AARHUS
UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2011-35

Flat-Top Realized Kernel Estimation of Quadratic Covariation with Non-Synchronous and Noisy Asset Prices

Rasmus Tangsgaard Varneskov

Department of Economics and Business Aarhus University Bartholins Allé 10 DK-8000 Aarhus C Denmark	Email: oekonomi@au.dk Tel: +45 8942 1610

Flat-Top Realized Kernel Estimation of Quadratic Covariation with Non-Synchronous and Noisy Asset Prices*

Rasmus Tangsgaard Varneskov†

Aarhus University and CREATES

First draft: September 27th 2011; This version: April 30th 2013

Abstract

This paper develops a general, multivariate additive noise model for synchronized asset prices and provides a multivariate extension of the generalized flat-top realized kernel estimators, analyzed in Varneskov (2013), to estimate its quadratic covariation. The additive noise model allows for α-mixing dependent exogenous noise, random sampling, and an endogenous noise component that encompasses synchronization errors, lead-lag relations, and diurnal heteroskedasticity. The various components may exhibit polynomially decaying autocovariances. In this setting, the class of estimators is consistent, asymptotically unbiased, and mixed Gaussian at the optimal rate of convergence, $n^{1/4}$. A simple finite sample correction based on projections of symmetric matrices ensures positive definiteness without altering the asymptotic properties of the estimators. The finite sample correction admits non-linear transformations of the estimated covariance matrix such as correlations and realized betas, which inherit the desirable asymptotic properties of the flat-top realized kernels. An empirically motivated simulation study assesses the choice of sampling scheme and projection rule, and it shows that flat-top realized kernels have a desirable combination of robustness and efficiency relative to competing estimators. Last, an empirical analysis of correlations and market betas for a portfolio of six stocks of varying size and liquidity illustrates the use and properties of the new estimators.

Keywords: Bias Reduction, Non-Synchronicity, Nonparametric Estimation, Market Microstructure Noise, Quadratic Covariation.

JEL classification: C14, C15, C58

*Some of this paper was written while I was visiting the Oxford-Man Institute, University of Oxford, whose hospitality I am grateful for. I wish to thank Neil Shephard for hosting me and for providing helpful comments. Furthermore, I wish to give thanks to Asger Lunde for providing cleaned high frequency data and to Bent Jesper Christensen, the joint editor, Jonathan Wright, an associate editor, two anonymous referees along with seminar participants at CREATES for many useful comments and suggestions. All the coding was done in Gauss 9.0 and 11.0. Financial support from Aarhus School of Business and Social Sciences, Aarhus University, CREATES, funded by the Danish National Research Foundation, and from Købmand Ferdinand Sallings Mindefond is gratefully acknowledged.

†Department of Economics and Business, Aarhus University, 8210 Aarhus V., Denmark. Email: rvarneskov@creates.au.dk.

1 Introduction

The covariation of asset returns is central to many aspects of financial economics such as, e.g., correlation analysis, portfolio selection, and risk management. Hence, with the availability of high-frequency financial data, the literature has seen a surge in the development of various econometric techniques, in particular a shift from parametric, conditional covariance estimation based on lower frequency observations (e.g. daily) towards virtually model-free measures and realized estimators, to extract gains from the vast amount of intra-daily information. The underlying idea is to use quadratic covariation as an ex-post covariance measure, whose increments can be studied to learn about of the dependence of asset returns. Andersen, Bollerslev, Diebold & Labys (2003) and Barndorff-Nielsen & Shephard (2004) pioneered the work on non-parametric estimation of quadratic covariation by introducing the realized covariance estimator, which in a Brownian semimartingale setting with synchronous trading is both a consistent and efficient estimator of quadratic covariation as the sampling interval progressively shrinks. Two empirical phenomena, however, render sampling at the highest frequencies undesirable and complicate inference about the covariance of asset returns. First, tick-by-tick trading of various assets occur randomly and non-synchronously. Second, the observed prices are affected by market microstructure (MMS) noise, which summarizes an array of market imperfections such as bid-ask bounce effects, asymmetric information and strategic learning, execution of block trades, etc.

Since the comprehensive study of stock returns by Epps (1979), non-synchronous trading has been recognized to bias empirical correlation statistics towards zero when sampling beyond the intra-hour level, the so-called Epps effect. Large (2007), however, provides a simple model showing that part of this bias may be due to delayed information transmission in financial markets, causing temporal lead-lag relations among correlated assets, and that this may be interpreted as a source of additive, endogenous MMS noise. Hence, when Martens (2004) and Hayashi & Yoshida (2005) propose refresh time sampling and a pseudo-aggregation algorithm of non-synchronous observations, respectively, to correct the realized covariance estimator for the Epps effect, they are, first of all, only accounting for one potential source of the bias, and secondly, they do not consider other forms of MMS noise.

Several recent papers have proposed techniques for simultaneously dealing with non-synchronicity and MMS noise models of varying complexity. These include non-parametric estimators such as the inconsistent, but bias-corrected Hayashi-Yoshida estimator in Voev & Lunde (2007), the two-scale realized covariance (TSRC) estimator in Zhang (2011), the realized kernel in Barndorff-Nielsen, Hansen, Lunde & Shephard (2011), the pre-averaged realized covariance estimator in Christensen, Kinnebrock & Podolskij (2010), and the parametric quasi maximum likelihood (QML) estimators in Aït-Sahalia, Fan & Xiu (2010) and Shephard & Xiu (2013). Among these, the pre-averaged realized covariance estimator and the QML estimators are shown to be consistent, asymptotically unbiased, and mixed Gaussian at the optimal rate convergence, $n^{1/4}$, n being the number of synchronized observations, whereas the TSRC estimator is consistent at the suboptimal rate $n^{1/6}$. However, these estimators all rely on high-level MMS noise assumptions, cf. exogenous and i.i.d., to achieve their asymptotic properties. Hansen & Lunde (2006) show that this is not too damaging if sampling occurs

around every minute (or every 15 ticks), but they also stress that the high-level noise assumption is violated for tick-by-tick data where the MMS noise is characterized by being serially dependent, endogenous, and exhibiting diurnal heteroskedasticity. Similar findings have been made by de Jong & Nijman (1997), Large (2007), Voev & Lunde (2007), Kalnina & Linton (2008), Ubukata & Oya (2009), Aït-Sahalia, Mykland & Zhang (2011), and Kalnina (2011). Diebold & Strasser (2012), in a comprehensive econometric analysis of MMS models, further argue that a general noise model, allowing for both exogenous and endogenous components with polynomially decaying autocovariances, is needed to avoid concerns about the underlying MMS mechanisms. Aware of these empirical regularities at tick-by-tick frequencies, the realized kernel is the first consistent estimator that allows for both non-synchronous trading and general assumptions on the MMS noise. However, the estimator also suffers from a bias in the asymptotic distribution and a suboptimal rate of convergence $n^{1/5}$. This is unfortunate as it implies a bias in non-linear transformations of the estimated covariance matrix, e.g. realized correlations and regression coefficients, and moreover, since Laurent, Rombouts & Violante (2012) show that consistent ranking of multivariate volatility models requires an unbiased estimate of the asset return covariance matrix, thus limiting the potential applications of the realized kernel. Hence, in its present state, there is room in the literature for an asymptotically unbiased and rate-optimal estimator of quadratic covariation, which accounts for non-synchronous trading and is valid under general MMS noise assumptions such that it may utilize all available high-frequency observations.

In a contemporaneous paper, Varneskov (2013) analyzes a generalized class of univariate flat-top realized kernels under weak assumptions on the MMS noise and equally spaced observations, and he establishes optimal asymptotic properties such as consistency, asymptotic unbiasedness, and mixed Gaussianity at the optimal rate of convergence, $n^{1/4}$. If optimally designed, the estimators are also efficient in a Cramér-Rao sense. This paper extends the analysis to a multivariate setting where it provides the following incremental theoretical contributions in addition to the simulation study and empirical analysis: (1) It develops a general, multivariate additive noise model for synchronized asset prices, which allows for α-mixing dependent exogenous noise, random sampling, and an endogenous noise component that encompasses synchronization errors, asymmetric lead-lag relations, and diurnal heteroskedasticity. The various components may exhibit polynomially decaying autocovariances. (2) It extends the generalized class of univariate flat-top realized kernels to the multivariate case and establishes similar optimal asymptotic properties in the present setting. (3) It analyzes element-wise estimation of the covariance matrix using various synchronization schemes, and the impact of the latter on the properties of the noise process. (4) It proposes a matrix regularization technique to ensure positive definiteness in finite samples, which affects neither consistency nor the asymptotic distribution. (5) It considers non-linear transformations of the estimated covariance matrix.

Ikeda (2011, 2013) analyzes a two-scale realized kernel, which may be interpreted as a realized kernel estimator with a generalized jack-knife kernel function. Under general conditions on the MMS noise, though stronger than in the present paper, the two-scale realized kernel is shown to posses first-order asymptotic properties similar to those of the flat-top realized kernels. The latter, however, has a higher-

order advantage in terms of bias reduction. Inspired by the underlying jack-knife bias-correction, a new bias-corrected pre-averaged realized covariance estimator is proposed, which is robust to serial dependence in the MMS noise. The finite sample performance of the flat-top realized kernels relative to these two robust competitors and the realized kernel is studied in a three-asset simulation setup, which is inspired by the high-frequency data used in the empirical analysis. Though not uniformly efficient for all combinations of noise, non-synchronicity, and transformations of the covariance matrix, the flat-top realized kernel is uniformly efficient for estimators that offer a stable bias control for all data generating processes, thus showing a desirable combination of robustness and efficiency, which illustrates both its optimal asymptotic properties and higher-order advantage. An empirical analysis of correlations and market beta estimates for a portfolio of six stocks with varying liquidity and size shows exactly the same pattern as the simulation study, thus reinforcing the conclusions.

The paper proceeds as follows. Section 2 develops the additive noise model. Section 3 analyzes the theoretical properties of flat-top realized kernels, and Section 4 studies a finite sample correction and non-linear transformations. Section 5 and 6 contain the simulation and empirical study, respectively. Last, Section 7 concludes. The appendix contains additional mathematical details and proofs. Furthermore, the following notation is used throughout: $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{N}$ denote the set of real numbers, integers and natural numbers; $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$ and $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$; $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function; $\|\cdot\|$ denotes the Euclidean (matrix) norm; $\otimes$ denotes the Kronecker product; $O(\cdot)$, $o(\cdot)$, $O_p(\cdot)$ and $o_p(\cdot)$ denote the usual (stochastic) orders of magnitude; "$\to$", "$\xrightarrow{\mathbb{P}}$", "$\xrightarrow{d}$" and "$\xrightarrow{d_s}$" indicate the limit, the probability limit, convergence in law and stable convergence in law, respectively.

2 Theoretical Setup

This section develops the additive noise model and provides the necessary assumptions to conduct the theoretical analysis. Moreover, the MMS noise specification is motivated by interpreting synchronization errors as additive noise.

2.1 The Semimartingale Process

Let a finite, d-dimensional vector of no-arbitrage logarithmic asset prices, $\boldsymbol{p}^*$, be defined on some filtered probability space $(\mathcal{O}, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where $(\mathcal{F}_t) \subseteq \mathcal{F}$ is an increasing family of σ-fields satisfying $\mathbb{P}$-completeness, right continuity and is assumed to be generated by other filtrations $\mathcal{P}_t$, the σ-algebra generated by $\boldsymbol{p}_t^*$, $\mathcal{H}_t$, the σ-algebra generated by $\boldsymbol{p}_t^c = (\boldsymbol{p}_t^*, \breve{\boldsymbol{p}}_t)'$ where $\boldsymbol{p}_t^*$ and $\breve{\boldsymbol{p}}_t$ are uncorrelated such that $\mathcal{P}_t \subset \mathcal{H}_t$, and $\mathcal{G}_t$ where $\mathcal{H}_t \perp\!\!\!\perp \mathcal{G}_s \ \forall t, s$ as $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{G}_t$, the smallest σ-field containing $\mathcal{H}_t \cup \mathcal{G}_t$. The triplet $(\mathcal{P}_t, \mathcal{H}_t, \mathcal{G}_t)$ is used to define a space for observable prices that allows an efficient price process to be contaminated by both endogenous and exogenous MMS noise components, defined on $\mathcal{H}_t$ and $\mathcal{G}_t$, respectively. Throughout the remainder of the paper, and without loss of generality, let $t \in [0, 1]$, which can be thought of as the passing of an economic event, e.g. a trading day. Formally, and following literature standards, let $\boldsymbol{p}^*$ follow a continuous time Brownian semimartingale process

of the form

$$\boldsymbol{p}_t^* = \boldsymbol{p}_0^* + \int_0^t \boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{\Sigma}_s^{1/2} d\boldsymbol{W}_s \tag{1}$$

where $\boldsymbol{W}_t \in \mathbb{R}^k$ is a vector of independent standard Brownian motions, $\boldsymbol{\mu}_t \in \mathbb{R}^{d\times 1}$ is a $(\mathcal{P}_t)$-predictable stochastic process, and $\boldsymbol{\Sigma}_t^{1/2} \in \mathbb{R}^{d\times k}$ is a $(\mathcal{P}_t)$-adapted matrix-valued stochastic volatility process. The theoretical analysis requires the following additional structure on (1):

Assumption 1. *Let* $vec(\boldsymbol{\Sigma}_t^{1/2}) \in \mathbb{R}^{dk\times 1}$ *follow a continuous time Brownian semimartingale process of the form* $vec(\boldsymbol{\Sigma}_t^{1/2}) = vec(\boldsymbol{\Sigma}_0^{1/2}) + \int_0^t \tilde{\boldsymbol{\mu}}_s ds + \int_0^t \tilde{\boldsymbol{\Sigma}}_s^{1/2} dW_s$ *where* $\tilde{\boldsymbol{\mu}}_t \in \mathbb{R}^{dk\times 1}$ *is* $(\mathcal{P}_t)$*-predictable,* $\tilde{\boldsymbol{\Sigma}}_t^{1/2} \in \mathbb{R}^{dk\times k}$ *is* $(\mathcal{P}_t)$*-adapted, and both* $\tilde{\boldsymbol{\mu}}_t$ *and* $\tilde{\boldsymbol{\Sigma}}_t^{1/2}$ *are càdlàg processes.*

Assumption 2. $\forall (t,w) \in [0,1]\times\mathcal{O}\ \exists \Lambda_1 > 0 : \|\boldsymbol{\mu}_t(w)\| + \|\tilde{\boldsymbol{\mu}}_t(w)\| + \|\boldsymbol{\Sigma}_t^{1/2}(w)\| + \|\tilde{\boldsymbol{\Sigma}}_t^{1/2}(w)\| \leq \Lambda_1$.

