots, d\}, p \neq q$*, that*

$$\mathbb{E}\Bigg[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\Big)^2\Bigg] = \nu_p \big(F_p^{-1}\big)' \eta_p n^{-1} + \mathcal{O}\big(n^{-1}\big)\,, \tag{8a}$$

$$\mathbb{E}\Bigg[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\Big)\Big(\sum_{i=1}^{n_q} \epsilon_i^{(q)} \varphi_{jk}\big(t_i^{(q)}\big) \frac{t_{i+1}^{(q)} - t_{i-1}^{(q)}}{2}\Big)\Bigg] = 0\,, \tag{8b}$$

$$\mathbb{E}\Bigg[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\Big)^4\Bigg] = \nu_p^2 \big(\big(F_p^{-1}\big)'\big)^2 3\,\eta_p^2 n^{-2} + \mathcal{O}\big(n^{-2}\big)\,, \tag{8c}$$

$$\begin{aligned}\mathbb{E}\Bigg[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\Big)^2\Big(\sum_{i=1}^{n_q} \epsilon_i^{(q)} \varphi_{jk}\big(t_i^{(q)}\big) \frac{t_{i+1}^{(q)} - t_{i-1}^{(q)}}{2}\Big)^2\Bigg] \\ = \nu_p \big(F_p^{-1}\big)' \eta_p \nu_q \big(F_q^{-1}\big)' \eta_q n^{-2} + \mathcal{O}\big(n^{-2}\big)\,.\end{aligned} \tag{8d}$$

Proof. Using (7) and an analogous estimate, we infer that

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

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 $\delta_{i,v}^R = \mathbb{1}_{\{|i-v| \leq R\}}$ for the following calculations. Accordingly, the fourth moments yield

$$\begin{aligned}&\mathbb{E}\Bigg[\Big(\sum_{i=1}^{n_p} \epsilon_i^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2}\Big)^4\Bigg] = \\ &\mathbb{E}\Bigg[\sum_{i,v,u,r=1}^{n_p} \epsilon_i^{(p)} \epsilon_v^{(p)} \epsilon_u^{(p)} \epsilon_r^{(p)} \varphi_{jk}\big(t_i^{(p)}\big) \varphi_{jk}\big(t_v^{(p)}\big) \varphi_{jk}\big(t_u^{(p)}\big) \varphi_{jk}\big(t_r^{(p)}\big) \frac{t_{i+1}^{(p)} - t_{i-1}^{(p)}}{2} \frac{t_{v+1}^{(p)} - t_{v-1}^{(p)}}{2}\end{aligned}$$

[2] Assumptions and Theorem 1 always refer to the main paper.

$$\times \tfrac{t_{u+1}^{(p)}-t_{u-1}^{(p)}}{2} \tfrac{t_{r+1}^{(p)}-t_{r-1}^{(p)}}{2}\Big]$$

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

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

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

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:

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

Theorem 1 is implied by Proposition 1.1, Proposition 1.2 and Proposition 1.3 below.

Proposition 1.1. *On the assumptions of Theorem 1, for any $s \in (0,1)$ it holds true that*

$$n^{\beta/2}\operatorname{vec}\big(\hat\Sigma_s^{or}(\tilde X+\epsilon) - \Sigma_s\big) \stackrel{d-(st)}{\longrightarrow} \mathbf{N}\Big(0, 2\big(\Sigma\otimes\Sigma_H^{1/2} + \Sigma_H^{1/2}\otimes\Sigma\big)_s\mathcal{Z}\Big)\,. \tag{9}$$

Proof. We consider the estimator (12*) under observations of the simplified process $\tilde X + \epsilon$ and associated spectral statistics $\tilde S_{jk}$ in (2). By virtue of the identity

$$\mathbb{C}\mathrm{OV}\big(\operatorname{vec}\big(\tilde S_{jk}\tilde S_{mk}^\top\big)\big|\mathcal{F}_{kh_n}\big) = I_{jk}^{-1}\mathcal{Z}(\delta_{jm} + \mathcal{O}(1))\,, \tag{10}$$

we deduce the conditional variance-covariance matrix

$$\begin{aligned}\mathbb{C}\mathrm{OV}\big(\mathrm{vec}\,\big(\hat{\Sigma}_s^{or}\big)\big|\mathcal{F}_{U_{s,n}}\big) &= \sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor+K_n}(2K_n+1)^{-2}\sum_{j=1}^{J_n}W_{jk}\mathbb{C}\mathrm{OV}\Big(\mathrm{vec}\,\big(\tilde{S}_{jk}\tilde{S}_{jk}^{\top}\big)\big|\mathcal{F}_{kh_n}\Big)W_{jk}^{\top}\\ &\qquad + \mathcal{O}(K_n^{-1})\\ &= \sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor+K_n}(2K_n+1)^{-2}I_k^{-1}\sum_{j=1}^{J_n}I_{jk}I_{jk}^{-1}\mathcal{Z}I_{jk}I_k^{-1}+\mathcal{O}(K_n^{-1})\\ &= \sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor+K_n}(2K_n+1)^{-2}I_k^{-1}\mathcal{Z}+\mathcal{O}(K_n^{-1})\,,\end{aligned}$$

using the tower property of conditional expectation and 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 (10), the asymptotic form of the conditional variance-covariance matrix follows by bounding $\mathbb{E}\big[\|I_l^{-1}-I_m^{-1}\|\big]$ for different blocks with $|l-m|\le 2K_n+1$ and a computation analogously to the proof of Corollary 4.3 of Bibinger *et al.* (2014). There, the asymptotic theory is pursued for continuous-time observations, but, once we have the conforming illustration for I_k^{-1}, the analysis is similar. Using block-wise transformations which diagonalize Σ_{kh_n} and transfer the noise level (5*) to the identity matrix, i.e.,

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

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

$$\begin{aligned}\mathbb{C}\mathrm{OV}\Big(\,\mathrm{vec}\,\big(\hat{\Sigma}_s^{or}\big)\big|\mathcal{F}_{U_{s,n}}\Big) &= \sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor+K_n}(2K_n+1)^{-2}\big(O_kH_k^{-1}\big)^{-\otimes 2}\tilde{I}_k^{-1}\big(H_k^{-1}O_k^{\top}\big)^{-\otimes 2}\mathcal{Z}\\ &\qquad + \mathcal{O}(K_n^{-1}),\end{aligned}$$

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 then derive that

$$\begin{aligned}\mathbb{C}\mathrm{OV}\Big(\,\mathrm{vec}\,\big(\hat{\Sigma}_s^{or}\big)\big|\mathcal{F}_{U_{s,n}}\Big) = (2+\mathcal{O}(1))&\sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor+K_n}(2K_n+1)^{-2}\\ &\times\big(\Sigma_{kh_n}\otimes\big(\Sigma_H^{kh_n}\big)^{1/2}+\big(\Sigma_H^{kh_n}\big)^{1/2}\otimes\Sigma_{kh_n}\big)\mathcal{Z}\,,\end{aligned}$$

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 by approximating Σ and H constant over the smoothing window using the smoothness of Σ and H granted by Assumptions 2-4. Therefore, to deduce the variance-covariance structure, it remains to prove that (10) is indeed valid. The computations rely on the preliminaries above, namely summation by parts, the orthogonality relations (5a) and (5b), as well as (6) and (7). For the signal term, we apply Lemma 4.4 of Altmeyer and Bibinger (2015), 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 $t_i, i = 0, \ldots, n$, that

$$\mathbb{C}\mathrm{OV}\big(\,\mathrm{vec}\,(\pi^2 j^2 h_n^{-2}\mathcal{X}\mathcal{X}^\top)\big|\mathcal{F}_{kh_n}\big) = \frac{1}{n}\Big(\Sigma_{kh_n}^{\otimes 2}\mathcal{Z}\Big)\,,$$

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

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

with $\mathbf{H}_k^n$ as defined in (9*), as well as by Lemma 1, (5b) and (6). The multiplication with $\mathcal{Z}/2$ when evaluating the variance-covariance matrix of the vectorized diagonal matrix is performed by elementary matrix calculus. While we consider in general random *conditional* covariances of signal terms, we may work with deterministic time-varying noise levels under Assumption 4 that sampling times are exogenous (condition in case of independent random sampling). 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 (10). Considering the statistics

$$\zeta_s^n = n^{\beta/2}\sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n}(2K_n+1)^{-1}\sum_{j=1}^{J_n} W_j\big(\mathbf{H}_k^n, \Sigma_{kh_n}\big)\tilde{Z}_{jk}\,, \tag{11a}$$

$$\tilde{Z}_{jk} = \mathrm{vec}\,\Big(\tilde{S}_{jk}\tilde{S}_{jk}^\top - \pi j^2 h_n^{-2}\mathbf{H}_k^n - \Sigma_{kh_n}\Big)\,, \tag{11b}$$

we obtain the following convergences:

$$\mathbb{E}\left[\zeta_s^n\big|\mathcal{F}_{U_{s,n}}\right] \xrightarrow{p} 0\,, \tag{11c}$$

$$\mathbb{E}\left[\left(\zeta_s^n(\zeta_s^n)^\top\right)\big|\mathcal{F}_{U_{s,n}}\right] \stackrel{p}{\to} 2\big(\Sigma\otimes\Sigma_H^{1/2}+\Sigma_H^{1/2}\otimes\Sigma\big)_s\mathcal{Z}\,, \tag{11d}$$

$$\mathbb{E}\left[\left(\zeta_s^n(\zeta_s^n)^\top\zeta_s^n(\zeta_s^n)^\top\right)\big|\mathcal{F}_{U_{s,n}}\right] \stackrel{p}{\to} 0\,. \tag{11e}$$

In case that $\alpha \leq 1/2$ in Assumption 2, volatility jumps can occur. However, a standard estimate gives

$$\mathbb{P}\Big(\sup_{t\in[(\lfloor sh_n^{-1}\rfloor-K_n)h_n,(\lfloor sh_n^{-1}\rfloor+K_n)h_n]}\|\sigma_t-\sigma_{t-}\|>0\Big)\leq C\,K_nh_n\to 0 \text{ as } n\to\infty\,, \tag{12}$$

with $\sigma_{t-}=\lim_{\tau\to t,\tau<t}\sigma_\tau$ and a constant C. This shows that for any fixed time the probability that the semi-martingale component $\sigma_s^{(1)}$ exhibits jumps in $[(\lfloor sh_n^{-1}\rfloor-K_n)h_n,(\lfloor sh_n^{-1}\rfloor+K_n)h_n]$ tends to zero. From the explicit form of I_k^{-1} from (13*) the difference $\mathbb{E}\big[\|I_l^{-1}-I_m^{-1}\|\big],|l-m|\leq 2K_n+1$, is determined by $\mathbb{E}\big[\|\sigma_{lh_n}-\sigma_{mh_n}\|\big],|l-m|\leq 2K_n+1$ on Assumption 2. By (12), the order $(K_nh_n)^\alpha$ of the approximation in the variance is readily obtained and is of the same order as the bias of the estimator. The above analysis for the variance thus implies (11d). The bias condition (11c) is readily obtained using Lemma 1, summation by parts, Itô isometry for the signal part, as well as (5a), (5b) and (12). The Lyapunov condition (11e) follows using analogous bounds as for the stable CLT of the integrated version in Section 6.2.3 of Altmeyer and Bibinger (2015). We are left to prove that $\alpha_n = n^{\beta/2}\,\mathrm{vec}\big(\hat{\Sigma}_s^{or}(\tilde{X}+\epsilon)-\Sigma_s\big)$ satisfy

$$\mathbb{E}\left[Zg(\alpha_n)\right]\to\mathbb{E}\left[Zg(\alpha)\right]=\mathbb{E}[Z]\mathbb{E}\left[g(\alpha)\right]\,, \tag{13}$$

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

$$\alpha=\big(\Sigma^{1/2}\otimes\Sigma_H^{1/4}\big)_s\mathcal{Z}U+\big(\Sigma_H^{1/4}\otimes\Sigma^{1/2}\big)_s\mathcal{Z}U' \tag{14}$$

and $U,U'\in\mathbb{R}^{d^2}$ being two independent standard normally distributed vectors independent of $\mathcal{F}$. (13) ensures that the central limit theorem (9) is $\mathcal{F}$-stable. The limit α gives indeed the asymptotic law of (9) as $\mathcal{Z}^2=2\mathcal{Z}$ and

$$\big(\Sigma^{1/2}\otimes\Sigma_H^{1/4}\big)_s\mathcal{Z}\Big(\big(\Sigma^{1/2}\otimes\Sigma_H^{1/4}\big)_s\mathcal{Z}\Big)^\top=2\big(\Sigma\otimes\Sigma_H^{1/2}\big)_s\mathcal{Z}\,,$$

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

The proof of (13) 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, $\big(\mathcal{H}_n\big)_{n\in\mathbb{N}}$ 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 $\Delta_i\tilde{X}^{(p)}, p=1,\ldots,d$, of $\tilde{X}(n)_t$ and uncorrelated from those of $\bar{X}(n)_t$. For all $n\ge n_0$, 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 (13) and completes the proof of Proposition 1.1. □

Proposition 1.2. *On the assumptions of Theorem 1, for any $s\in(0,1)$ it holds true that*

$$n^{\beta/2}\,\mathrm{vec}\left(\hat{\Sigma}_s^{or}-\hat{\Sigma}_s^{or}(\tilde{X}+\epsilon)\right)\xrightarrow{p}0\,. \tag{15}$$

Proof. The left-hand side above equals

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

Using that

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

and the order of the weights provided by (3), a crude estimate suffices here:

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

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

Proposition 1.3. *On the assumptions of Theorem 1, for any $s\in(0,1)$ it holds true that*

$$n^{\beta/2}\big(\operatorname{vec}\big(\hat\Sigma_s-\hat\Sigma_s^{or}\big)\big)\xrightarrow{p}0\,. \tag{16}$$

Proof. We write the left-hand side above

$$n^{\beta/2}\big(\operatorname{vec}\big(\hat\Sigma_s-\hat\Sigma_s^{or}\big)\big)=n^{\beta/2}\big(2K_n+1\big)^{-1}\sum_{k=\lfloor sh_n^{-1}\rfloor-K_n}^{\lfloor sh_n^{-1}\rfloor-K_n}\sum_{j=1}^{J_n}\Big(\hat W_j\big(\hat\Sigma_{kh_n}\big)-W_j\big(\Sigma_{kh_n}\big)\Big)Z_{jk},$$

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

By Proposition 1.2, 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 A 1 and (7). We decompose the difference between adaptive and oracle estimator using

$$\begin{aligned}\hat W_j\big(\hat\Sigma_{kh_n}\big)-W_j\big(\Sigma_{kh_n}\big)&=\hat W_j\big(\hat\Sigma_{m(2K_n+1)h_n}\big)-W_j\big(\Sigma_{m(2K_n+1)h_n}\big)+\\&\quad+\hat W_j\big(\hat\Sigma_{kh_n}\big)-\hat W_j\big(\hat\Sigma_{m(2K_n+1)h_n}\big)\\&\quad+W_j\big(\Sigma_{m(2K_n+1)h_n}\big)-W_j\big(\Sigma_{kh_n}\big)\,,\end{aligned}$$

where $\hat\Sigma_{m(2K_n+1)h_n}, m=0,\ldots,\lfloor(2K_n+1)^{-1}h_n^{-1}\rfloor-1$, 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

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

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 (3), 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 Delta-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 $\sum_j W_j = I_{d^2 \times d^2}$, 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:

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

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 $I_k^{-1}\big(\Sigma_{kh_n}\big)$. $\square$

1.3 Microstructure Noise: Endogeneity and Serial Correlation

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 in this Section 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

$$\hat{\eta}_R = (2n)^{-1}\sum_{i=1}^{n}\big(\Delta_i Y\big)^2 + n^{-1}\sum_{r=1}^{R}\sum_{i=1}^{n-r}\Delta_i Y \Delta_{i+r} Y\,, \tag{17a}$$

$$\hat{\eta}_r - \hat{\eta}_{r+1} = (2n)^{-1}\sum_{i=1}^{n}\big(\Delta_i Y\big)^2 + n^{-1}\sum_{u=1}^{r}\sum_{i=1}^{n-u}\Delta_i Y \Delta_{i+u} Y\,,\; 0 \leq r \leq R-1\,. \tag{17b}$$

In particular, this includes

$$\hat{\eta}_0 - \hat{\eta}_1 = (2n)^{-1}\sum_{i=1}^{n}\big(\Delta_i Y\big)^2\,, \tag{17c}$$

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 $q, r, r' \in \{0, \ldots, R\}$

$$\tilde{\Gamma}_q^{rr'} = n^{-1}\sum_{i=1}^{n-(r\vee(q+r'))}\Delta_i Y \Delta_{i+r} Y, \Delta_{i+q} Y \Delta_{i+q+r'} Y - [\hat{\eta}_r - \hat{\eta}_{r+1} - (\hat{\eta}_{r-1} - \hat{\eta}_r)] \tag{18}$$
$$\times\, [\hat{\eta}_{r'} - \hat{\eta}_{r'+1} - (\hat{\eta}_{r'-1} - \hat{\eta}_{r'})]\,,$$

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 (17b). Then, the variance of $\hat{\eta}_r, 0 \leq r \leq R$, is consistently estimated by

$$\widehat{\mathbb{V}\mathrm{ar}}(\hat{\eta}_r) = n^{-1}\left(V_{r+1}^n + V_r^n + 2C_{r,r+1}^n\right) , \tag{19a}$$

$$\text{with } C_{r,r+1}^n = \left(\frac{\hat{\Gamma}_0^{00}}{4} + \frac{1}{2}\sum_{u=1}^{r}\hat{\Gamma}_u^{00} + \sum_{u=0}^{r}\sum_{u'=1}^{r+1}\left(\hat{\Gamma}_0^{uu'} + 2\sum_{q=1}^{R}\hat{\Gamma}_q^{uu'}\right)\right), \tag{19b}$$

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 A 1. *Under Assumption 3, the following central limit theorem applies to the estimators defined by* (17a) *and* (17b)*:*

$$\sqrt{n}(V_r^n + V_{r+1}^n + 2C_{r,r+1}^n)^{-1/2}\left(\hat{\eta}_r - \eta_r\right) \xrightarrow{d} \mathbf{N}(0,1)\,. \tag{20}$$

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

$$T_Q^n(Y) = \sqrt{n/V_Q^n}\,\hat{\eta}_Q \xrightarrow{d} \mathbf{N}(0,1)\,. \tag{21}$$

Proof. Observe that

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

for $r \geq 1, 1 \leq i \leq (n-r)$, and

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

The remainders stem from the signal terms X, which are of smaller order. Hence, by the definition of the estimators in (17b), 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 (17a)-(17c), denote by

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

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

$$\begin{aligned}
\mathbb{V}\text{ar}(\hat{\eta}_R) &= (2n)^{-2}\sum_{i=1}^{n}\mathbb{V}\text{ar}\big((\Delta_i Y)^2\big) + (2n)^{-2}\sum_{i=1}^{n-R}2\sum_{u=1}^{R}\mathbb{C}\text{ov}\big((\Delta_i Y)^2, (\Delta_{i+u}Y)^2\big) \\
&\quad + n^{-2}\sum_{i=1}^{n-R}\sum_{r=1}^{R}\sum_{r'=1}^{R}\Big(\mathbb{C}\text{ov}\big(\Delta_i Y\Delta_{i+r}Y, \Delta_i Y\Delta_{i+r'}Y\big) \\
&\qquad\qquad + 2\sum_{u=1}^{R}\mathbb{C}\text{ov}\big(\Delta_i Y\Delta_{i+r}Y, \Delta_{i+u}Y\Delta_{i+u+r'}Y\big)\Big) \\
&\quad + n^{-2}\sum_{i=1}^{n-R}\sum_{r=1}^{R}\Big(\mathbb{C}\text{ov}\big((\Delta_i Y)^2, \Delta_i Y\Delta_{i+r}Y\big) \\
&\qquad\qquad + 2\sum_{u=1}^{R}\mathbb{C}\text{ov}\big((\Delta_i Y)^2, \Delta_{i+u}Y\Delta_{i+u+r}Y\big)\Big) \\
&= n^{-1}\left(\tfrac{1}{4}\Gamma^{00}_0 + \tfrac{1}{2}\sum_{u=1}^{R}\Gamma^{00}_u + \sum_{r=1}^{R}\sum_{r'=1}^{R}\left(\Gamma^{rr'}_0 + 2\sum_{u=1}^{R}\Gamma^{rr'}_u\right)\right. \\
&\qquad\qquad \left. + \sum_{r=1}^{R}\left(\Gamma^{0r}_0 + 2\sum_{u=1}^{R}\Gamma^{0r}_u\right)\right) + \mathcal{O}\big(n^{-1}\big) \\
&= n^{-1}\left(\tfrac{1}{4}\left(\Gamma^{00}_0 + 2\sum_{u=1}^{R}\Gamma^{00}_u\right) + \sum_{r=0}^{R}\sum_{r'=1}^{R}\left(\Gamma^{rr'}_0 + 2\sum_{u=1}^{R}\Gamma^{rr'}_u\right)\right) + \mathcal{O}\big(n^{-1}\big) .
\end{aligned}$$