This setup follows recent contributions in the literature such as Barndorff-Nielsen et al. (2011), Christensen et al. (2010), and Ikeda (2011), and it nests most continuous time models in financial economics, see Andersen & Benzoni (2012) for a review. The Brownian semimartingale model admits the presence of leverage effects, i.e. correlations between the no-arbitrage vector price process and the matrix-valued stochastic volatility process, but excludes discontinuous movements, or (co-)jumps, in either. Including the latter is an interesting direction for further research, see e.g. the analyses in Jacod & Todorov (2009) and Mancini & Gobbi (2012) in a noiseless setup with synchronous prices, but this is beyond the scope of this paper. The aim here is to estimate the quadratic covariation matrix of the d-valued no-arbitrage asset return vector,

$$[\boldsymbol{p}^*] = \int_0^1 \boldsymbol{\Sigma}_t dt \in \mathbb{R}^{d\times d}, \qquad \boldsymbol{\Sigma}_t = (\boldsymbol{\Sigma}_t^{1/2})(\boldsymbol{\Sigma}_t^{1/2})', \tag{2}$$

accounting for empirical regularities at tick-by-tick frequencies such as MMS noise and non-synchronous trading, formally introduced in the next subsections.

2.2 Synchronization Schemes

The individual observable prices for assets $q = 1,\ldots,d$ are recorded irregularly and non-synchronously at times $t_i^{(q)} \in [0,1]$, $i = 0,\ldots,N(1,q)$, where $N(t,q)$ is the counting process of asset q. They are comprised of a signal, the no-arbitrage price process, and an additive noise $\tilde{u}_{q,t_i^{(q)}}$, which summarizes an array of market imperfections and is formally defined later. Hence, the model reads

$$p_{q,t_i^{(q)}} = p^*_{q,t_i^{(q)}} + \tilde{u}_{q,t_i^{(q)}}, \qquad i = 0,\ldots,N(1,q), \quad q = 1,\ldots,d. \tag{3}$$

This signal-plus-noise model involves two challenges. First, non-synchronously observed prices create problems for high frequency estimators of the covariance of asset returns, since they induce stale prices and lead-lag relations among correlated assets, two recognized sources of the Epps effect. Second, the inclusion of MMS noise in the model leads to explosive behavior (will be formalized later) in the

estimated (co)variances when the sampling frequency increases unless properly accounted for, see e.g. Hansen & Lunde (2006), Voev & Lunde (2007), and Zhang (2011). There are two main approaches of dealing with non-synchronous observations. (1) Apply a synchronization scheme and estimate all elements of the covariance matrix jointly, e.g. Barndorff-Nielsen et al. (2011) and Ikeda (2011). (2) Align the observations by a pseudo-aggregation algorithm and estimate the covariance matrix element-wise, e.g. Hayashi & Yoshida (2005) and Voev & Lunde (2007). As such, both paradigms may be interpreted as pseudo-alignment of the observations and, as illustrated below, both may be analyzed within the same signal-plus-noise framework for the synchronized price (or equivalently, return) vector, provided that the MMS noise is specified appropriately. Following the aforementioned literature, the preferred synchronization scheme is *refresh time sampling*, since it includes the largest amount of data among all generalized sampling schemes, see Aït-Sahalia et al. (2010).

Definition 1 (Refresh Time Sampling)**.** *The refresh time common to all security* ***prices*** *is defined by: (1)* $t_0 = \max_{q=1,\dots,d} t_0^{(q)}$*, if* $i = 0$*. (2)* $t_i = \max_{q=1,\dots,d} t^{(q)}_{N(t_{i-1}^{(q)},q)+1}$*, if* $i \geq 1$*.*

Definition 2 (Hayashi-Yoshida Sampling)**.** *For two assets* $a, b \in 1, \dots d$*, denote the sets* $\Pi_a = \{t_i^{(a)} : i = 0, \dots, N(1,a)\}$ *and* $\Pi_b = \{t_i^{(b)} : i = 0, \dots, N(1,b)\}$*, respectively, and for* $N(1,a) \leq N(1,b)$ *denote* $\underline{t}_{i-1}^{(b)} = \max\{t \in \Pi_b : t \leq t_{i-1}^{(a)}\}$ *and* $\bar{t}_i^{(b)} = \min\{t \in \Pi_b : t \geq t_i^{(a)}\}$*. Define the set of overlapping* ***return*** *times as* $\Pi_{a,b} = \{t_i - t_{i-1} : \bar{t}_i^{(b)} - \underline{t}_{i-1}^{(b)} > 0\}$*. Now, for* $\Pi_{a,b} \neq \emptyset$*, the common sample of asset returns is defined:*

$$\Delta p_{a,t_i} = p_{a,t_i^{(a)}} - p_{a,t_{i-1}^{(a)}}, \qquad \Delta p_{b,t_i} = p_{b,\bar{t}_i^{(b)}} - p_{b,\underline{t}_{i-1}^{(b)}} \qquad \text{for} \quad \Delta t_i \in \Pi_{a,b}.$$

Let the synchronized sample size of returns be denoted by N regardless of sampling scheme. Refresh time sampling adjust to the trading frequency of the assets by sampling at times where all assets have been traded at least once. It facilitates direct estimation of the whole covariance matrix, but as the sampling scheme is universe-dependent, i.e. adapting to the underlying frequency of the slowest traded asset and depending on the degree of non-synchronous trading, the information loss is potentially great for heterogeneous assets. Contrary, the Hayashi-Yoshida sampling scheme only permits element-wise analysis of the covariance matrix, but as it operates with a 2-dimensional return vector, the pseudo-aggregation algorithm guarantees that the pairwise synchronized sample size is

$$N_{a,b} = i_N(a,b) - i_1(a,b) + 1 \qquad \text{for any} \quad a, b \in 1, \dots, d,$$

where $i_1(a,b)$ and $i_N(a,b)$ are two functions defined as

$$\begin{aligned} i_1(a,b) &= \mathbf{1}_{\{t_0^{(a)} \geq t_0^{(b)}\}} + N(t_0^{(b)}, a)\mathbf{1}_{\{t_0^{(a)} < t_0^{(b)}\}}, \\ i_N(a,b) &= N(1,a)\mathbf{1}_{\{t_{N(1,a)}^{(a)} \leq t_{N(1,b)}^{(b)}\}} + N(t_{N(1,b)}^{(b)}, a)\mathbf{1}_{\{t_{N(1,a)}^{(a)} > t_{N(1,b)}^{(b)}\}}. \end{aligned}$$

The functions $i_1(a,b)$ and $i_N(a,b)$ state that the (potential) loss of information occurs only when one asset, the fastest trading asset in this case, is not trading in the beginning and/or in the end of a

trading day. All relevant information about the covariance between the two assets is, however, kept, which may be interpreted as a maximum likelihood property, see Mykland (2010).

2.3 An Encompassing Additive Noise Model

The additional error induced by irregularly spaced and non-synchronous observations takes the following form for the refresh time sampling scheme:

$$\Delta p_{q,t_i} = \Delta p^*_{q,t_i} + \underbrace{\tilde{u}_{q,t_i^{(q)}} + p^*_{q,t_i^{(q)}} - p^*_{q,t_i}}_{u_{q,t_i}} - \underbrace{\tilde{u}_{q,t_{i-1}^{(q)}} + p^*_{q,t_{i-1}} - p^*_{q,t_{i-1}^{(q)}}}_{u_{q,t_{i-1}}} \tag{4}$$

where $t_i^{(q)} \leq t_i \ \forall q, i$, and a similar form for the Hayashi-Yoshida sampling scheme with u_{q,t_i} and $u_{q,t_{i-1}}$ defined slightly differently for the two assets a and b, respectively. As the synchronization error may be treated as an additional source of MMS noise, the vector noise component of the synchronized price vector must be specified to capture (cross) temporal dependence in the noise along with (cross) contemporaneous and serial correlation with $\boldsymbol{p}^*$ such that itself and additional microstructure features, e.g. bid-ask bounce effects (Roll 1984), asymmetric information (Glosten & Milgrom 1985), and lead-lag information transmission (Large 2007), are accounted for.

Assumption 3. *Let the synchronized and jittered d-dimensional return vector be defined by the signal-plus-noise model*

$$\Delta \boldsymbol{p}_{t_i} = \Delta \boldsymbol{p}^*_{t_i} + \boldsymbol{U}_{t_i} - \boldsymbol{U}_{t_{i-1}}, \qquad i = 1, \ldots, n$$

where $n, m \in \mathbb{N}^+$ with $n - 1 + 2m = N$ and $m \propto n^{\xi}$, $\xi \in (1/4, 1)$, the sample end-points are redefined as $\boldsymbol{p}_{t_0} = m^{-1} \sum_{i=1}^{m} \boldsymbol{p}_{t_{i-1}}$ and $\boldsymbol{p}_{t_n} = m^{-1} \sum_{i=1}^{m} \boldsymbol{p}_{t_{N-m+1}}$, and $\boldsymbol{p}_{t_i} = \boldsymbol{p}_{t_{m+i}}$ for $i \in [1, n-1]$. Further, let $\boldsymbol{U}_{t_i} \in \mathbb{R}^d$ be defined by the component model: $\boldsymbol{U}_{t_i} = \boldsymbol{e}_{t_i} + \boldsymbol{u}_{t_i}$.

Assumption 4. *$\exists r_e \in \mathbb{N}^+$ such that $\alpha_e = O(1)\mathbf{1}_{\{|g| \leq 1\}} + O(|g|^{-(1+r_e+\epsilon)})\mathbf{1}_{\{|g|>1\}} \in \mathbb{R}^+$ for some $\epsilon > 0$. Further, define the matrix of functions $\boldsymbol{\Theta}_t(g) : t \in [0,1] \to \mathbb{R}^{d \times k}$ and the vector of standard Brownian motions, $\boldsymbol{B}_t \in \mathbb{R}^k$. Then, $\boldsymbol{e}_{t_i}$, $i = 1, \ldots, N$, has representation*

$$\boldsymbol{e}_{t_i} = \sum_{g=-\infty}^{\infty} \boldsymbol{\Theta}(t_i, g) \left(\Delta t_{i-g}\right)^{-1/2} \Delta \boldsymbol{B}_{t_{i-g}},$$

which satisfies the following conditions: (1) $d[\boldsymbol{W}, \boldsymbol{B}]_t = \boldsymbol{\Upsilon}_t dt$ where $\boldsymbol{\Upsilon}_t \in \mathbb{R}^{k \times k}$ is continuous and $\forall (t, w) \in [0,1] \times \mathcal{O}\ \exists \Lambda_2 > 0 : \|\boldsymbol{\Upsilon}_t(w)\| \leq \Lambda_2$, (2) $\sup_{t_i \in [0,1]} \|\boldsymbol{\Theta}(t_i, g)\| \leq \alpha_e(g)$, (3) $\sup_{t \in [0,1]} \|\boldsymbol{\Theta}_t(g)\| \leq \alpha_e(g)$, (4) for some $\Lambda_3 \in (0, \infty)$, $\sup_g \sum_{i=1}^{N} \|\boldsymbol{\Theta}(t_i, g) - \boldsymbol{\Theta}_{t_i}(g)\| \leq \Lambda_3$, (5) $\sum_{i=1}^{N} \left\|\boldsymbol{\Theta}_{t_i}(g) - \boldsymbol{\Theta}_{t_{i-1}}(g)\right\| \leq \alpha_e(g)$, and (6) $\boldsymbol{\Upsilon}_t$ is $\mathcal{H}_t$-adapted, and $\boldsymbol{\Theta}(t, g)$, and $\boldsymbol{\Theta}_t(g)$ are both $\mathcal{H}_1$-measurable for all g.