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

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

where $\mathcal{V}^n_r$ are defined analogously to $\mathcal{V}^n_R$ except replacing R by $r < R$ and

$$\mathcal{C}^n_{r,r+1} = \left(\frac{\Gamma^{00}_0}{4} + \frac{1}{2}\sum_{u=1}^{R}\Gamma^{00}_u + \sum_{u=0}^{r}\sum_{u'=1}^{r+1}\left(\Gamma^{uu'}_0 + 2\sum_{q=1}^{R}\Gamma^{uu'}_q\right)\right) .$$

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). $\square$

Table 1: 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 $\theta_{\epsilon,2} = 0.3$. $\hat{R}$ and $\hat{\eta}$ are computed with (17b) and following the tests based on (21), 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	$\bar{\hat{R}}$	STD$\big(\hat{R}\big)$	BIAS($\hat{\eta}$)	STD($\hat{\eta}$)
0	0.00	0.05	0.000	3.467
1	1.11	0.61	-0.026	3.186
2	2.37	1.34	-0.073	2.678

The statistic $T^n_Q(Y)$ 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 $T^n_Q(Y)$ 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 $\hat{\eta}_0, \ldots, \hat{\eta}_{\hat{R}}$.

Finally, we study the precision of the procedure for estimating the long-run noise variance η in simulations and the same setup as in Section 4 of the paper. As the estimation of the long-run noise variance using the above procedure is unrelated to the problem dimension, we consider the processes in (24a), (24b) and (25) with $d = 1$ to evaluate the aforementioned approach and set $n = 23,400$. Further, we employ microstructure noise processes based on different orders of serial dependence R. Accordingly, we assume $\epsilon_i = \Theta_R(L)\, u_i$, $\Theta_R(L) := \sum_{r=0}^{R} \theta_{\epsilon,r} L^r$, $\theta_{\epsilon,0} = 1$, $u_i | \{(X, \sigma)\} \sim N\big(0, \eta/\Theta_R(1)^2\big)$, $i = 1, \ldots, n$. We consider 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 1 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 slightly over-estimates R, resulting in more conservative estimates of the order of serial dependence in the noise process. Generally, we can conclude that the proposed approach provides a satisfactory precision in a realistic scenario.

Spectral estimation under endogenous noise

Assumption A 1. *Let* $\mathbb{C}\mathrm{ov}(\epsilon_i, \Delta_j X)$ *be possible non-zero covariances for* $1 \vee (i - M) \leq j \leq (i + M) \wedge n$*, between the market microstructure noise and the efficient returns for some* $M < \infty$*.*

The time-invariant quantity

$$\varrho(\epsilon, X) = \sum_{i=1}^{2M+1} \mathbb{C}\mathrm{ov}(\epsilon_{M+1}, \Delta_i X) \tag{22}$$

measures the total long-run covariance between efficient returns and noise.

From a conceptual point of view, the exogeneity of noise in Assumption 3 appears natural, since it reflects the original motivation of the model with dominant microstructure noise that is ancillary in terms of not providing any information about the volatility. If there was a significant contribution in the ϵ_is depending on $\Delta_i X$ or σ_{t_i}, this would lead to a different model in which the information carried about volatility in the noise could play a role for statistical inference on the volatility. A model with microstructure that is informative about volatility, however, appears unrealistic. Still, endogeneities could occur as reported in Hansen and Lunde (2006) and should be considered as misspecification. In the following, we shall prove that our estimators attain exactly the same asymptotic properties under Assumption A 1 replacing exogenous noise in Assumption 3 of the paper.
Unbiasedness of the estimator (12*) readily follows from the summation by parts decomposition

$$S_{jk} = \left(\sum_{i=1}^{n} \Delta_i X \Phi_{jk}\big(\frac{t_{i-1}+t_i}{2}\big) - \sum_{i=1}^{n-1} \epsilon_i \, \varphi_{jk}(t_i) \, \frac{t_{i+1}-t_{i-1}}{2} \right) \big(1 + \mathcal{O}(1)\big)$$

with the identities

$$\int \Phi_{jk}(t) \, \varphi_{uk}(t) \, dt = \frac{(1-\cos(\pi j)\cos(\pi u))2h_n}{\pi^2(j^2-u^2)} \,, \tag{23}$$

which give $4h_n/(\pi^2(j^2-u^2))$ whenever j is odd and u even, or the other way round. Importantly, the above integral vanishes in the case that $j = u$. Integral approximations of sums similar as in the proof of Theorem 1 above then give $\mathbb{E}[S_{jk}^2]$ is equal as in the case of exogenous noise. Furthermore, $\mathbb{V}\mathrm{ar}(S_{jk}^2)$ is invariant as well and the weights derived under exogenous noise are still optimal. Nevertheless, the endogeneity manipulates the laws of the S_{jk}s and we need to consider possible covariances $\mathbb{C}\mathrm{ov}(S_{jk}, S_{uk})$, at least for u and j with different parities, for the asymptotic distribution of the overall estimator.
Under exogenous noise all spectral statistics (S_{jk}) were (approximately) uncorrelated by orthogonality relations of $(\Phi_{jk})_{j\ge1}$ and $(\varphi_{jk})_{j\ge1}$ even on the same block k. Now, we reconsider possible covariances under Assumption A 1. We derive that

$$\frac{h_n^4}{\pi^4 u^2 j^2} \, \mathbb{C}\mathrm{ov}(S_{jk}^2, S_{uk}^2) =$$

$$
\begin{aligned}
&\mathbb{C}\mathrm{ov}\Bigg(\Big(\sum_{i=1}^{n}\Delta_i X\Phi_{jk}\big(\frac{t_{i-1}+t_i}{2}\big)\Big)^2-2\sum_{i=1}^{n-1}\Delta_i X\Phi_{jk}\big(\frac{t_{i-1}+t_i}{2}\big)\epsilon_i\,\varphi_{jk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\\
&\qquad\qquad\qquad\qquad\qquad\qquad+\Big(\sum_{i=1}^{n-1}\epsilon_i\,\varphi_{jk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\Big)^2,\\
&\Big(\sum_{i=1}^{n}\Delta_i X\Phi_{uk}\big(\frac{t_{i-1}+t_i}{2}\big)\Big)^2-2\sum_{i=1}^{n-1}\Delta_i X\Phi_{uk}\big(\frac{t_{i-1}+t_i}{2}\big)\epsilon_i\,\varphi_{uk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\\
&\qquad\qquad\qquad\qquad\qquad\qquad+\Big(\sum_{i=1}^{n-1}\epsilon_i\,\varphi_{uk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\Big)^2\Bigg)\\
&=\big(1+\mathcal{O}(1)\big)\Bigg(\mathbb{C}\mathrm{ov}\Big(\Big(\sum_{i=1}^{n}\Delta_i X\Phi_{jk}\big(\frac{t_{i-1}+t_i}{2}\big)\Big)^2,\Big(\sum_{i=1}^{n-1}\epsilon_i\,\varphi_{uk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\Big)^2\Big)\\
&+\mathbb{C}\mathrm{ov}\Big(\Big(\sum_{i=1}^{n}\Delta_i X\Phi_{uk}\big(\frac{t_{i-1}+t_i}{2}\big)\Big)^2,\Big(\sum_{i=1}^{n-1}\epsilon_i\,\varphi_{jk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\Big)^2\Big)\\
&+4\mathbb{C}\mathrm{ov}\Big(\sum_{i=1}^{n-1}\Delta_i X\Phi_{jk}\big(\frac{t_{i-1}+t_i}{2}\big)\epsilon_i\,\varphi_{jk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2},\\
&\qquad\qquad\qquad\qquad\sum_{i=1}^{n-1}\Delta_i X\Phi_{uk}\big(\frac{t_{i-1}+t_i}{2}\big)\epsilon_i\,\varphi_{uk}(t_i)\,\frac{t_{i+1}-t_{i-1}}{2}\Big)\Bigg)\\
&=\frac{(\varrho(\epsilon,X))^2}{n}\Bigg(2\Big(\int\Phi_{jk}(t)\varphi_{uk}(t)\,dt\Big)^2+2\Big(\int\Phi_{uk}(t)\varphi_{jk}(t)\,dt\Big)^2\\
&\qquad\qquad\qquad\qquad+4\int\Phi_{jk}(t)\varphi_{uk}(t)\,dt\int\Phi_{uk}(t)\varphi_{jk}(t)\,dt\Bigg)\big(1+\mathcal{O}(1)\big)\\
&=\frac{(\varrho(\epsilon,X))^2}{n}\,\frac{16h_n^2}{\pi^4}\,\Big(4(j^2-u^2)^2-4(j^2-u^2)^2\Big)\big(1+\mathcal{O}(1)\big)\,.
\end{aligned}
$$

In fact, the leading term vanishes. If the term in parenthesis was not zero, the above term would contribute to the asymptotic variance of the estimator, while the rate of convergence would remain the same. Thus, the effect of endogenous noise on our estimator is asymptotically negligible at first order. We study the finite-sample effect of endogenous noise in a simulation. For simplicity, we repeat the one-dimensional simulation study from Altmeyer and Bibinger (2015). Additionally to the setup with exogenous noise, we implement 'extremely high' endogenous noise of the form

$$\epsilon_i\sim\sqrt{3/4}\cdot Z_i+\sqrt{0.001\cdot n/4}\cdot\Delta_i X,\ i=1,\ldots,n\,,\ \epsilon_0\sim Z_0\,.$$

$Z_i\sim N(0,\eta^2)$ are i.i.d. standard normal random variables with $\eta=0.01$. Compared to the results under exogenous noise, see Section 5 of Altmeyer and Bibinger (2015), the finite-sample variance slightly increases by ca. 10%, while the estimator remains perfectly unbiased. This shows that endogeneity could slightly increase the finite-sample variance, but does not harm the main properties of the estimator. Also, we expect the finite-sample effect not to be relevant under a realistic magnitude of correlation between signal and noise.