Assumption 5. *$\boldsymbol{u}_{t_i}$ is a strictly stationary, $(\mathcal{G}_{t_i})$-measurable α-mixing sequence of random vectors with mixing coefficient $\alpha_u(h)$ such that $\exists r_u \in \mathbb{N}^+ : \sum_{j=1}^{\infty} j^{r_u} \alpha(j) < \infty$. Further, $\forall i = 1, \ldots, N$: $\mathbb{E}[\boldsymbol{u}_i] = \mathbf{0}$ and $\exists v > 4 : \sup_i \mathbb{E}[\|\boldsymbol{u}_i\|^v] < \infty$. Last, denote the h-th autocovariance $\boldsymbol{\Omega}^{(uu)}(h)$, the long run variance*

$\boldsymbol{\Omega}^{(uu)} = \sum_{h\in\mathbb{Z}} \boldsymbol{\Omega}^{(uu)}(h)$ *and for* $j,k,l \in \mathbb{Z}$*, let the third and fourth order cumulants,* $\kappa_3(0,j,k)$ *and* $\kappa_4(0,j,k,l)$*, respectively, satisfy* $\sum_{j,k\in\mathbb{Z}} |\kappa_3(0,j,k)| < \infty$ *and* $\sum_{j,k,l\in\mathbb{Z}} |\kappa_4(0,j,k,l)| < \infty$*.*

Assumption 6. *The durations between observation times are defined as* $t_i - t_{i-1} = D_{n,i}/n \ \forall i = 1,\dots,N$*, where* $D_{n,i} \in \mathbb{R}_+$ *satisfies: (1)* $D_{n,i} \perp\!\!\!\perp (\boldsymbol{e}_{t_j}, \boldsymbol{p}^*_{t_k})' \ \forall j,i,k = 1,\dots,N$*. (2)* $\mathbb{E}[D^s_{n,i}|\mathcal{H}_1] \xrightarrow{\mathbb{P}} \chi_s(t_i)$*,* $0 < s \leq 2$ *as* $n \to \infty$*, where* $\chi_s(t) \in \mathbb{R}_+$ *is càdlàg,* $(\mathcal{H}_t)$*-adapted and bounded. (3)* $\mathbb{V}[D_{n,i}|\mathcal{H}_1] \xrightarrow{\mathbb{P}} \varrho(t_i) \in \mathbb{R}_+$*, which is càdlàg,* $(\mathcal{H}_t)$*-adapted and bounded. (4)* $\max_{i=1,2,\dots,n} D_{n,i} = o_p(n^\delta)$ *for* $\delta \in (0,1)$*.*

Assumption 3 formalizes the synchronized return vector, the additive structure of the MMS noise model, and defines the (for kernel-based estimators) well-known jittering of the end-points to avoid end-effects influencing the asymptotic distribution. Assumption 4-5 is related to their counterparts in Barndorff-Nielsen et al. (2011) and Ikeda (2011), and they are multivariate generalizations of the framework in Varneskov (2013). The MMS noise is comprised of an exogenous component, $\boldsymbol{u}_{t_i}$, which captures α-mixing dependence in the noise, and an endogenous component, $\boldsymbol{e}_{t_i}$, that allows for a two-sided, non-degenerate endogenous correlation structure to capture, among others, asymmetric lead-lag relations, de Jong & Nijman (1997), Voev & Lunde (2007) and Griffin & Oomen (2011), and the correlations caused by sampling errors in (4) and/or its Hayashi-Yoshida sampling equivalent. Even in the case where $\boldsymbol{W}$ and $\boldsymbol{B}$ are uncorrelated, Assumption 4 allows the noise to exhibit diurnal heteroskedasticity, e.g. Kalnina & Linton (2008). Clearly, (4) and Assumption 4 illustrate that the treatment of non-synchronicity related challenges is embedded in the problem of accounting for MMS noise in a synchronized sample. Relative to the work of Barndorff-Nielsen et al. (2011) and Ikeda (2011), the polynomial mixing rate is milder than the exponential mixing rate in Ikeda (2011), and the endogenous noise model, which is inspired by the work of Dahlhaus & Polonik (2009) and Dahlhaus (2009) on spectral analysis of locally stationary processes, generalizes the endogenous noise models of the former two by allowing for increased persistence and greater flexibility in the noise generating process. Assumption 6 is inspired by Phillips & Yu (2008). Instead of making explicit statements about the arrival times of individual observations, it imposes a mild structure on the durations between synchronized observations. The durations are allowed to be random, but exogenous, and they nest exogenous Poisson arrivals as the latter implies $\max_{i=1,2,\dots,n} D_{n,i} = O_p(\ln(n))$. Note that the exogeneity assumption is not strict as any endogenous effects are picked up by Assumption 4.

Remark 1. *Let* $\boldsymbol{\Omega}^{(ee)}_t(h) = \sum_{j=-\infty}^{\infty} \boldsymbol{\Theta}_t(j+h)\boldsymbol{\Theta}_t(j)$ *and* $\boldsymbol{\Omega}^{(ee)}_t = \sum_{h\in\mathbb{Z}} \boldsymbol{\Omega}^{(ee)}_t(h)$ *denote the local* h*-th autocovariance and long run covariance of* $\boldsymbol{e}_t$*. Similarly, let* $\boldsymbol{\Omega}^{(ep)}_t(h) = \boldsymbol{\Theta}_t(h)\boldsymbol{\Upsilon}_t(\boldsymbol{\Sigma}^{1/2}_t)'$ *and* $\boldsymbol{\Omega}^{(ep)}_t = \sum_{h\in\mathbb{Z}} \boldsymbol{\Omega}^{(ep)}_t(h)$ *be the corresponding covariance quantities for* $\boldsymbol{e}_t$ *and* $\Delta\boldsymbol{p}^*_t$*. Assumptions 1 and 4, then, imply* $\sup_{t\in[0,1]} \|\boldsymbol{\Omega}^{(ee)}_t(h)\| \leq O(\alpha_e(h))$ *and* $\sup_{t\in[0,1]} \|\boldsymbol{\Omega}^{(ep)}_t(h)\| \leq O(\alpha_e(h))$*, thereby ensuring* $\sum_{h\in\mathbb{Z}} |h|^{r_e} \|\boldsymbol{\Omega}^{(ee)}_t(h)\| < \infty$ *and* $\sum_{h\in\mathbb{Z}} |h|^{r_e} \|\boldsymbol{\Omega}^{(ep)}_t(h)\| < \infty$ *such that the asymptotic mean and variance of the estimators may be established. Note that none of these quantities are progressively measurable with respect to* $\mathcal{H}_t$*, but they are, however, measurable with respect to* $\mathcal{H}_1$*.*

Remark 2. *Preliminary bounds have been imposed on the jittering rate,* $\xi \in (1/4, 1)$*, and the random duration between observations,* $\delta \in (0,1)$*. Both will be strengthened to derive the asymptotic results.*

Remark 3. *From Definitions 1-2, it is apparent that refresh time samples and Hayashi-Yoshida samples are constructed similarly. The latter differs from the former, however, by replacing previous-tick interpolation with next-tick interpolation at the end-points of observation increments. This impacts the persistency of the MMS noise for the leading asset in Definition 2 as* $\Delta p_{b,t_{i+1}}$ *and* $\Delta p_{b,t_i}$ *will be positively correlated by construction. Depending on the strength of this correlation, the difference in sample scheme may have detrimental finite sample effects on realized kernel estimators.*

3 Flat-Top Realized Kernel Estimation

The generic structure of the proposed estimators follows the realized kernel framework,

$$RK(\boldsymbol{p}) = \Gamma_0(\boldsymbol{p}) + \sum_{h=1}^{n-1} k\left(\frac{h}{H}\right)\{\Gamma_h(\boldsymbol{p}) + \Gamma_{-h}(\boldsymbol{p})\} \tag{5}$$

where $k(\cdot)$ is a weighting function, H is the bandwidth, and $\Gamma_h(\cdot)$ is the realized covariation, which for any two equally dimensioned vector processes $\boldsymbol{X}$ and $\boldsymbol{Z}$ is defined as

$$\Gamma_h(\boldsymbol{X}, \boldsymbol{Z}) = \sum_{i=|h|+1}^{n} \Delta \boldsymbol{X}_{t_i} \Delta \boldsymbol{Z}'_{t_{i-h}} \qquad \text{for} \quad h \geq 0, \tag{6}$$

$\Gamma_h(\boldsymbol{X}, \boldsymbol{Z}) = \Gamma_{-h}(\boldsymbol{X}, \boldsymbol{Z})'$ for $h < 0$, and $\Gamma_h(\boldsymbol{X}, \boldsymbol{X}) = \Gamma_h(\boldsymbol{X})$. The realized covariance estimator is defined within the realized autocovariation structure as $RC = \Gamma_0(\boldsymbol{p})$. In a continuous Brownian semimartingale model, $RC \xrightarrow{\mathbb{P}} \int_0^1 \boldsymbol{\Sigma}_t dt$ as $\sup_i \Delta t_i \to 0$ for $n \to \infty$, and its asymptotic distribution theory is studied in Barndorff-Nielsen & Shephard (2004). When MMS noise is included in the model,

$$RC = \Gamma_0(\boldsymbol{p}^*) + \Gamma_0(\boldsymbol{U}) + \Gamma_0(\boldsymbol{p}^*, \boldsymbol{U}) + \Gamma_0(\boldsymbol{U}, \boldsymbol{p}^*) = O_p(1) + O_p(n) + O_p(n^{1/2}),$$

using the Cauchy-Schwartz inequality. Hence, as the sampling frequency increases, the estimated (co)variances explode. To circumvent this problem, the realized kernel estimators use appropriate weighting of higher-order realized autocovariations $\Gamma_h(\boldsymbol{p})$ $h \neq 0$ to reduce the MMS noise-induced bias and variance, including the implicit effects of non-synchronicity, sufficiently to achieve consistent estimators. The asymptotic and finite sample properties of the estimators in (5), however, are greatly dependent on the choice of $k(\cdot)$. Consider the following two classes of kernel functions:

Definition 3. *$\mathcal{K}$ is a set of functions k: $\mathbb{R} \to [-1, 1]$. Define $k^{(j)}(x) \equiv \partial^j k(x)/\partial x^j$, $k_{\tilde{a}}^{(2)} = \lim_{x \to 0} |x|^{-\tilde{a}}(k^{(2)}(0) - k^{(2)}(x)) < \infty$, $\exists \tilde{a} \geq 1$, $\tilde{q} = \max_{\tilde{a} \in \mathbb{N}^+}\{\tilde{a} \geq 1 : k_{\tilde{a}}^{(2)} \in (-\infty, 0)\}$, and let $k(x)$ satisfy the following five regularity conditions: (1) $k(x)$ is twice continuously differentiable, $k^{(2)}(x)$ is differentiable at all but a finite number of points. (2) $k(x) = k(-x)$. (3) $k(0) = 1$, $k^{(1)}(0) = 0$, $k^{(2)}(0) < 0$. (4) $k^{(jj)} \equiv \int_0^\infty [k^{(j)}(x)]^2 dx < \infty$ for $j = 0, 1, 2$, and for $j = 3$ almost everywhere. (5) $\int_{-\infty}^{\infty} k(x) e^{-ix\lambda} \geq 0$, $\forall \lambda \in \mathbb{R}$.*

Definition 4. *Let $c = H^{-\gamma}$ for some $\gamma \in [0,1]$ and $\lambda(x) \in \mathcal{K}$ and define $\mathcal{K}^*$ as the set of functions k: $\mathbb{R} \to [-1,1]$ characterized by $k(x) = \mathbf{1}_{\{|x| \leq c\}} + \lambda(|x| - c)\mathbf{1}_{\{|x| > c\}}$.*

Barndorff-Nielsen et al. (2011) analyze realized kernels using the $\mathcal{K}$ class of kernel functions. These are characterized by being second-order smooth, condition (1), where $\tilde{q}$ measures the smoothness around the origin of $k^{(2)}(x)$. This excludes the Bartlett kernel, which, as Barndorff-Nielsen, Hansen, Lunde & Shephard (2008) show, cannot achieve consistency faster than the suboptimal rate $n^{1/6}$. Conditions (3) and (5) guarantee the class of estimators to be positive semi-definite. The class of flat-top realized kernels, $\mathcal{K}^*$, differs from $\mathcal{K}$ by having a shrinking flat-top support $[-c, c]$ in the neighborhood of the origin, which, as will be laid out in the next subsections, is crucial for obtaining rate-optimal estimators.