2 Simulation Study

2.1 Setting

We consider a high-dimensional setting with $d = 15$. For 15 assets, we estimate a 120-dimensional volatility matrix and the estimator utilizes weight matrices with 7260 entries. To ensure parsimony in this framework, we assume that the efficient log-price process follows a simple factor structure as employed, e.g., in Barndorff-Nielsen *et al.* (2011). We extend the latter to incorporate both a flexible 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

$$dX_t^{(p)} = \mu\, dt + \phi_t^{(p)} \tilde{\sigma}_t^{(p)} dV_t^{(p)}, \quad dV_t^{(p)} := \tilde{\rho}\, dW_t + \sqrt{1-\tilde{\rho}^2}\, dB_t^{(p)}, \tag{24a}$$

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

where W_t and $B_t^{(p)}$ are independent standard Brownian motions, while $\tilde{\sigma}_t^{(p)}$, $p = 1, \ldots, d$, are the stochastic volatility components. Hence, the correlation between two processes is $d\big[X^{(p)}, X^{(q)}\big]_t / \sqrt{d\big[X^{(p)}\big]_t d\big[X^{(q)}\big]_t} = \tilde{\rho}^2$, $p \neq q$, which is tantamount to an equicorrelation structure. The stochastic volatility components $\tilde{\sigma}_t^{(p)}$, $p = 1, \ldots, d$, are assumed to follow the two-factor model introduced by Chernov *et al.* (2003) as it allows for both volatility persistence and pronounced volatility of volatility. Hence,

$$\tilde{\sigma}_t^{(p)} = \text{s–exp}\left[\beta_0 + \beta_1 v_{1,t}^{(p)} + \beta_2 v_{2,t}^{(p)}\right], \quad p = 1, \ldots, d, \tag{25}$$
$$dv_{1,t}^{(p)} = \alpha_1 v_{1,t}^{(p)} dt + dW_{1,t}^{(p)}, \quad dv_{2,t}^{(p)} = \alpha_2 v_{2,t}^{(p)} dt + \left(1 + \beta_v v_{2,t}^{(p)}\right) dW_{2,t}^{(p)},$$

where $W_{1,t}^{(p)}$ and $W_{2,t}^{(p)}$ are independent standard Brownian motions with

$$d\big[W_1^{(p)}, V^{(q)}\big]_t / \sqrt{d\big[W_1^{(p)}\big]_t d\big[V^{(q)}\big]_t} = \rho_1 \delta_{pq}, \tag{26}$$
$$d\big[W_2^{(p)}, V^{(q)}\big]_t / \sqrt{d\big[W_2^{(p)}\big]_t d\big[V^{(q)}\big]_t} = \rho_2 \delta_{pq}, \quad p, q = 1, \ldots, d,$$

thereby allowing for leverage effects, and with the s–exp-function defined as in Appendix A of Chernov *et al.* (2003). The latter ensures that a unique solution to (25) exists by splicing the exponential function with appropriate growth conditions at an extremely high volatility level.

The drift term in (24a) is set to $\mu = 0.03$. The seasonality component is normalized such that $\int_0^1 \phi_t^{(p),2} dt = 1$, $p = 1, \ldots, d$, while to choose parameter values, we divide the 15 efficient price processes into three groups. For series $1-5$, parameters are set to the median estimates for mid-quote revisions of the 10 most liquid constituents of the Nasdaq 100 in 2010-14. For series $6-10$ and $11-15$, medians of the 10 second and third most liquid Nasdaq 100 constituents are considered, respectively. See Section 2 in the paper for a summary of the dataset. In this context, we employ the estimation procedure by Andersen and Bollerslev (1997) estimating the daily volatility component based on sub-sampled five-minute realized variances instead of a parametric GARCH approach. To calibrate the spot (equi-)correlation between the efficient price processes, we set $\tilde{\rho} = \sqrt{\rho_{\text{emp}}}$, where $\rho_{\text{emp}} = 0.312$ is the cross-sectional median of the (across-day averages of) realized correlations based on the Nasdaq data from our empirical application. These are computed using the multivariate realized kernel by Barndorff-Nielsen *et al.* (2011). For the stochastic volatility components (25), we follow Huang and Tauchen (2005), setting $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.137e^{-2}$, $\alpha_2 = -1.386$, $\beta_v = 0.25$, $\rho_1 = \rho_2 = -0.3$. The multivariate process in (24a), (24b) and (25) is then simulated by a Euler discretization scheme based on a $1/5$-second grid assuming $23,400$ seconds per trading day.

We dilute the observations of the efficient log-price process by serially dependent microstructure noise with $R = 1$, i.e., $\epsilon_i^{(p)} = \theta_\epsilon \epsilon_{i-1}^{(p)} + u_i^{(p)}$, $u_i^{(p)} | \left\{ \left(X^{(p)}, \sigma^{(p)} \right) \right\} \sim N\!\left(0, \eta_p / \left(1 + \theta_\epsilon\right)^2\right)$, $i = 1, \ldots, n_p$, $p = 1, \ldots, d$. To ensure that the absolute noise level is in line with the variation in the volatility process, we determine its long-run variance η_p by choosing the noise-to-signal ratio per trade $\xi_p^2 := n_p \eta_p / \sqrt{\int_0^1 \phi_t^{(p),4} \tilde{\sigma}_t^{(p),4} dt}$, $p = 1, \ldots, d$. The latter specification implies endogenous noise as the long-run variance of the latter depends on the given volatility path of the efficient price process. We set ξ_p, $p = 1, \ldots, d$, to the deciles of the respective estimates from the empirical study. See Table 2 for details. Here, η_p, $p = 1, \ldots, d$, is estimated following the procedure from Section 1.3, while the integrated quarticity is approximated by the squared sub-sampled five-minute realized variance. We choose $\theta_\epsilon = 0.6$, yielding a first-order autocorrelation of $\eta_1 = 0.441$, which is the median estimate for the underlying Nasdaq data. Finally, asynchronicity effects are introduced by drawing the observation times $t_i^{(p)}$, $i = 1, \ldots, n_p$, from independent Poisson processes with intensities λ_p, $p = 1, \ldots, d$. The latter are set to the cross-sectional deciles of the (across-day) average number of mid-quote revisions per second in the Nasdaq data with the exact values reported in Table 2.

Table 2: Noise-to-signal ratios and observation intensities used in the simulation study. ξ_p denotes the square root of the noise-to-signal ratio per trade, i.e., $\xi_p^2 := n_p \eta_p / \sqrt{\int_0^1 \phi_t^{(p),4} \tilde{\sigma}_t^{(p),4} dt}$, $p = 1, \ldots, d$. λ_p denotes the intensity of the Poisson process generating the observation times $t_i^{(p)}$, $i = 1, \ldots, n_p$, $p = 1, \ldots, d$. Both ξ_p and λ_p represent the minimum and cross-sectional deciles of estimates based on the Nasdaq data described in Section 2 of the paper. η_p, $p = 1, \ldots, d$, is estimated by the procedure from Section 1.3, while the integrated quarticity is approximated by the squared sub-sampled five-minute realized variance. η_p, $p = 1, \ldots, d$, is estimated as the average number of mid-quote revisions per second.

p	1	2	3	4	5	6–10	11	12	13	14	15
ξ_p	0.381	0.485	0.692	0.938	1.181	1.585	1.907	1.992	2.143	2.305	2.495
λ_p	3.765	1.259	0.793	0.696	0.589	0.460	0.382	0.221	0.138	0.099	0.063

2.2 Additional Results

To examine whether ensuring positive semi-definite spot covariance matrix estimates yields an improved finite sample performance, Table 3 reports simulation results analogous to Table 1 in the paper for the LMM estimator augmented by a truncation of negative eigenvalues of the estimates at zero. Most importantly, finite sample precision improves, indeed, with MIFB being somewhat lower than above for the corresponding optimal values of the input parameters. The latter also imply that the number of blocks (on average) decreases to 69, while the spectral cutoff increases to 22. Interestingly, for the (co-)variance estimates, the MIFB-optimal values of the inputs now yield a performance that is closer to the one implied by MISE$_c$- and MISE$_v$-optimal inputs than before.

While the above setting may represent "regular" trading day days, we additionally consider an augmented framework to proxy "unusual" days as analyzed in Section 5.3 in the paper. For that purpose, we extend the process in (24a), (24b) and (25) to allow for volatility (co-)jumps in a parsimonious way, modifying the first factor of the stochastic volatility components (25) to

$$dv_{1,t}^{(p)} = \alpha_1 v_{1,t}^{(p)} dt + dW_{1,t}^{(p)} + \mathcal{I}_{t^*}^{(p)} \mathcal{J}^{(p)}, \quad p = 1, \ldots, d, \tag{27}$$

where $\mathcal{I}_{t^*}^{(p)}$ is an indicator, taking the value one if the p-th stochastic volatility component jumps at one certain time t^*. For each component, we simulate such a jump with probability $q_{\mathcal{J}} = 0.75$, while the jump sizes $\mathcal{J}^{(p)}$ satisfy for all p, $\mathcal{J}^{(p)} \sim \text{Exp}(\beta_1/\ln(5))$. We set $t^* = 0.6$, corresponding to around 1:30 pm, while the expected jump size implies an (expected) five-fold increase in volatility due to the jump compared to the prevailing level. The latter is motivated by the empirical findings in Section 5.3 in the paper. Table 4 provides the simulation results for the modified process based on the LMM spot covariance matrix estimator with truncation of negative eigenvalues. Despite the considerable (expected) size of the jumps, finite sample precision suffers only a little compared to Table 3 as MIFB implied by the optimal input values

Table 3: Performance of LMM spot covariance matrix estimator with truncation of negative eigenvalues depending on θ_h, θ_J and θ_K. Positve semi-definiteness of spot covariance matrix estimates is ensured by truncation of negative eigenvalues at zero. See caption of Table 1 in the paper for further definitions.

	Full Covariance Matrix				
	θ_h	θ_J	θ_K	RMIFB	% PSD
Opt	0.175	7.000	2.000	23.995	100.000
	0.175	1.000	2.000	37.655	100.000
	0.175	10.000	2.000	24.006	100.000
	0.175	7.000	1.200	24.701	100.000
	0.175	7.000	4.800	26.099	100.000
	0.025	7.000	2.000	37.911	100.000
	0.250	7.000	2.000	24.657	100.000
	Covariances				
	θ_h	θ_J	θ_K	RMISE$_c$	
Opt	0.150	9.000	2.000	23.881	
Opt*	0.175	7.000	2.000	23.935	
	Variances				
	θ_h	θ_J	θ_K	RMISE$_v$	
Opt	0.225	5.000	1.600	16.854	
Opt*	0.175	7.000	2.000	17.584	

increases by only around 1.3 percentage points. The optimal input values themselves remain almost unchanged with only the spectral cutoff reducing to 19. These findings suggests that the LMM estimator is able to retrieve spot covariance matrix paths in a setting with considerable volatility (co-)jumps.

Table 4: Performance of LMM spot covariance matrix estimator with truncation of negative eigenvalues depending on θ_h, θ_J and θ_K in setting with volatility (co-)jumps. Positve semi-definiteness of spot covariance matrix estimates is ensured by truncation of negative eigenvalues at zero. The simulated process is given by (24a), (24b) and with (25) modified according to (27). See caption of Table 1 in the paper for further definitions.

	Full Covariance Matrix				
	θ_h	θ_J	θ_K	RMIFB	% PSD
Opt	0.175	6.000	2.000	25.211	100.000
	0.175	1.000	2.000	36.301	100.000
	0.175	10.000	2.000	25.238	100.000
	0.175	6.000	1.200	25.689	100.000
	0.175	6.000	4.800	27.545	100.000
	0.025	6.000	2.000	41.122	100.000
	0.250	6.000	2.000	26.116	100.000
	Covariances				
	θ_h	θ_J	θ_K	RMISE_c	
Opt	0.150	7.000	2.000	24.824	
Opt*	0.175	6.000	2.000	24.907	
	Variances				
	θ_h	θ_J	θ_K	RMISE_v	
Opt	0.250	4.000	1.600	19.132	
Opt*	0.175	6.000	2.000	19.985	

3 Summary Statistics for Quote Data

Table 5: 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 $\tilde{Q} = 50$ according to the procedure outlined in Section 1.3. $\overline{\hat{\xi}^*}$: avg. noise-to-signal ratio per observation, where $\hat{\xi}^{2*} = n\hat{\eta}^*/\mathrm{RV}^{ss}_{5m}$ with RV^{ss}_{5m} denoting the subsampled five-minute realized variance. $\overline{\hat{R}^*}$: avg. estimate of order of serial dependence in the noise process.

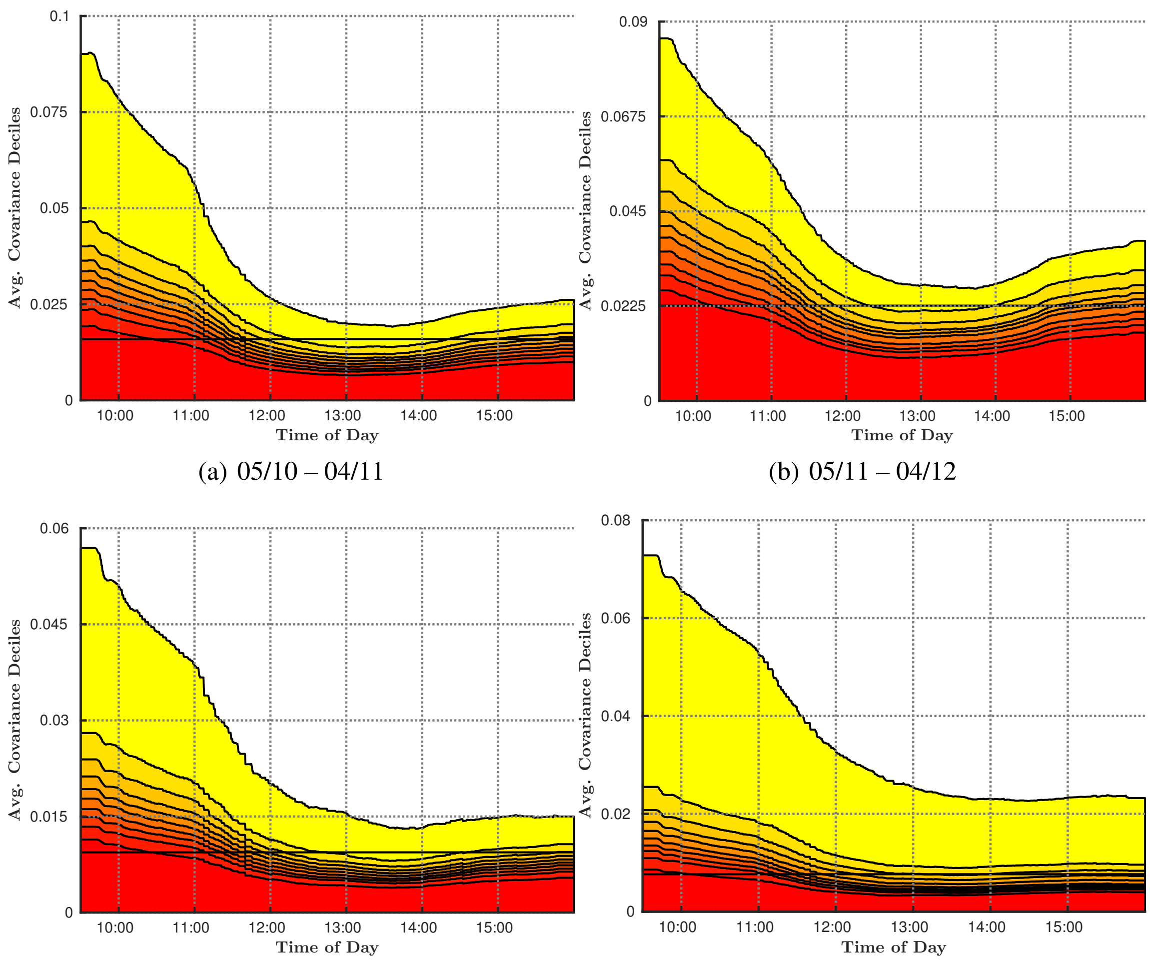
Symbol	$\overline{n}$	$\overline{\Delta t}$	$\%_{(\|\Delta Y\|>0)}$	$\overline{\Delta t_{(\|\Delta Y\|>0)}}$	$\overline{\sqrt{\hat{\eta}}}$	$\overline{\hat{\xi}}$	$\overline{\hat{R}}$
AAPL	184792.224	0.180	51.864	0.344	0.004	0.582	12.246
MSFT	444262.806	0.063	0.913	7.499	0.116	0.835	3.015
GOOG	62548.169	0.451	53.530	0.870	0.038	1.364	11.407
CSCO	295700.063	0.092	0.748	13.845	0.070	0.463	1.727
ORCL	326370.371	0.088	1.670	5.711	0.217	1.222	4.056
INTC	352591.320	0.078	0.818	10.344	0.054	0.442	1.921
QCOM	338279.341	0.082	4.925	1.952	0.280	2.425	8.880
AMGN	142855.142	0.232	14.910	1.863	0.193	2.170	6.364
TEVA	117498.898	0.286	9.559	3.275	0.183	2.116	5.133
GILD	192972.910	0.145	7.927	2.453	0.268	2.219	7.099
AMZN	102312.847	0.326	49.525	0.655	0.052	0.983	11.744
EBAY	283542.140	0.097	4.103	2.941	0.392	1.992	6.707
NWSA	201062.913	0.156	1.385	11.859	0.253	0.894	1.950
DTV	165357.795	0.188	7.270	2.988	0.226	1.920	4.745
CELG	82505.222	0.359	27.136	1.458	0.178	1.869	6.823
INFY	65839.135	0.555	26.248	2.318	0.137	2.495	4.880
YHOO	253667.347	0.132	1.268	12.627	0.181	0.688	2.770
COST	77445.868	0.390	23.667	1.807	0.121	2.392	6.311
SPLS	139319.424	0.198	1.406	16.492	0.113	0.507	1.516
SBUX	180134.971	0.157	7.213	2.769	0.295	1.992	6.460
ADP	115795.086	0.272	8.680	3.660	0.148	2.175	3.906
ESRX	138732.103	0.232	13.287	1.869	0.172	1.792	5.784
ADBE	169761.579	0.165	5.233	4.020	0.199	1.377	3.686
FSLR	78021.596	0.411	33.772	1.407	0.731	1.895	8.253
BIIB	53285.482	0.562	39.584	1.515	0.160	2.007	4.916
SYMC	154504.801	0.180	1.436	13.961	0.137	0.696	1.677
JNPR	168600.474	0.194	5.513	5.147	0.271	1.141	3.399
BRCM	250452.381	0.114	4.366	3.398	0.412	1.980	6.123
AMAT	156708.366	0.173	0.905	21.231	0.082	0.381	0.943
CMCSA	270851.740	0.104	1.823	6.515	0.179	1.033	3.327
Avg.	185525.750	0.222	13.689	5.560	0.195	1.468	5.259

4 Summary Statistics for Inputs

Table 6: 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 3.4 of the paper, using the input parameters that are optimal in the simulation study of Section 4 of the paper: $\theta_h = 0.175$, $\theta_J = 7$ and $\theta_K = 2$.