3.1 Motivation for Flat-Top Kernels

The choice of kernel function impacts the asymptotic distribution of the realized kernels. To see this, define $\alpha(h) = \max(\alpha_e(h), \alpha_u(h))$ and $r = \min(r_e, r_u) \in \mathbb{N}^+$, and consider, initially, the first two $\mathcal{H}_1$-conditional moments for $k(x) \in \mathcal{K}$.

Lemma 1. *Let Assumptions 1-6 be satisfied and let $H \propto n^{\nu}$, $\nu \in (1/3, 1)$, $\delta \in (0, 1-\nu)$, $\xi \in (1/4, 1/(2+\delta))$, and $k(x) \in \mathcal{K}$ with $\tilde{q} \leq r$, then the first two $\mathcal{H}_1$-conditional moments are:*

$$\mathbb{E}[RK(\boldsymbol{p})|\mathcal{H}_1] = \int_0^1 \boldsymbol{\Sigma}_t dt + nH^{-2}|k^{(2)}(0)|\boldsymbol{\Omega} + nH^{-(2+\tilde{q})}k_{\tilde{q}}^{(2)} \sum_{h \in \mathbb{Z}} |h|^{\tilde{q}} \boldsymbol{\Omega}(h) + O_p\left(n^{1/2}H^{-2}\right) + o_p(1),$$

$$\mathbb{V}[RK(\boldsymbol{p})|\mathcal{H}_1] = 4Hn^{-1}k^{(00)}\boldsymbol{\mathcal{Q}} + 4nH^{-3}k^{(22)}\boldsymbol{\mathcal{N}} + 8H^{-1}k^{(11)}\boldsymbol{\mathcal{C}} + o_p(1),$$

where, if $\boldsymbol{N}_d$ is the symmetrizer matrix (see Appendix A.1), $\boldsymbol{\Omega}(h) = \boldsymbol{\Omega}^{(uu)}(h) + \int_0^1 \boldsymbol{\Omega}_t^{(ee)}(h)\chi_1^{-1}(t)dt$ and $\boldsymbol{\Omega} = \sum_{h \in \mathbb{Z}} \boldsymbol{\Omega}(h)$ is the h-th average autocovariance and long run variance of $\boldsymbol{U}$, and

$$\boldsymbol{\mathcal{Q}} = \int_0^1 (\boldsymbol{\Sigma}_t \otimes \boldsymbol{\Sigma}_t)\, \chi_2(t)\chi_1^{-1}(t)dt,$$

$$\boldsymbol{\mathcal{N}} = \boldsymbol{\Omega}^{(uu)} \otimes \boldsymbol{\Omega}^{(uu)} + \int_0^1 \left(\boldsymbol{\Omega}_t^{(ee)} \otimes \boldsymbol{\Omega}_t^{(ee)}\right) \chi_1^{-1}(t)dt + 2\boldsymbol{N}_d \left(\boldsymbol{\Omega}^{(uu)} \otimes \int_0^1 \boldsymbol{\Omega}_t^{(ee)}\chi_1^{-1}(t)dt\right),$$

$$\boldsymbol{\mathcal{C}} = \boldsymbol{N}_d \left(\int_0^1 \left(\boldsymbol{\Sigma}_t \otimes \left(\boldsymbol{\Omega}^{(uu)} + \boldsymbol{\Omega}_t^{(ee)}\right)\chi_1(t) + 2\boldsymbol{\Omega}_t^{(ep)} \otimes \boldsymbol{\Omega}_t^{(ep)}\right)\chi_1^{-1}(t)dt\right).$$

From Lemma 1, it follows that the MMS noise biases the realized kernel estimator of Barndorff-Nielsen et al. (2011) such that its bandwidth cannot be set $H \propto n^{1/2}$, which otherwise balances the contributions from discretization ($\boldsymbol{\mathcal{Q}}$), MMS noise ($\boldsymbol{\mathcal{N}}$), and cross-products ($\boldsymbol{\mathcal{C}}$) to the asymptotic covariance matrix, thus preventing the estimator from achieving the optimal rate of convergence, $n^{1/4}$. Instead, the bandwidth must be over-smoothened ($\nu > 1/2$) to eliminate the dominant bias component, which, as a consequence, leads to the bias-variance balancing choice $H \propto n^{3/5}$ and a suboptimal rate

of convergence, $n^{1/5}$. Non-synchronicity related problems such as lead-lag relations, sampling errors, etc. impact both the bias and variance of the realized kernels through $\boldsymbol{\Omega}_t^{(ee)}$, $\boldsymbol{\Omega}_t^{(ep)}$, and $\chi_s(t)$. Thus, an estimator which corrects the leading (and smaller order) bias will simultaneously account for both MMS noise and non-synchronicity. To motivate the flat-top correction, rewrite the contribution of $\boldsymbol{U}$ on the asymptotic distribution as

$$\begin{aligned}\sum_{i=1}^{n}\Delta\boldsymbol{U}_{t_i}\Delta\boldsymbol{U}_{t_i}'+\sum_{h=1}^{n-1}k\left(\frac{h}{H}\right)\sum_{i=h+1}^{n}\left(\Delta\boldsymbol{U}_{t_i}\Delta\boldsymbol{U}_{t_{i-h}}'+\Delta\boldsymbol{U}_{t_{i-h}}\Delta\boldsymbol{U}_{t_i}'\right)\\ \approx\frac{n}{H^2}a(0)\sum_{i=1}^{n}\boldsymbol{U}_{t_i}\boldsymbol{U}_{t_i}'+\frac{n}{H^2}\sum_{h=1}^{n-1}a\left(\frac{h}{H}\right)\frac{1}{n}\sum_{i=h+1}^{n}\left(\boldsymbol{U}_{t_i}\boldsymbol{U}_{t_{i-h}}'+\boldsymbol{U}_{t_{i-h}}\boldsymbol{U}_{t_i}'\right)\end{aligned}\tag{7}$$

where $a(h/H)$ is the finite sample analog of $-k^{(2)}(h/H)$ and the approximation error is due to end-effects of order $O_p(m^{-1})$. Clearly, (7) shows that the problem of estimating quadratic covariation resembles that of spectral analysis (or HAC estimation), and, more importantly, that extending the flat-top region of the kernel function by $c=H^{-\gamma}$ exactly eliminates the bias-contribution from the first $H^{1-\gamma}$ autocovariances, $\boldsymbol{\Omega}(h)$, whose implications are formalized in the following lemma:

Lemma 2. *Let Assumptions 1-6 be satisfied and let $H\propto n^{\nu}$, $\nu\in(1/3,1)$, $\delta\in(0,1-\nu)$, $\xi\in(1/4,1/(2+\delta))$, and $k(x)\in\mathcal{K}^*$, then the first two $\mathcal{H}_1$-conditional moments are:*

$$\begin{aligned}\mathbb{E}[RK(\boldsymbol{p})|\mathcal{H}_1]&=\int_0^1\boldsymbol{\Sigma}_t dt+O_p\left(\alpha\left(H^{1-\gamma}\right)nH^{-2}\right)+O_p\left(\alpha_e\left(H^{1-\gamma}\right)n^{1/2}H^{-1}\right),\\ \mathbb{V}[RK(\boldsymbol{p})|\mathcal{H}_1]&=4Hn^{-1}\left(\lambda^{(00)}+H^{-\gamma}\right)\boldsymbol{\mathcal{Q}}+4nH^{-3}\lambda^{(22)}\boldsymbol{\mathcal{N}}+8H^{-1}\lambda^{(11)}\boldsymbol{\mathcal{C}}+o_p(1).\end{aligned}$$

Lemma 2 shows that if $\gamma\in(0,1)$, the flat-top realized kernels may have $H\propto n^{1/2}$ and still eliminate both the dominant and smaller order bias components asymptotically with no implications for the asymptotic variance, thus enabling consistency at the optimal rate, $n^{1/4}$. The finite sample bias and variance, however, depends on the choice of γ and, r, the smoothness of the MMS noise.

Remark 4. *The bounds on jittering $\xi\in(1/4,1/(2+\delta))$ and the random duration $\delta\in(0,1-\nu)$ sharpen similar bounds in Barndorff-Nielsen et al. (2011) and Ikeda (2011), who do not treat the end-averaged no-arbitrage returns, $\Delta\boldsymbol{p}_{t_1}^*$ and $\Delta\boldsymbol{p}_{t_n}^*$ as triangular arrays, and they prevent end-effects from impacting the asymptotic distribution. While jittering and end-effects are important for the theoretical analysis, (Barndorff-Nielsen et al. 2011, Section 6.4) dismiss their practical relevance.*

Remark 5. *The flat-top kernel functions in $\mathcal{K}^*$ are subtly different from the flat-top kernels analyzed by Politis (2011) in the context of spectral analysis, who fixes $c\in(0,1]$. As seen from Lemma 2, fixing c leads to a strictly larger asymptotic variance of $RK(\boldsymbol{p})$. Furthermore, Barndorff-Nielsen et al. (2008) consider kernel functions from $\mathcal{K}^*$ with $\gamma=1$ under the assumption that the MMS noise is i.i.d. and exogenous. However, as seen from Lemma 2, these are inconsistent in the present setting unless $\nu>1/2$, similar to kernel functions from $\mathcal{K}$.*

3.2 Central Limit Theory

The asymptotic elimination of the dominant bias in Lemma 1 has great implications for the statistical properties of the flat-top realized kernels, which are summarized in the following theorem, where, to avoid confusion from this point on, $RK^*(\boldsymbol{p})$ is used to denote (flat-top) realized kernels with $k(x) \in \mathcal{K}^*$.

Theorem 1. *Let Assumptions 1-6 be satisfied and let $H = \kappa n^{1/2}$ where $\kappa > 0$, $\delta \in (0, 1/2)$, $\xi \in (1/4, (3/8)/(1+\delta/2))$, and $\gamma \in (0, (1/2+r)/(1+r))$, and define $\boldsymbol{\mathcal{B}}(\lambda, \kappa) = \lim_{n\to\infty} n^{1/2}\mathbb{V}[RK^*(\boldsymbol{p})|\mathcal{H}_1]$, then*

$$n^{1/4}\left(RK^*(\boldsymbol{p}) - \int_0^1 \boldsymbol{\Sigma}_t dt\right) \overset{d_s(\mathcal{H}_1)}{\to} MN\left(\mathbf{0}, \boldsymbol{\mathcal{B}}(\lambda, \kappa)\right).$$

Here, $\overset{d_s(\mathcal{H}_1)}{\to}$ stands for $\mathcal{H}_1$-stable convergence and "MN" abbreviates a mixed Gaussian distribution, see Appendix A.1 for a definition and, e.g., (Barndorff-Nielsen et al. 2008, Appendix A) and Podolskij & Vetter (2010) for details. Theorem 1, and similarly for Lemmas 1-2, generalizes the univariate result in (Varneskov 2013, Theorem 1) to a multivariate setting, which allows for non-synchronous trading, random durations between observations and asymmetric lead-lag dependencies that impact the asymptotic variance of the estimators through $\boldsymbol{\mathcal{Q}}$, $\boldsymbol{\mathcal{N}}$, and $\boldsymbol{\mathcal{C}}$. It shows that by imposing suitable conditions on the flat-top shrinkage, γ, the flat-top realized kernels are consistent, asymptotically unbiased and mixed Gaussian at the optimal rate of convergence, $n^{1/4}$. The lower bound, $\gamma > 0$, prevents the asymptotic variance from being inflated (see Lemma 2), while the upper bound, $\gamma < (1/2+r)/(1+r)$, which follows from scaling the remaining bias of order $\alpha\left(H^{(1-\gamma)}\right) = O\left(H^{-(1+r+\epsilon)(1-\gamma)}\right)$ by $n^{1/4}$, guarantees the flat-top realized kernels to be asymptotically unbiased and consistent at the optimal rate. It is exactly the slower flat-top shrinkage relative to the realized kernels in Barndorff-Nielsen et al. (2008, 2011), which provides the stronger theoretical result, $n^{1/4}$ vs. $n^{1/5}$-consistency.

The decomposition of the asymptotic variance, $\boldsymbol{\mathcal{B}}(\lambda, \kappa)$, is similar to the decomposition in (Barndorff-Nielsen et al. 2008, Theorem 4), who consider the special case with $d = 1$, and i.i.d. and exogenous MMS noise. This implies that bandwidth selection (see Section 5.2 for details) and efficiency analysis of the flat-top realized kernels mimics its counterparts in (Barndorff-Nielsen et al. 2008, Sections 4.3-4.5), see also (Ikeda 2013, Section 3.3) and (Varneskov 2013, Section 3.2). Most importantly, $\boldsymbol{\mathcal{B}}(\lambda, \kappa)$ illustrates that the intrinsic efficiency of $\lambda(x)$ controls the asymptotic efficiency of the flat-top realized kernels, and for the parametric version of the univariate problem ($d = 1$, constant volatility and i.i.d. MMS noise), setting $\lambda(x) = (1+x)e^{-x}$ in conjunction with an optimally selected bandwidth allows the flat-top realized kernels to achieve the Cramér-Rao efficiency bound.