Sample	Input	$q_{0.05}$	Mean	$q_{0.95}$	Std.
05/10	$\lceil h_n^{-1} \rceil$	21.000	27.230	35.000	4.849
–	J_n	43.000	47.957	53.000	3.223
04/14	K_n	5.000	5.946	7.000	0.702
05/10	$\lceil h_n^{-1} \rceil$	22.000	28.566	37.000	5.443
–	J_n	44.000	48.864	54.000	3.350
04/11	K_n	5.000	6.136	7.500	0.754
05/11	$\lceil h_n^{-1} \rceil$	22.000	28.299	37.000	5.678
–	J_n	44.000	48.660	54.000	3.573
04/12	K_n	5.000	6.100	7.500	0.822
05/12	$\lceil h_n^{-1} \rceil$	20.000	24.509	31.000	3.181
–	J_n	42.000	45.954	51.000	2.557
04/13	K_n	5.000	5.544	6.500	0.495
05/13	$\lceil h_n^{-1} \rceil$	23.000	27.273	32.000	2.972
–	J_n	45.000	48.180	52.000	2.189
04/14	K_n	5.500	5.966	6.500	0.457

5 Intraday Behavior of Spot (Co-)Variances Year-by-Year



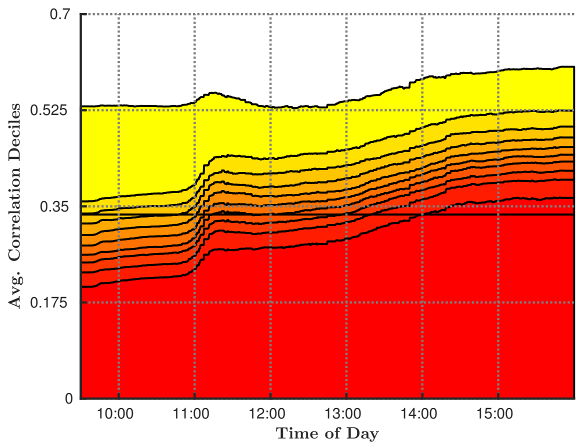
(a) 05/10 – 04/11

(b) 05/11 – 04/12

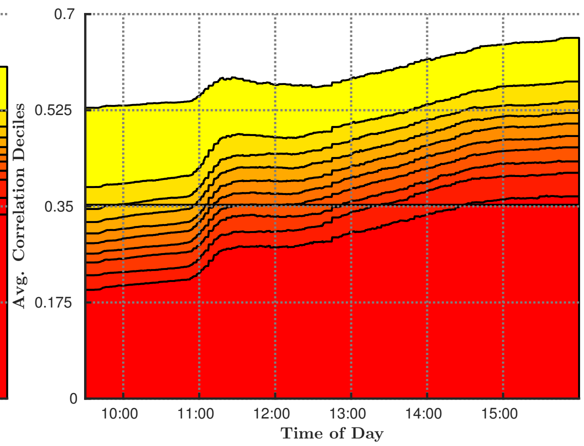
(c) 05/12 – 04/13

(d) 05/13 – 04/14

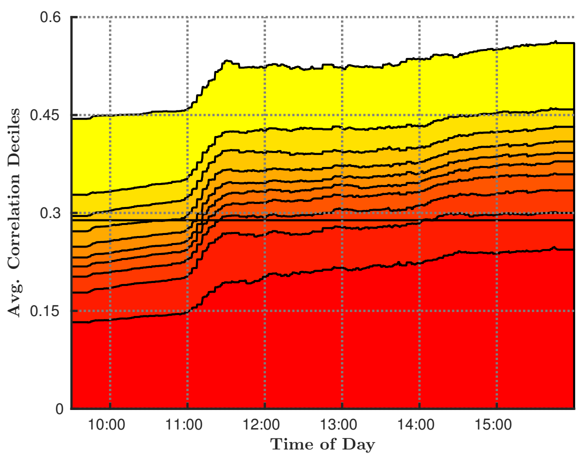
Figure 1: Cross-sectional deciles of across-day averages of spot covariances. Spot estimates are first averaged across days for each asset pair. Subsequently, cross-sectional sample deciles of the across-day averages are computed. Solid horizontal line corresponds to the cross-sectional median of the across-day averages of *integrated* covariance 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. “Unusual days” discussed in Section 5.3 of the paper as well as days with scheduled FOMC announcements are removed. Covariances are annualized.



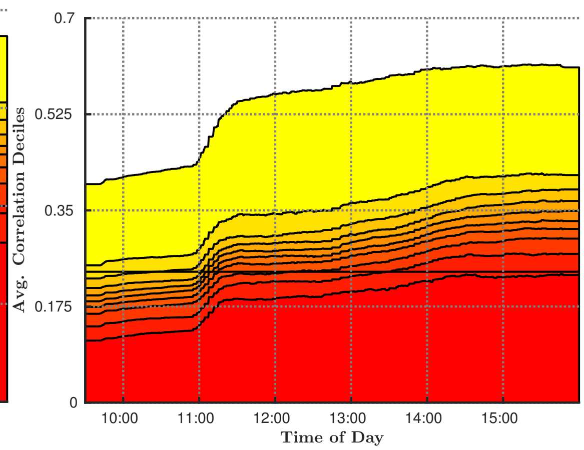
(a) 05/10 – 04/11



(b) 05/11 – 04/12

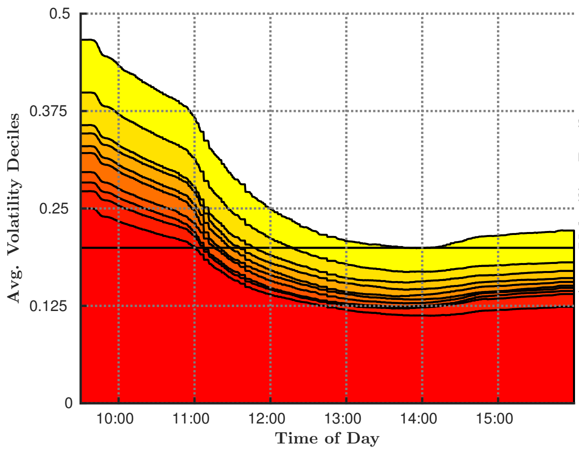


(c) 05/12 – 04/13

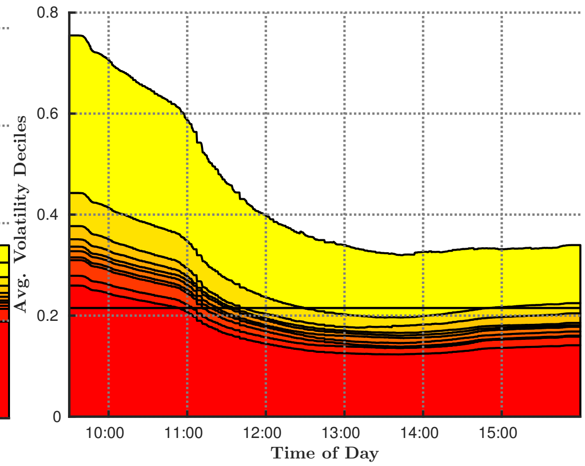


(d) 05/13 – 04/14

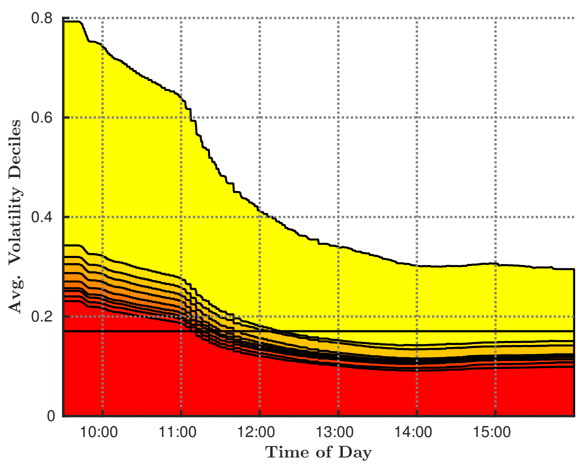
Figure 2: Cross-sectional deciles of across-day averages of spot correlations. Spot estimates are first averaged across days for each asset pair. Subsequently, cross-sectional sample deciles of the across-day averages are computed. Solid horizontal line corresponds to the cross-sectional median of the across-day averages of *integrated* 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. “Unusual days” discussed in Section 5.3 of the paper as well as days with scheduled FOMC announcements are removed.



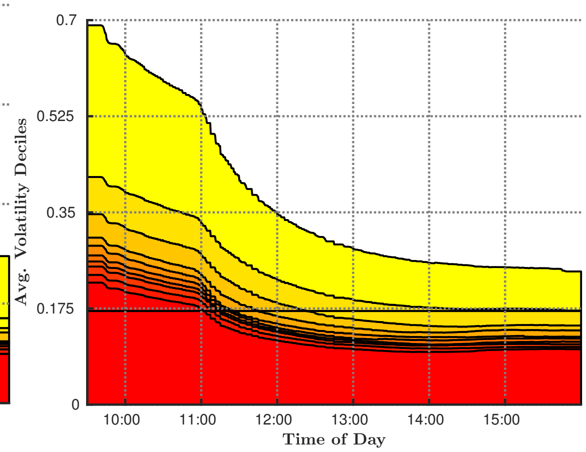
(a) 05/10 – 04/11



(b) 05/11 – 04/12

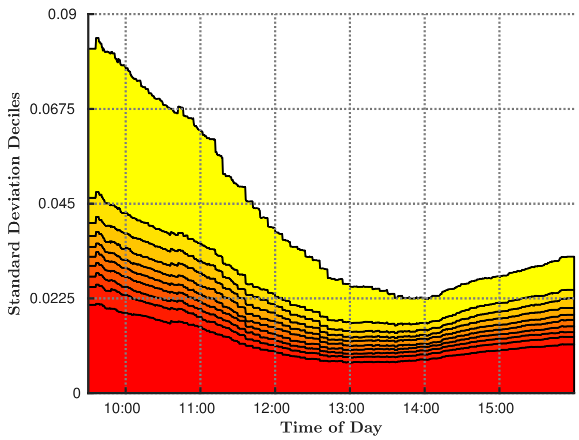


(c) 05/12 – 04/13

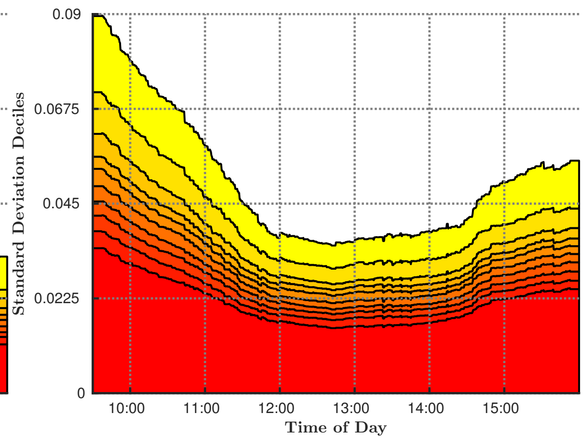


(d) 05/13 – 04/14

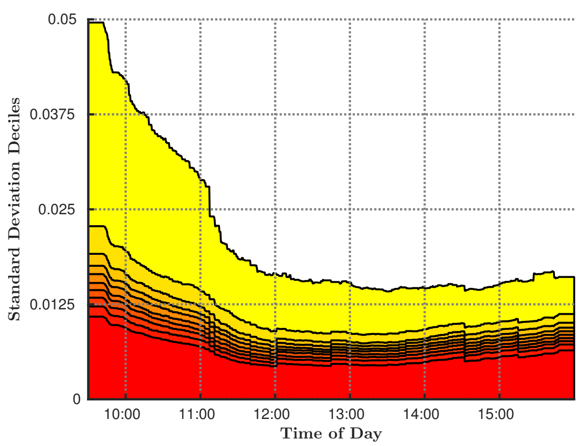
Figure 3: Cross-sectional deciles of across-day averages of spot volatilities. Spot estimates are first averaged across days for each asset. Subsequently, cross-sectional sample deciles of the across-day averages are computed. Solid horizontal line corresponds to the cross-sectional median of the across-day averages of *integrated* 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. "Unusual days" discussed in Section 5.3 of the paper as well as days with scheduled FOMC announcements are removed. Volatilities are annualized.