Remark 6. *Ikeda (2011) provides an inference strategy for the two-scale realized kernel (discussed in Section 5.5) based on the subsampling scheme of Kalnina (2011), which is also applicable for flat-top realized kernels: (1) Divide the n synchronized observations into $L_n = \lfloor n/M_n \rfloor$ subsamples of successive observations with size M_n, where the l-th subsample is given by observations $i = (l-1)M_n, \cdots, lM_n$. (2) Let lJ_n denote a smaller, centered subsample of lM_n, and define the flat-top realized kernel for subsample l of size k_n as $RK^*(\boldsymbol{p}, l, k_n)$ for $k = (J, M)$ where the bandwidth is selected as $H = \kappa k_n^{1/2}$.*

(3) Let $\Delta_{l,k} = n^{-1}\sum_{i=1}^{k_n} D_{n,(l-1)k_n+i}$ *be the time span covered by a given subsample and define the* $d^2 \times d^2$ *flat-top asymptotic variance (FTAV) estimator as*

$$FTAV(\boldsymbol{p}) = J_n^{1/2}\sum_{l=1}^{L_n}\{\boldsymbol{V}_{l,n}\otimes\boldsymbol{V}_{l,n}\}\,\Delta_{l,M}, \qquad \boldsymbol{V}_{l,n} = \Delta_{l,J}^{-1}RK^*(\boldsymbol{p},l,J_n) - \Delta_{l,M}^{-1}RK^*(\boldsymbol{p},l,M_n).$$

If the conditions of Theorem 1 are satisfied, where the jittering $m_k \propto k_n^{\xi}$ *occurs in all subsamples,* $M_n \to \infty$*,* $J_n \to \infty$*,* $J_n/M_n = o(1)$*,* $(J_nM_n)/n = o(1)$*, then* $FTAV(\boldsymbol{p}) \xrightarrow{\mathbb{P}} \boldsymbol{\mathcal{B}}(\lambda,\kappa)$ *for some* $\kappa > 0$ *follows from (Ikeda 2011, Proposition 1), providing a feasible central limit theory.*

3.3 Element-wise Estimation of the Covariance Matrix

The theoretical results presented so far are valid for any finite d in a synchronized sample. However, the Hayashi-Yoshida sampling scheme is no longer valid when $d > 2$. Additionally, the loss of information from refresh time sampling may potentially be great when d is large. Hence, an alternative strategy is to estimate the $d(d+1)/2$ unique elements of the covariance matrix separately. Let $n_{a,b}$, $H_{a,b}$ and $RK^*_{a,b}(\boldsymbol{p})$ denote the (jittered) synchronized sample size, bandwidth and flat-top realized kernel, respectively, for any pair $a, b \in 1,\dots d$, and define the element-wise flat-top realized kernel estimator as

$$ERK^*(\boldsymbol{p}) = \begin{pmatrix} RK^*_{1,1}(\boldsymbol{p}) & RK^*_{1,2}(\boldsymbol{p}) & \dots & RK^*_{1,d}(\boldsymbol{p}) \\ RK^*_{1,2}(\boldsymbol{p}) & RK^*_{2,2}(\boldsymbol{p}) & \dots & RK^*_{2,d}(\boldsymbol{p}) \\ \vdots & \vdots & \ddots & \vdots \\ RK^*_{1,d}(\boldsymbol{p}) & RK^*_{2,d}(\boldsymbol{p}) & \dots & RK^*_{d,d}(\boldsymbol{p}) \end{pmatrix}. \tag{8}$$

The element-wise flat-top realized kernels, using pair-wise refresh time sampling, is similar to the composite realized kernels of Lunde, Shephard & Sheppard (2011). They differ, however, by not splitting covariance estimation into separate estimation of volatilities and correlations and by using flat-top realized kernels, $k(x) \in \mathcal{K}^*$, instead of realized kernels, $k(x) \in \mathcal{K}$.

Corollary 1. *Let the conditions of Theorem 1 hold,* $n/n_{a,b} \to k_{a,b}^2 \in (0,1]$*,* $H_{a,b} = \kappa_{a,b}n_{a,b}^{1/2}$*, and* $\boldsymbol{\mathcal{B}}_{a,b}(\lambda,\kappa_{a,b}) \in \mathbb{R}_+$ *be the asymptotic variance, defined via* $\boldsymbol{\mathcal{B}}(\lambda,\kappa)$*, whose exact form is provided in Appendix A.2, then the* (a,b)*-th element of* $ERK^*(\boldsymbol{p})$ *has the following marginal distribution*

$$n^{1/4}\left(RK^*_{a,b}(\boldsymbol{p}) - \int_0^1 \boldsymbol{\Sigma}_t^{a,b}dt\right) \xrightarrow{d_s(\mathcal{H}_1)} MN\left(0, k_{a,b}\boldsymbol{\mathcal{B}}_{a,b}(\lambda,\kappa_{a,b})\right).$$

Proof. Follows directly from Theorem 1. □

Corollary 1 shows that all estimated elements have optimal asymptotic properties (asymptotically unbiased, consistent at the optimal rate, and reaches the Cramér-Rao efficiency bound for the parametric problem) as a consequence of using $k(x) \in \mathcal{K}^*$, which is not the case for the composite realized kernels, whose elements have asymptotic properties characterized by Lemma 1. Furthermore, the

element-wise flat-top realized kernel estimator may potentially have two sources of finite sample efficiency gains relative to the flat-top realized kernels. The first is through element-wise tailoring of the bandwidth (hence through $\kappa_{a,b}$, see Section 5.2 for details). The second is from the additional information maintained by doing pair-wise synchronization instead of global synchronization (hence through $k_{a,b}$). In fact, the diagonal elements are estimated using all available observations.

4 Finite Sample Adjustments and Applications

Despite their attractive asymptotic properties, neither the flat-top realized kernels nor the element-wise version are guaranteed to produce positive semi-definite estimates of quadratic covariation, which, in its strict form, is important for many non-linear transformations, see e.g. Section 4.2. As this is a recurring problem among rate-optimal estimators, cf. the bias-corrected pre-averaging estimator in Christensen et al. (2010) and the two-scale realized kernel in Ikeda (2011), Section 4.1 provides a simple correction to ensure positive definiteness of such estimators.

4.1 A Positive Definite Projection

The positive definite projection is based a unitary decomposition $RK^*(\boldsymbol{p}) = \boldsymbol{M}'\boldsymbol{K}\boldsymbol{M}$ where $\boldsymbol{M}$ is a matrix of orthonormal eigenvectors and $\boldsymbol{K} = \text{diag}(k_1, \ldots, k_d)$ is a diagonal matrix of eigenvalues. Let $\hat{\boldsymbol{K}} = \text{diag}(\hat{k}_1, \ldots, \hat{k}_d)$ where $\hat{k}_q = \max(k_q, \epsilon_n)$, $q = 1, \ldots, d$, for some $\epsilon_n = o(n^{-1/4}) \in \mathbb{R}_+$ be a diagonal matrix of truncated eigenvalues, and use this to define a positive definite flat-top realized kernel estimator as $RK^\epsilon(\boldsymbol{p}) = \boldsymbol{M}'\hat{\boldsymbol{K}}\boldsymbol{M}$. The optimal asymptotic properties of $RK^*(\boldsymbol{p})$ comes with the sacrifice of the positive semi-definiteness, and while incidents of negative definite estimates may be rare due to the fast rate of convergence, $n^{1/4}$, the use of $\hat{\boldsymbol{K}}$ provides an easy fix of this event.

Theorem 2. *Let the conditions of Theorem 1 hold and assume $\int_0^1 \boldsymbol{\Sigma}_t dt$ is positive definite, then*

$$RK^\epsilon(\boldsymbol{p}) = RK^*(\boldsymbol{p}) + o_p(n^{-1/4}).$$

Theorem 2 differs notably from (Politis 2011, Corollary 4.1), who studies a similar eigenvalue truncation for spectral estimates and shows that while the truncation does not alter the rate of consistency, it may change the asymptotic distribution, i.e. his result replaces $o_p(n^{-1/4})$ with $O_p(n^{-1/4})$. The stronger result in Theorem 2 is attributed to the flat-top realized kernels being asymptotically unbiased, which is not the case in Politis (2011). In fact, Theorem 2 may be generalized:

Theorem 3. *Under the conditions of Theorem 2, let $\boldsymbol{V} \in \mathbb{R}^{d\times d}$ be an estimator satisfying $n^{1/4}(\boldsymbol{V} - \int_0^1 \boldsymbol{\Sigma}_t dt) \overset{d_s(\mathcal{H}_1)}{\to} MN(\boldsymbol{0}, \boldsymbol{\mathcal{B}}(\boldsymbol{V}))$ where $\boldsymbol{\mathcal{B}}(\boldsymbol{V}) \in \mathbb{R}_+$ is $\mathcal{H}_1$-measurable and bounded, and let $\boldsymbol{V}^\epsilon$ differ from $\boldsymbol{V}$ only by having its eigenvalues truncated by $\epsilon_n = o(n^{-1/4}) \in \mathbb{R}_+$, then $\boldsymbol{V}^\epsilon = \boldsymbol{V} + o_p(n^{-1/4})$.*

Proof. Same as for Theorem 2. □

Theorems 2-3 show that asymptotically unbiased and rate-optimal estimators of quadratic covariation may also enjoy positive definiteness via a simple asymptotically negligible correction of the eigenvalues. Hence, the conjecture of (Ikeda 2011, p. 15) that (Politis 2011, Corollary 4.1) may be applied directly in the context of quadratic covariation estimation overstates the impact of eigenvalue truncation. While the selection of $\epsilon_n = o(n^{-1/4}) \in \mathbb{R}_+$ is asymptotically irrelevant, it may have finite sample implications not only for the estimates of quadratic covariation, but also for non-linear transformations thereof, and its impact is, thus, analyzed in the simulation study.

4.2 Non-Linear Transformations

Two applications in financial economics that depend on transformations of quadratic covariation estimates are the realized correlation and realized regression coefficients, respectively,

$$\rho_{ab} = \left(\int_0^1 \boldsymbol{\Sigma}_t^{a,a} dt \int_0^1 \boldsymbol{\Sigma}_t^{b,b} dt \right)^{-1/2} \int_0^1 \boldsymbol{\Sigma}_t^{a,b} dt, \qquad \beta_{ab} = \left(\int_0^1 \boldsymbol{\Sigma}_t^{a,a} dt \right)^{-1} \int_0^1 \boldsymbol{\Sigma}_t^{a,b} dt.$$

If asset a is the market portfolio, then β_{ab} measures the average market beta over a period $[0, 1]$, e.g. a trading day, and may be used to price risk in a one-factor conditional CAPM model, whose importance in financial economics has been highlighted by Ferson & Harvey (1991), Jagannathan & Wang (1996), and Andersen, Bollerslev, Diebold & Wu (2006). Define RK^{ϵ}_{ab} as the (a, b)-th element of $RK^{\epsilon}(\boldsymbol{p})$, then ρ_{ab} and β_{ab} may be estimated robustly against non-synchronicity and MMS noise as $\rho^{RK^{\epsilon}}_{ab} = (RK^{\epsilon}_{aa} RK^{\epsilon}_{bb})^{-1/2} RK^{\epsilon}_{ab}$ and $\beta^{RK^{\epsilon}}_{ab} = (RK^{\epsilon}_{aa})^{-1} RK^{\epsilon}_{ab}$, respectively.

Corollary 2. *Under the conditions of Theorem 2, then for $a, b \in 1, \dots d$,*

$$n^{1/4} \left(\rho^{RK^{\epsilon}}_{ab} - \rho_{ab} \right) \xrightarrow{d_s(\mathcal{H}_1)} MN(0, \boldsymbol{\mathcal{B}}_{\rho_{ab}}(\lambda, \kappa)),$$
$$n^{1/4} \left(\beta^{RK^{\epsilon}}_{ab} - \beta_{ab} \right) \xrightarrow{d_s(\mathcal{H}_1)} MN(0, \boldsymbol{\mathcal{B}}_{\beta_{ab}}(\lambda, \kappa)),$$

where $\boldsymbol{\mathcal{B}}_{\rho_{ab}}(\lambda, \kappa)$ and $\boldsymbol{\mathcal{B}}_{\beta_{ab}}(\lambda, \kappa)$ are provided in Appendix A.3.

Proof. Follows by Theorems 1-2 in conjunction with the delta method. □

The realized correlation and regression coefficient estimates inherit optimal asymptotic properties from the flat-top realized kernels, thereby highlighting the importance of correcting for the bias caused by MMS noise and non-synchronicity when estimating quadratic covariation to avoid (Epps 1979)-type biases in non-linear transformations thereof as the sampling interval progressively shrinks. A feasible inference strategy for these quantities is directly available by applying the procedure in Remark 6.