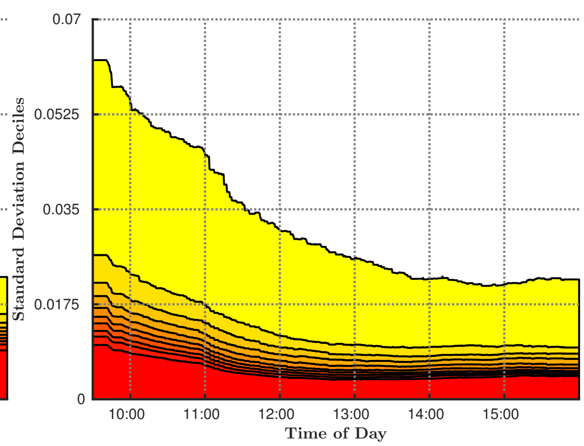
(a) 05/10 – 04/11



(b) 05/11 – 04/12

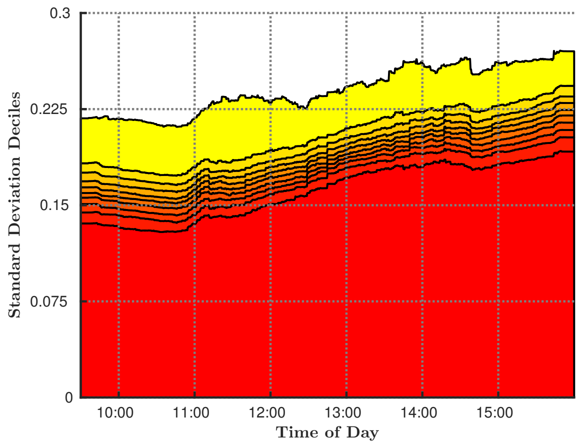


(c) 05/12 – 04/13

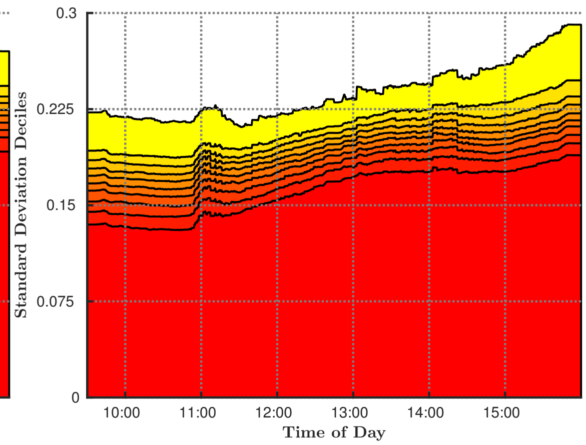


(d) 05/13 – 04/14

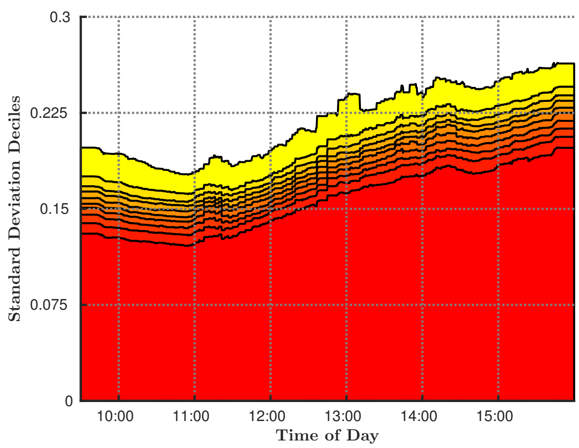
Figure 4: Cross-sectional deciles of across-day standard deviations of spot covariances. First, sample standard deviations of spot estimates are computed across days for each asset pair. Subsequently, cross-sectional sample deciles of the across-day standard deviations are computed. “Unusual days” discussed in Section 5.3 of the paper as well as days with scheduled FOMC announcements are removed. Covariances are annualized.



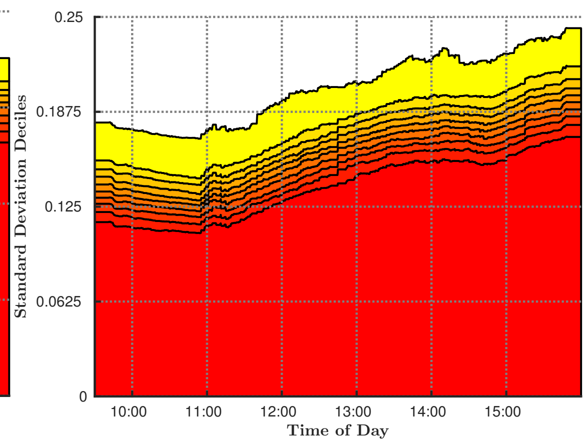
(a) 05/10 – 04/11



(b) 05/11 – 04/12

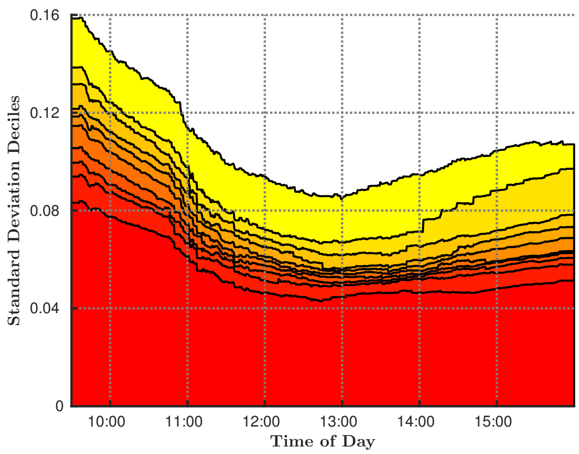


(c) 05/12 – 04/13

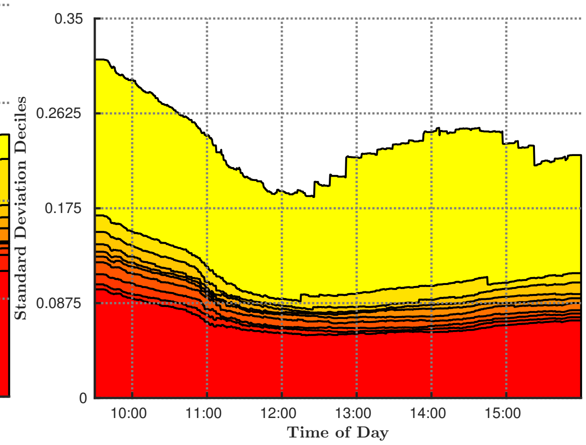


(d) 05/13 – 04/14

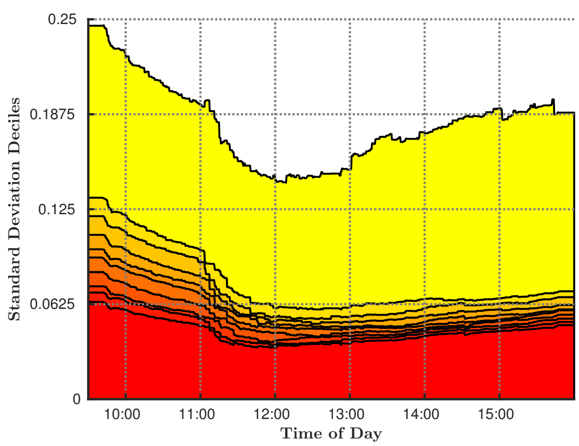
Figure 5: Cross-sectional deciles of across-day standard deviations of spot correlations. First, sample standard deviations of spot estimates are computed across days for each asset pair. Subsequently, cross-sectional sample deciles of the across-day standard deviations are computed. “Unusual days” discussed in Section 5.3 of the paper as well as days with scheduled FOMC announcements are removed.



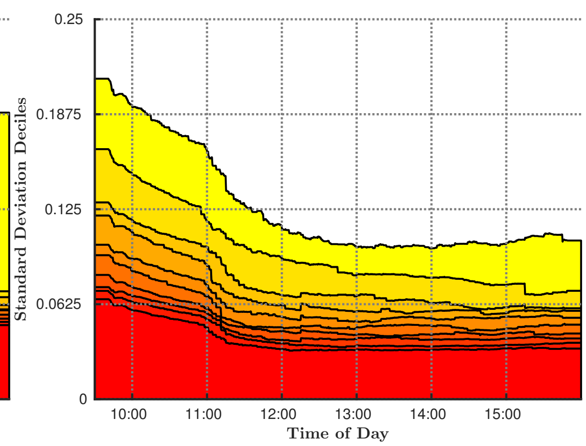
(a) 05/10 – 04/11



(b) 05/11 – 04/12



(c) 05/12 – 04/13

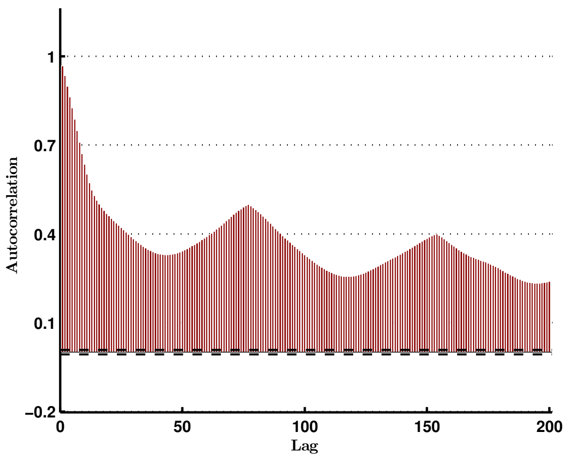


(d) 05/13 – 04/14

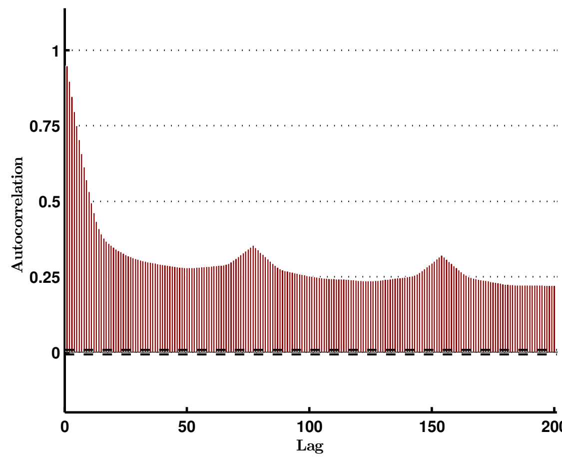
Figure 6: Cross-sectional deciles of across-day standard deviations of spot volatilities. First, sample standard deviations of spot estimates are computed across days for each asset. Subsequently, cross-sectional sample deciles of the across-day standard deviations are computed. “Unusual days” discussed in Section 5.3 of the paper as well as days with scheduled FOMC announcements are removed. Volatilities are annualized.

6 Sample Autocorrelation Functions of Spot Estimates

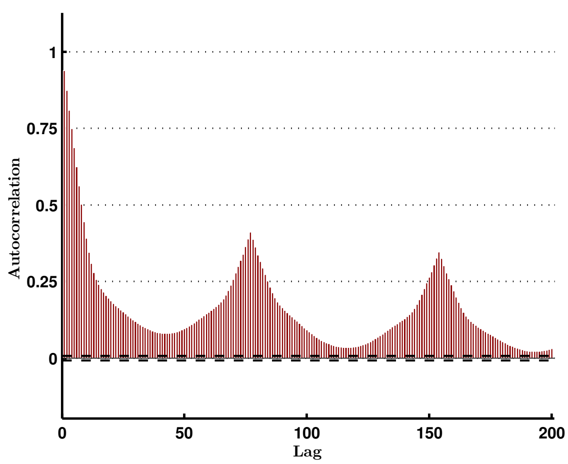
Figure 7 reports the (averaged) autocorrelation functions (ACFs) of spot covariances, correlations and volatilities with the figures being constructed such that one lag corresponds to approximately five minutes. We observe that all (co-)variability measures are strongly serially correlated across short 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, this pattern is less pronounced. Here, long-term autocorrelations stabilize around 0.25 and decay very slowly.



(a) Covariance



(b) Correlation



(c) Volatility

Figure 7: Avg. ACFs of spot covariance, correlation 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 ($\pm 1.96/\sqrt{n}$).

7 Another Unusual Day

We analyze intraday risk on 12/27/12 using the adaptive one-sided version of the estimator (12*) as outlined in the first paragraph of Section 5.3 of the paper. On this day, 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.[3] As shown in Figure 8, 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. Figure 9 shows that, 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. Consequently, the positive “fiscal cliff news” lead to a significant increase in spot correlations that lasted until the end of the trading day. The estimated spot covariance, correlation and volatility paths for AAPL and AMZN in Figure 10 demonstrate that analogous patterns can be observed for a particular asset pair.

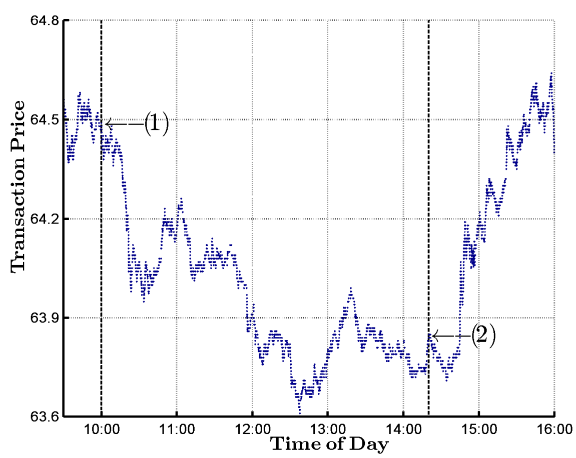
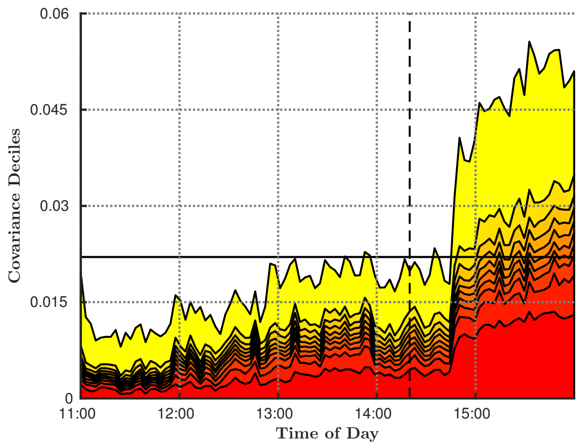
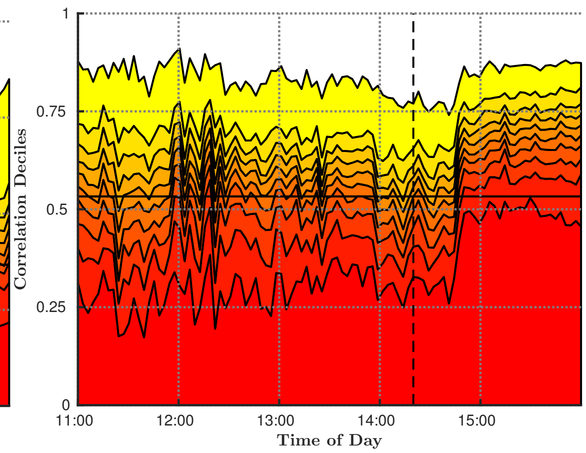


Figure 8: QQQ transaction prices (12/27/12). (1): Senate Majority Leader states that resolution to “fiscal cliff” crisis before January 1, 2013, unlikely. (2): News that the House of Representatives will convene on the following Sunday in an attempt to end the “fiscal cliff” crisis.

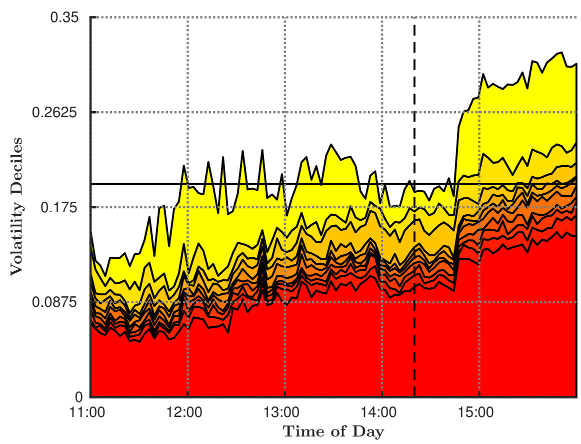
[3]See http://money.msn.com/now/post.aspx?post=73878f29-4fdb-45c5-baf6-0f83f40-b821c&_p=986b65a2-3eea-479c-a4a0-ba9d3988b0e0.



(a) Covariances

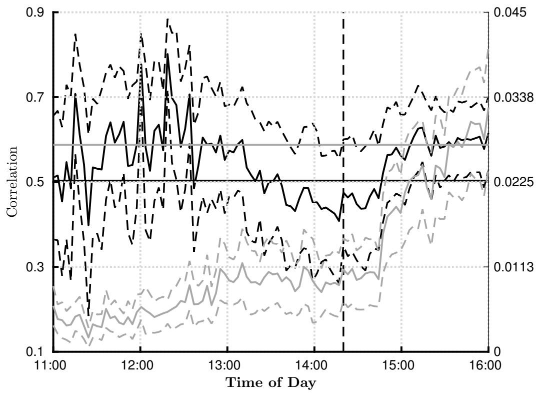


(b) Correlations

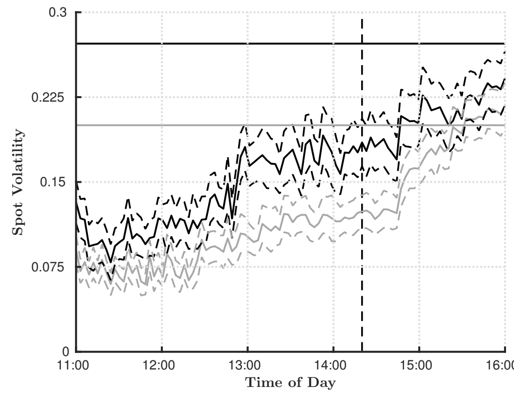


(c) Volatilities

Figure 9: Cross-sectional deciles of spot covariances, correlations and volatilities (12/27/12). Solid horizontal line corresponds to the cross-sectional median of *integrated* covariance, correlation 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. Covariances and volatilities are annualized.



(a) Covariances/Correlations



(b) Volatilities

Figure 10: Spot covariances, correlations and volatilities for AAPL and AMZN (04/23/13). In left plot, black lines (and left y-axis) represent correlations, grey lines (and right y-axis) covariances. In right plot, black lines are for AMZN, grey lines for AAPL. Dashed lines correspond to approximate pointwise 95% confidence intervals according to Corollary 1 in the paper. Horizontal lines correspond to the cross-sectional median of *integrated* covariance, correlation 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. Covariances and volatilities are annualized.

8 Code

The web supplement contains the MATLAB code developed for this paper in the archive `Code.zip`. The text file `README.txt` details the scripts required to obtain the main results of the paper as well as the order in which they should be executed.

Note that we cannot share the message and order book data used for the empirical analysis as it is proprietary. A LOBSTER subscription can be obtained at `https://lobsterdata.com/`.

References

Altmeyer, R. and Bibinger, M., 2015. Functional stable limit theorems for quasi-efficient spectral covolatility estimators, *Stochastic Processes and their Applications*, 125 (12), 4556–4600.

Andersen, T.G. and Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance*, 4 (2-3), 115–158.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., and Shephard, N., 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 (2), 149–169.

Bibinger, M., Hautsch, N., Malec, P., and Reiß, M., 2014. Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency, *Annals of Statistics*, 42 (4), 1312–1346.

Billingsley, P., 1991. *Probability and Measure*, Springer, New York, 2nd ed.

Chernov, M., Gallant, A.R., Ghysels, E., and Tauchen, G., 2003. Alternative models for stock price dynamics, *Journal of Econometrics*, 116 (12), 225 – 257.

Gallant, A.R., 1981. On the bias in flexible functional forms and an essentially unbiased form : The fourier flexible form, *Journal of Econometrics*, 15 (2), 211–245.

Hansen, P.R. and Lunde, A., 2006. Realized variance and market microstructure noise, *Journal of Business & Economic Statistics*, 24 (2), 127–161.

Huang, X. and Tauchen, G., 2005. The relative contribution of jumps to total price variance, *Journal of Financial Econometrics*, 3 (4), 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*.

Zhang, L., Mykland, P.A., and Ait-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, 100 (472), 1394–1411.