5 Simulation Study

This section presents a simulation study to uncover how the choice of sampling scheme, refresh time sampling vs. Hayashi-Yoshida sampling, impacts the properties of the MMS noise in a synchronized

sample, which sampling scheme to choose when implementing the element-wise flat-top realized kernel, how eigenvalue truncation impacts non-positive semi-definite estimators, and, finally, it studies the relative finite sample performance of the flat-top realized kernels in comparison with other rate-optimal estimators in the literature and the realized kernel.

5.1 Simulation Design

The simulation design follows Barndorff-Nielsen et al. (2011) and Christensen et al. (2010). A standard 6.5-hour trading day on the NYSE is normalized to the unit interval, $t \in [0,1]$, such that 1 second corresponds to an increment of size $1/23400$. The efficient price diffusion is, then, simulated by a d-variate stochastic volatility model,

$$\begin{aligned} dp^*_{q,t} &= \mu_1 dt + \sigma_{q,t} dV_{q,t}, \quad \text{where} \quad \sigma_{q,t} = \exp(\beta_0 + \beta_1 f_{q,t}), \\ df_{q,t} &= \mu_2 f_{q,t} dt + dW_{q,t}, \quad dV_{q,t} = \varphi dW_{q,t} + \sqrt{1-\varphi^2} dB_t \quad \text{and} \quad W_{q,t} \perp\!\!\!\perp B_t, \end{aligned}$$

for $q = 1, \ldots, d$. Here, B_t and $W_{q,t}$ captures common and idiosyncratic uncertainty, respectively, and φ measures the leverage between $p^*_{q,t}$ and $f_{q,t}$. The parameter values are set in accordance with the literature as $(\mu_1 = 0.03, \beta_1 = 0.125, \mu_2 = -0.025, \varphi = -0.3, \beta_0 = \beta_1^2/(2\mu_2))'$, and the process is restarted on each "trading day" by drawing the initial observation from its stationary distribution $f_{q,t} \sim N(0, -1/(2\mu_2))$. To capture the effects of non-synchronicity, the observation times $t_i^{(q)}$, $i = 1, \ldots, N(1,q)$, are modeled by q independent Poisson processes with $\zeta = (\zeta_1, \ldots, \zeta_d)'$ controlling the average duration between observations. Further, let $\tilde{\eta}_{q,t_i^{(q)}} \sim N(0, \omega_{q,\eta})$, $\omega_{q,\eta} = \psi^2 (N^{-1} \sum_{i=1}^N \sigma^4_{q,t})^{1/2}$, be a sequences of i.i.d. normal variables, whose variance is determined by the noise-to-signal ratio, ψ^2, formally introduced in the next subsection. For all simulations, however, $\psi^2 = 0.005$ is fixed, which is consistent with the noise-to-signal ratios measured in Hansen & Lunde (2006) for DJIA stocks and in the empirical analysis below. The MMS noise is added through (3) using three data generating processes: **(DGP 1)** $\tilde{u}_{q,t_i^{(q)}} = \phi_q \tilde{u}_{q,t_{i-1}^{(q)}} + \tilde{\eta}_{q,t_i^{(q)}}$ where $\phi_q < 0$ $\forall q$, **(DGP 2)** $\tilde{u}_{q,t_i^{(q)}} = \tilde{\eta}_{q,t_i^{(q)}} + \theta_q \tilde{\eta}_{q,t_{i-1}^{(q)}}$ where $\theta_q < 0$ $\forall q$, and **(DGP 3)** is similar to DGP 1 with $\phi_q > 0$ $\forall q$. Negative AR(1) processes are consistent with the findings of (Aït-Sahalia et al. 2011, Figure 4) and (Ikeda 2013, Figure 2), positive AR(1) processes with the strategic learning models discussed and analyzed in Diebold & Strasser (2012) and/or clustering of order flow, see the discussions in Bandi & Russell (2006) and Ubukata & Oya (2009), and, finally, the impact of negative MA(1) and AR(1) processes is similar, but the former has shorter lasting effects on the efficient prices process, see also the discussion in Hansen, Large & Lunde (2008). The three DGP's generate non-trivial dependence in the observable log-returns, which is consistent with the empirical study. All simulations are performed using 1000 replications.

5.2 The Choice of Kernel and Tuning Parameters

Optimal bandwidth selection has been studied in the univariate case for asymptotically unbiased and rate-optimal kernel-based estimators with variance of the form (8) and MMS noise models of varying

complexity, see Barndorff-Nielsen et al. (2008), Ikeda (2013), and Varneskov (2013). Inspired by this, the advocated bandwidth selection method is of the form, $H = \kappa^* n^{1/2}$,

$$\kappa^* = f(\psi_q)\sqrt{\frac{\lambda^{(11)}}{\lambda^{(00)}}\left(1 + \sqrt{\frac{3\lambda^{(00)}\lambda^{(22)}}{(\lambda^{(11)})^2}}\right)}, \qquad \psi_q^2 = \frac{\boldsymbol{\Omega}^{q,q}}{\int_0^1 \boldsymbol{\Sigma}_t^{q,q} dt}, \tag{9}$$

where, e.g., $f(\psi_q) = \min_{q=1,\ldots,d} \psi_q$, $f(\psi_q) = \max_{q=1,\ldots,d} \psi_q$ or $f(\psi_q) = d^{-1}\sum_{q=1}^d \psi_q$, i.e. the global bandwidth is a function of the univariate mean squared error (MSE) optimal bandwidths using two approximations $\boldsymbol{\mathcal{Q}} = \int_0^1 \boldsymbol{\Sigma}_t dt \otimes \int_0^1 \boldsymbol{\Sigma}_t dt$ and $\boldsymbol{e}_{t_i} = \mathbf{0}$, $\forall i$. The first approximation provides an upward Jensen's inequality bias in the noise-to-signal ratio, ψ_q^2, while the second, excluding diurnal heteroskedasticity and endogeneity in the MMS noise, provides a downward bias in κ^*, (Varneskov 2013, Corollary 1). To accommodate these features, and following the empirical recommendations of Barndorff-Nielsen et al. (2009, 2011), $\boldsymbol{\Omega}^{q,q}$ and $\int_0^1 \boldsymbol{\Sigma}_t^{q,q} dt$ may, then, be estimated conservatively to balance the negative bias. Hence, $\boldsymbol{\Omega}^{q,q}$ is estimated using the upward biased, $n^{1/3}$-consistent estimator $\hat{\Omega}(\boldsymbol{p}, q) = (|\lambda^{(2)}|nG^{-2})^{-1} RK_{q,q}(\boldsymbol{p})$ of Ikeda (2013) where $G = n^{1/3}$ is the bandwidth, and a pilot estimate of $\int_0^1 \boldsymbol{\Sigma}_t^{q,q} dt$ is provided by the 20-minute sampled, subsampled and averaged realized variance estimator $RC_{20,q}^{sub}(\boldsymbol{p}, 1)$, which, in the d-variate case may be written as

$$RC_{20}^{sub}(\boldsymbol{p}, d) = \frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{18} \Delta \boldsymbol{p}_{t_{k+K(i-1)}} \Delta \boldsymbol{p}'_{t_{k+K(i-1)}} \tag{10}$$

where $K = 1200$ ensures the maximal degree of subsampling. Sparse sampling ameliorates the effects of MMS noise, and subsampling increases efficiency of the estimator. Additionally, the rule $f(\psi_q) = \max_{q=1,\ldots,d} \psi_q$ is selected along with the Parzen kernel,

$$\lambda(x) = (1 - 6x^2 + 6|x|^3)\mathbf{1}_{\{0 \leq |x| \leq 1/2]\}} + 2(1 - |x|)^3 \mathbf{1}_{\{1/2 \leq |x| \leq 1]\}},$$

for all kernel-based estimators. Last, γ, is fixed to balance the bias-variance tradeoff in Lemma 2. If $r = \tilde{q}$, (Varneskov 2013, Corollary 2) shows that $\gamma = (1/2 + \tilde{q})/(3/2 + \tilde{q}) = 3/5$ is a conservative MSE optimal flat-top shrinkage, which will be used throughout along with $\gamma = 2/5$ for robustness. The latter places higher emphasis on bias-reduction. Notice, the Parzen kernel has finite support on $x \in [-1, 1]$, which implies the support of the flat-top realized kernel is $x \in [-(1 + c), 1 + c]$.

5.3 Synchronization Effects and Return Autocorrelation

The impact of applying either refresh time sampling or Hayashi-Yoshida sampling to synchronize the "raw" series is assessed by examining the properties of the synchronized time series for the leading asset using the configuration, $d = 2$, $\zeta = (5, 30)'$, and letting all persistence parameters in the noise DGP's 1-3 be $\pm 1/2$. The choice of ζ corresponds to observing the two assets every 5, respectively, 30 seconds on average, suggesting that the properties of the synchronized series for the slowest trading asset are

almost unaffected by scheme since all observations are kept (on average). The time series properties for the leading asset is examined by estimating the autocorrelation function (ACF), the short-run variance of the MMS noise as $\hat{\omega} = (2n)^{-1} \sum_{i=1}^{n} \Delta p_{1,t_i}^2$, and the long run variance by $\hat{\Omega}(\boldsymbol{p}, 1) = \hat{\Omega}$. As the latter is an upward biased estimator, the long run variance of the MMS noise is also estimated by a bias-corrected version $\hat{\Omega}_{BC}(\boldsymbol{p}, 1) = \hat{\Omega}_{BC}$, (Ikeda 2011), which in the multivariate case is defined as

$$\hat{\Omega}_{BC}(\boldsymbol{p}, d) = (1 - \tau^2)^{-1} \left(|\lambda^{(2)}(0)| n G^{-2} \right)^{-1} (RK(\boldsymbol{p}, G) - RK(\boldsymbol{p}, H)) \tag{11}$$

where $\tau = G/H$, $H = \kappa n^{1/2}$ and $G = n^{\tilde{\nu}}$ for $\tilde{\nu} \in \left[(2\tilde{q} + 1)^{-1}, 1/2 \right]$. Here, $\tilde{\nu} = 1/2$ is selected to emphasize bias reduction. The ACF's and MMS noise-variance estimates are presented in Figure 1 and the top three rows of Table 1, respectively.

(Figure 1 and Table 1 around here)

A distinct pattern emerges in Figure 1. Both synchronization schemes seem to add positive persistence to the series, which is especially pronounced for Hayashi-Yoshida sampling. When compared to the properties of a raw series with AR(1) noise and increasing, positive persistence parameters, a similar pattern emerges as ϕ approaches unity. The patterns are not expected to perfectly coincide as synchronization errors is a form of endogenous MMS noise, while the latter is exogenous, but the comparison merely serves as an illustration. As alluded to in Remark 3, the stronger persistence pattern for the Hayashi-Yoshida sampling scheme is not surprising as next-tick interpolation at the end-points generates positive persistence by construction. The long run noise-variance estimates in Table 1 elaborate on these results by showing extremely upward biased estimates from $\hat{\Omega}$ and unstable, even negative, estimates from $\hat{\Omega}_{BC}$ when the series have been synchronized using Hayashi-Yoshida sampling, suggesting it may be favorable to use refresh time sampling for kernel-based estimators.

To validate this conjecture, the remaining parts of Table 1 show the relative bias and root mean squared error (RMSE) of the estimated covariance between asset 1 and 2, $\boldsymbol{\Sigma}_{12} = \int_0^1 \boldsymbol{\Sigma}_t^{1,2} dt$, using flat-top realized kernels in conjunction with either refresh time sampling or Hayashi-Yoshida sampling. Further, it shows the relative bias and RMSE of $\boldsymbol{\Sigma}_{12}^{\epsilon}$, the off-diagonal element from a corresponding element-wise estimate of the whole covariance matrix whose eigenvalues have been truncated by $\epsilon_n = n^{-1/2}$ to ensure positive definiteness. The specific choice of ϵ_n preludes the conclusion from the next subsection. A few remarks before describing the results. First, bandwidth selection for the diagonal elements of an element-wise covariance matrix estimator collapses to a univariate selection rule. Second, the bandwidth for Hayashi-Yoshida sampled series is selected using $f(\psi_q) = \min_{q=1,\ldots,d} \psi_q$ since $\hat{\Omega}$ is extremely upward biased thereby suggesting that $f(\psi_q) = \max_{q=1,\ldots,d} \psi_q$ may suffer from similar distortions. Changing $f(\psi_q)$, however, does not alter the conclusions. Last, eigenvalues are only truncated for negative semi-definite estimates. This is a special case of the truncation rule in Section 4.1, which will be used throughout. Table 1 shows that refresh time sampling delivers the best results, thus verifying the conjecture made above and in Remark 3 that the positive persistence generated by Hayashi-Yoshida sampling has detrimental finite sample effects on kernel-based estimators.

5.4 The Impact of Truncating Eigenvalues

While Section 4.1 provides a simple correction to ensure positive definiteness of asymptotically unbiased and rate-optimal estimators, the choice of truncation level $\epsilon_n = o(n^{-1/4}) \in \mathbb{R}_+$ offers little guidance for practical implementation. To investigate said choice, the simulation study is expanded to three assets, $d = 3$, where DGP's 1 and 3 are implemented with persistence parameters $(\pm 0.3, \pm 0.4, \pm 0.5)'$ and DGP 2 with $(-0.4, -0.5, -0.7)'$. This setup is fixed in the remaining parts of the simulation study whereas the level of non-synchronicity, $\zeta = (\zeta_1, \zeta_2, \zeta_3)'$, is varied. Here, $\zeta = (5, 10, \zeta_3)'$, $\zeta_3 = (20, 30)'$, to vary the fraction of kept data, ϑ. From the comparison of estimators in Section 5.6, the element-wise flat-top realized kernel with $\gamma = 3/5$, $ERK^*_{3/5}$, shows by far the most occurrences of negative semi-definite estimates, and is, thus, used as the subject for this investigation. Table 2 shows the number of binding eigenvalue truncations for $ERK^*_{3/5}$, the mean relative RMSE of the unique elements of the quadratic covariation matrix for various choice of ϵ_n, and similarly for four non-linear transformations thereof, namely the correlations and betas fixing asset 1 as the base asset.

(Table 2 around here)

Table 2 shows that the mean relative RMSE of the unique elements of the quadratic covariation matrix are decreasing as ϵ_n decreases, whereas it starts to increase for the non-linear transformations when $\epsilon_n = o(n^{-1/2})$ since the matrix becomes exceedingly unstable as its smallest eigenvalue approaches zero. Hence, the preferred truncation rule is $\epsilon_n = n^{-1/2}$, which seems to strike a balance between the two counteracting effects.

5.5 Related and Competing Estimators

The class of flat-top realized kernel estimators provides efficiency gains relative to the realized kernels. However, there are other noteworthy alternatives in the rapidly expanding literature, which provide interesting comparisons. Two of these are the two-scale realized kernel (Ikeda 2011) and the pre-averaged realized covariance estimator (Christensen et al. 2010) where a new bias-correction is proposed for the latter to accommodate more general forms of MMS noise (than i.i.d. dependence).

5.5.1 The Two-Scale Realized Kernel

The two-scale realized kernel (TSRK) shares the generic structure of the realized kernels, but uses a generalized jack-knife kernel function,

$$k(x, \tau) = \left(1 - \tau^2\right)^{-1} \left\{\lambda(x) - \tau^2 \lambda(x/\tau)\right\},$$

for $\lambda(x) \in \mathcal{K}$, where, again, $\tau = G/H$, $H = \kappa n^{1/2}$ and $G = n^{\tilde{\nu}}$ for $\tilde{\nu} \in \left[(2\tilde{q}+1)^{-1}, 1/2\right]$. The asymptotic similarities and differences between the TSRK and flat-top realized kernels are studied by Varneskov (2013) in the univariate case for the selections $\tilde{\nu} = \left((2\tilde{q}+1)^{-1}, 1/2\right)'$, which emphasizes

MSE and bias reduction, respectively. More generally, the former is conveyed by a conditional moment comparison of Lemma 2 against (under Assumptions 1-6 and $\tilde{q} \leq r$)

$$\mathbb{E}[TSRK(\boldsymbol{p})|\mathcal{H}_1] = \int_0^1 \boldsymbol{\Sigma}_t dt + nH^{-2}G^{-\tilde{q}}k_{\tilde{q}}^{(2)} \sum_{h\in\mathbb{Z}} |h|^{\tilde{q}} \boldsymbol{\Omega}(h) + o_p(1)$$
$$\operatorname*{plim}_{n\to\infty} n^{1/2}\mathbb{V}[TSRK(\boldsymbol{p})|\mathcal{H}_1] = \lim_{n\to\infty} \boldsymbol{\mathcal{B}}(\Phi,\kappa)$$

where $\Phi^{(jj)}(\tau) = \lambda^{(jj)} + f_j(\tau)$, $j = 0,1,2$, $f_j(\tau) \in R_+$ and $f_j(\tau) = O(\tau^2)$. Hence, the two estimators differ in terms of finite sample bias and variance, and potentially even in terms of asymptotic variance. By selecting the flat-top shrinkage suitably as $\gamma \in (0, (1+r-2\tilde{\nu}\tilde{q})/(1+r))$, the flat-top realized kernels achieve a higher-order advantage in terms of bias reduction. The difference in terms of finite sample variance, however, depends on the finite sample inflation of the characteristic parameters of orders $O(n^{-\gamma/2})$, see Lemma 2, $O(n^{2(\tilde{\nu}-1/2)})$, and on the relative weights of $\boldsymbol{\mathcal{Q}}$, $\boldsymbol{\mathcal{N}}$, and $\boldsymbol{\mathcal{C}}$. However, if $\tilde{\nu} = 1/2$ such that the TSRK places maximal emphasis on bias reduction, $f_j(\tau) = O(1)$ $\forall j = 0,1,2$, implying that the flat-top realized kernels have a strictly lower asymptotic variance.

5.5.2 Pre-averaged Realized Covariance

Following Christensen et al. (2010), let $M = \bar{\theta}n^{\bar{\nu}}$, $\bar{\nu} \in (0,1)$, be a sequence of integers, $g(x) = \min(x, 1-x)$ a non-zero weight function, and define the modulated realized covariance estimator,

$$MRC(\boldsymbol{p}) = \sum_{i=0}^{n-M} \bar{\boldsymbol{p}}_{t_i}\bar{\boldsymbol{p}}_{t_i}', \qquad \bar{\boldsymbol{p}}_{t_i} = \sum_{j=1}^{M} g\left(\frac{j}{M}\right)\Delta\boldsymbol{p}_{t_{i+j}}.$$

Then, under Assumptions 1-3, 5, 6, and $\boldsymbol{e}_{t_i} = \boldsymbol{0}$ $\forall i$, it follows by (Varneskov 2013, Lemma 2) in conjunction with the Cramér-Wold Theorem, e.g. (Davidson 2002, Theorem 25.6), that for $\bar{\nu} = 1/2$

$$\frac{1}{\varpi_2\bar{\theta}n^{1/2}} MRC(\boldsymbol{p}) \xrightarrow{\mathbb{P}} \int_0^1 \boldsymbol{\Sigma}_t dt + \frac{\varpi_1}{\bar{\theta}^2\varpi_2}\boldsymbol{\Omega}, \qquad \boldsymbol{\Omega} = \boldsymbol{\Omega}^{(uu)}, \tag{12}$$

where $\varpi_1 = \int_0^1 [g^{(1)}(x)]^2 dx$ and $\varpi_2 = \int_0^1 [g(x)]^2 dx$. The modulated realized covariance estimator successfully balances the asymptotic orders of $\Delta\boldsymbol{p}^*$ and $\Delta\boldsymbol{U}$ such that its consistency at the optimal rate, the choice of $\bar{\nu} = 1/2$ is analogous to $\nu = 1/2$, depends only on a bias-correction. Now, careful inspection of (Ikeda 2011, Lemmas 4-9) shows that under the same conditions and as long as $\tilde{q} \leq r$, the bias-corrected estimator of the long-run MMS noise-variance has the following properties

$$\hat{\Omega}_{BC}(\boldsymbol{p}, d) = \boldsymbol{\Omega} + O_p(G^{-\tilde{q}}) + \boldsymbol{\mathcal{Z}}_{\mathcal{N}}(1 + o_p(1)) \tag{13}$$

where $\boldsymbol{\mathcal{Z}}_{\mathcal{N}} \overset{d_s(\mathcal{H}_1)}{\to} MN(\boldsymbol{0}, \lim_{n\to\infty}\boldsymbol{\mathcal{B}}(\mathcal{N}))$, $\boldsymbol{\mathcal{B}}(\mathcal{N}) = O_p(Gn^{-1}) + O_p(HG^4n^{-3}) + O_p(G^3n^{-2})$. Hence, a generally corrected pre-averaged realized covariance estimator, $PARC(\boldsymbol{p})$, readily follows by combining

(12) and (13). Since it is based on a similar jack-knife bias-correction as the TSRK and since the weight function $g(x) = \min(x, 1-x)$ is equivalent to using a Parzen kernel, see Christensen et al. (2010), the finite sample performances of $PARC(\boldsymbol{p})$ and $TSRK(\boldsymbol{p})$ are expected to be similar. Note, however, that (12) excludes an endogenous noise component. Including the latter will necessitate a more complicated bias-correction that depends on $\boldsymbol{\Omega}^{(ep)}$.

5.6 Estimates of Quadratic Covariation and Non-linear Transformations

The class of (element-wise) flat-top realized kernel estimators is compared to the subsampled realized covariance estimator using 20-minute intervals, $RC_{20}^{sub}(\boldsymbol{p}, 3)$, the realized kernel with a bandwidth $H = 3.51f(\psi)n^{3/5}$ where $f(\psi)$ is computed as in Section 5.2, the TSRK, which is implemented with bandwidths $H = \max(\kappa^* n^{1/2}, G+1)$, $G = n^{\tilde{\nu}}$, $\tilde{\nu} = (1/3, 1/2)'$, and the pre-averaged realized covariance (PARC) estimator using $\bar{\theta} = 1$ and the same configuration as the TSRK for the jack-knife bias-correction. This follows along the lines of the respective authors. Neither the TSRK nor the PARC estimator is guaranteed to be positive semi-definite. However, Theorem 3 shows that the advocated truncation rule of Section 5.4 may be applied. The relative bias and RMSE of the estimated elements of the quadratic covariation matrix are shown for two non-synchronous configurations, $\zeta = (5, 10, 20)'$ and $\zeta = (3, 3, 30)'$ in Tables 3 and 4. Note that the liquidity and data loss for the two portfolios resemble the empirical analysis below.

(Tables 3 and 4 around here)

Tables 3 and 4 feature several noteworthy observations. First, $RC_{20}^{sub}(\boldsymbol{p}, 3)$ performs worst by all measures. Second, Table 3 shows that both flat-top realized kernels estimate all elements of the covariance matrix with smaller bias and RMSE compared to the realized kernel, and that the biggest efficiency gains are from the specification with $\gamma = 3/5$. This illustrates the gains from the flat-top bias-correction, thus complementing the asymptotic results in Lemmas 1-2 and Theorem 1. Table 4 shows similar results, though slightly worse for the shrinkage $\gamma = 2/5$. The relative emphasis on bias and variance for flat-top realized kernels with $\gamma = 3/5$ and $\gamma = 2/5$ are as expected from Lemma 2. Third, the TSRK and the PARC estimator perform similarly. When placing maximal emphasis on bias reduction, $\tilde{\nu} = 1/2$, both estimators successfully eliminates the MMS noise-induced bias, but at the expense of a variance inflation. In fact, both estimators have higher RMSE's than the realized kernel. For the MSE optimal choice, $\tilde{\nu} = 1/3$, the PARC estimator performs slightly better than the TSRK in terms of bias and vice versa for RMSE's. This difference is potentially caused by fixing $\bar{\theta} = 1$, as Christensen et al. (2010) suggest, instead of selecting it optimally. Both estimators, however, provide efficiency gains relative to the realized kernel, but their performance is also unstable as none of the two are able to control the bias for DGP 3, showing biases in the 7-15% range. Fourth, the difference between the rate-optimal estimators is quite suggestive. The flat-top realized kernel with $\gamma = 3/5$ places itself between the MSE optimal TSRK and PARC estimators in terms of RMSE, and it performs well for all DGP's. Hence, among estimators who offer stable bias control, counting

the two flat-top realized kernels and the TSRK and PARC estimators with $\tilde{\nu} = 1/2$, the flat-top realized kernel with $\gamma = 3/5$ is the most efficient, illustrating the higher-order advantage of the flat-top approach. Finally, the element-wise flat-top realized kernel offer efficiency gains for some elements of the covariance matrix by utilizing more observations. However, the estimator displays by far the most occurrences of negative semi-definite estimates, illustrating potential matrix instability problems.

To see the potential implications of these results for applications, Table 5 shows the relative bias and RMSE of four non-linear transformations of the covariance matrix, namely the correlations and betas fixing asset 1 as the base.

(Table 5 around here)

Once again, $RC_{20}^{sub}(\boldsymbol{p}, 3)$ performs worst for all transformations, and the flat-top realized kernel with $\gamma = 3/5$ performs better than the realized kernel in terms of both bias and RMSE. After the realized covariance estimator, the realized kernel has the largest bias among the estimators, except for a few cases where it is surpassed by the element-wise flat-top realized kernel, which, not surprisingly, delivers unstable results. The MSE optimal TSRK cannot control the bias for DGP 3, and while the PARC counterpart performs better, it still shows signs of unstable bias control for DGP 3 when the data loss is large. In terms of RMSE's, the flat-top realized kernel with $\gamma = 3/5$ is comparable to the MSE optimal TSRK and performs slightly better than TSRK emphasizing bias reduction. It has slightly higher RMSE's than both PARC estimators when estimating correlations, but its performance is on par with the MSE optimal PARC estimator and slightly better than the PARC estimator emphasizing bias reduction when estimating betas. Hence, it offers a desirable combination of robustness and efficiency across a variety of noise models.

Finally, note that synchronization errors is the only source of endogenous noise in the present simulation setup, and it is even an order of magnitude lower than the exogenous noise component. As noted in Section 5.5.2, if an endogenous noise component of the same order of magnitude is present in the DGP, the bias-correction for the PARC estimator may be insufficient, thus potentially leading to larger distortions than shown in Tables 3-5.

6 Empirical Analysis

To illustrate the relevance of the theoretical results, an empirical analysis of correlations and market betas is performed for a portfolio of six stocks using tick-by-tick trade data from 2007. A special thanks goes to Asger Lunde for providing the cleaned high-frequency data. The six stocks are: International Business Machines (IBM), Exxon Mobil (XOM), Intel (INTC), Microsoft (MSFT), Total System Services (TSS), an IT firm from Georgia, and Standard & Poor's Depository Receipt (SPY), an exchange traded fund that tracks the S&P 500. The portfolio is heavy on IT stocks of varying liquidity and size, and individual summary statistics are presented in Table 6.

(Table 6 around here)

Table 6 shows that the TSS stock is 5-10 times less liquid than the remaining stocks, and its inclusion in the portfolio leads to a dramatic data loss. Further, the estimated noise-to-signal ratios for refresh time sampled data illustrate how false reliance on an i.i.d. noise assumption understates the magnitude of the noise by an approximate factor 5 relative to using the bias-corrected long-run MMS noise-variance estimator (11), and that the selected noise-to-signal ratio $\psi^2 = 0.005$ in the simulation study seems appropriate. Additionally, Table 6 shows that the estimated 97.5% and 2.5% quantiles for the first five lags of the individual ACF's of intra-daily log-returns are much wider than a corresponding conservative confidence band under a white-noise null hypothesis, which is approximately $\pm 1/900 \approx \pm 0.0011$ on average, thereby suggesting the presence of a MMS noise component with non-trivial dependence. The signs of the respective quantiles also suggest that the noise should allow for a sign-alternating autocorrelation pattern on some trading days and positive persistence on others. Last, Figure 2 displays the 97.5% and 2.5% quantiles for the first ten lags of the cross-autocorrelation function between the pairs (IBM, TSS) and (SPY, TSS) where both stocks have been used as the base asset.

(Figure 2 around here)

Figure 2 shows a pronounced asymmetric lead-lag dependence pattern where TSS is led by IBM and the market proxy, SPY, similar to the findings of Large (2007), Voev & Lunde (2007), and Griffin & Oomen (2011) for London Stock Exchange and DJIA stocks.

The Brownian semimartingale plus i.i.d. noise model cannot explain the higher noise-to-signal ratios for long run noise-variance estimators, the wide confidence bands for the log-return ACF's, nor the asymmetric lead-lag patterns, thus motivating the development of a general additive noise model for asset prices and asymptotically unbiased and rate-optimal estimators of its quadratic covariation.

6.1 Correlation Analysis

Table 6 shows that the actual high-frequency data is similar to the simulated data in Section 5. There are large variations in liquidity of the six stocks, the data loss is substantial, the noise-to-signal ratio is comparable, and the MMS noise component is non-trivial. Hence, Table 7 shows the unconditional average of daily correlation estimates from the flat-top realized kernels with $\gamma = (2/5, 3/5)'$ and the biases and RMSE's of the 20-minute subsampled realized covariance estimator, the realized kernel, the TSRK with $\tilde{\nu} = (1/3, 1/2)'$, the PARC estimator with $\tilde{\nu} = 1/3$, and the outer product of open-to-close returns (OTOC) using the flat-top realized kernel with $\gamma = 3/5$ as a proxy for the true quadratic covariation matrix to analyze the relative performance of the estimators. The PARC estimator with $\tilde{\nu} = 1/2$ and the element-wise flat-top realized kernel are excluded since they do not provide positive definite estimates of quadratic covariation on all trading days.

(Table 7 around here)

The stocks are positively correlated with magnitudes that seem divided into three tiers. The strongest and weakest correlations are between the SPY and TSS, respectively, and the remaining stocks. The

latter illustrates the potential diversification benefits of including an asset outside of the DJIA in the portfolio. Table 7 also shows that the relative magnitudes of the biases and RMSE's for the various estimators are in line with the simulation results, which suggests that the flat-top realized kernels provide accurate information about the correlation structure of asset returns since they offer a desirable combination of robustness and efficiency. Two noteworthy features from Table 7 require a deeper investigation, however. These are the upward bias for OTOC and both the bias and relatively high RMSE for the PARC estimator. Hence, Table 8 shows the average quadratic covariation estimates of the two estimators in comparison with those from the flat-top realized kernel and the TSRK.

(Table 8 around here)

The average off-diagonal elements for the three high-frequency estimators are highly similar and the deviations from the OTOC equivalents are also relatively minor. The biggest difference comes from the average estimates of the diagonal elements where both the OTOC and the PARC estimators have pronounced negative biases. The OTOC estimates are very noisy and may not adequately capture intra-daily variation. The negative bias for the PARC estimator, on the other hand, may be caused by negative correlation between the efficient prices and the MMS noise, features which have already been documented by (Hansen & Lunde 2006, Fact I) and Diebold & Strasser (2012), and which the generalized bias-correction in (12)-(13) is unable to account for.

6.2 Realized Market Beta Estimation

As a second application, this subsection analyzes daily market beta estimates for TSS using SPY as a proxy of the market portfolio. Note, however, that even though TSS is used as an example, almost identical results are obtained for the other stocks in the portfolio. Figure 3 shows the estimated market betas for the flat-top realized kernel with $\gamma = 3/5$ along with a smoothed series from an ARMA(1,1) filter with estimated intercept $\mu = 0.84$ and persistence parameters $(\phi, \theta)' = (0.90, -0.79)'$ that are consistent with the findings of Barndorff-Nielsen et al. (2011) and Christensen et al. (2010). Further, it shows scatterplots of the market beta estimates from the flat-top realized kernel against equivalent ones from the remaining high-frequency estimators of the previous subsection.

(Figure 3 around here)

Figure 3 illustrates that that the daily market beta is time-varying, predictable, and it exhibits stationary fluctuations around its unconditional mean $\mu = 0.84$. Further, it shows that, except for the MSE optimal TSRK, the remaining estimators provide noisy estimates of the market betas, and that the 20-minute subsampled realized covariation estimator suffers from the largest dispersion. These results are elaborated upon in Table 9 by running the regressions corresponding to the scatterplots.

(Table 9 around here)

Table 9 confirms the visual analysis by showing the largest dispersion for the 20-minute subsampled realized covariation estimator, followed by the TSRK emphasizing bias reduction, the realized kernel, and the PARC estimator. With exception of the latter, this exact pattern is observed in Table 5. However, as explained previously, endogenous noise may distort the PARC estimator, potentially causing this dispersion. Last, Table 9 shows a strong coherence between the flat-top realized kernel and the MSE optimal TSRK. The Wald test of an unbiased estimator, however, is strongly rejected. Similarly, by subtracting the two beta series, i.e. by imposing a unity slope, the estimated difference in mean is positive, $\beta_0 = 0.0357$, and significant with a HAC robust t-statistic of $t_{HAC} = 5.59$. Hence, in conjunction with Table 5, this illustrates the higher-order advantage of the flat-top realized kernels over the TSRK. The former is able to maintain a small variance while still providing (asymptotically) unbiased estimates, thus illustrating the theoretical results in Sections 3 and 4.2.

7 Conclusion

The paper develops a general, multivariate additive noise model for synchronized asset prices and extends the generalized class of univariate flat-top realized kernels, analyzed in Varneskov (2013), to estimate its quadratic covariation. The noise model allows for α-mixing dependent exogenous noise, random sampling, and an endogenous noise component that encompasses synchronization errors, asymmetric lead-lag relations, and diurnal heteroskedasticity, thus accommodating a wide variety of empirical regularities at tick-by-tick frequencies. The flat-top realized kernels are shown to be consistent, asymptotically unbiased, and mixed Gaussian at the optimal rate of convergence, $n^{1/4}$. A simple eigenvalue correction guarantees the class of estimators to be positive definite without altering its asymptotic properties. This allows the computation of non-linear transformations of the estimated covariance matrix where the two leading examples in this paper are the realized correlation and regression coefficient. However, other potential applications include ranking of multivariate volatility models, Laurent et al. (2012), and dynamic mean-variance analysis in the spirit of Chiriac & Voev (2011) and Varneskov & Voev (2013). An empirically motivated simulation study shows that refresh time sampling is preferable to Hayashi-Yoshida sampling for kernel-based estimators, it provides empirical guidelines for eigenvalue truncation, and it shows that flat-top realized kernels have a desirable combination of robustness and efficiency relative to competing estimators in the literature. The latter is reinforced by an empirical analysis of correlations and market betas for a portfolio of six stocks of varying size and liquidity.

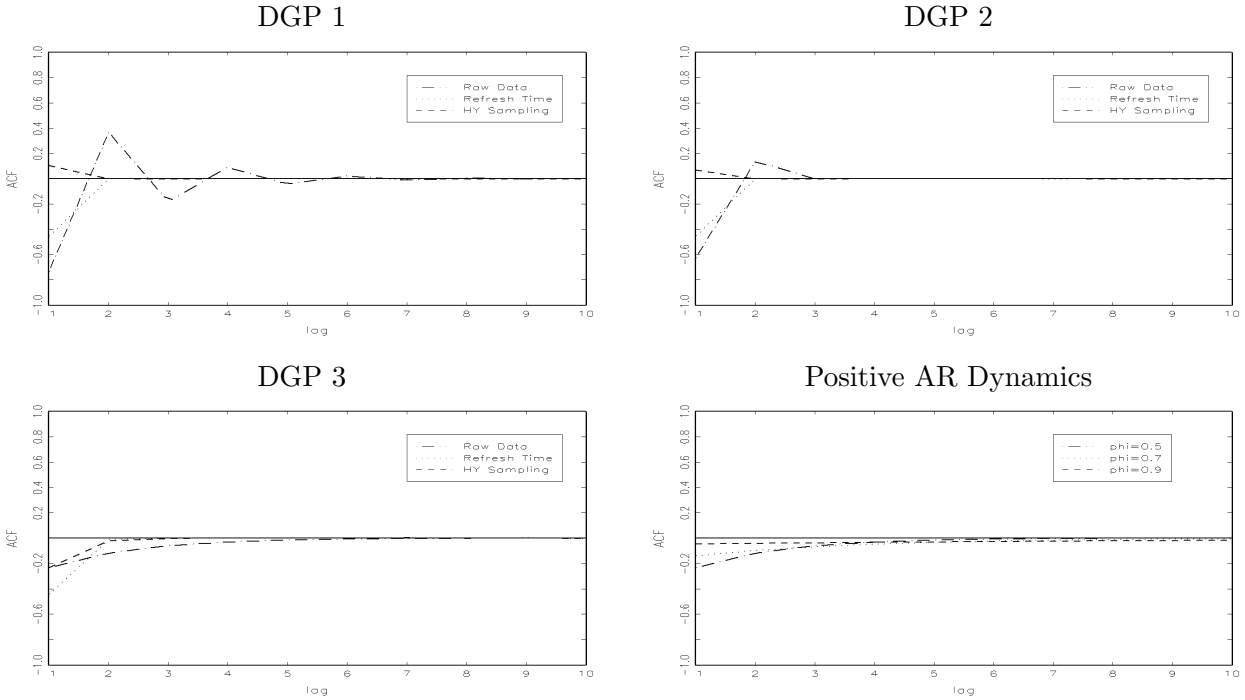

Figure 1: Panels 1-3 show the autocorrelation function (ACF) of the most liquid of two non-synchronously observed assets using three DGP's for observable returns, non-synchronous configuration $\zeta = (5, 30)$ and three synchronized series: The raw data (dashed, dotted), refresh time sampling (dotted), and Hayashi-Yoshida (HY) sampling (dashed). DGP 1 models the MMS noise as a negative AR(1) process, DGP 2 as a negative MA(1), and DGP 3 as a positive AR(1). All persistence parameters are $\pm 1/2$. Panel 4 shows the ACF of the raw return series where MMS is modeled as a positive AR(1) process with persistence parameters $\phi = (0.5, 0.7, 0.9)'$.

Impact of Synchronization Scheme

	RTS			**HYS**			**No Synchronization**		
	DGP 1	DGP 2	DGP 3	DGP 1	DGP 2	DGP 3	DGP 1	DGP 2	DGP 3
$\hat{\omega}$	0.72	0.68	0.70	0.73	0.69	0.72	1.00	0.88	0.34
$\hat{\Omega}_{BC}$	0.63	0.60	0.70	-0.46	-0.50	0.64	0.27	0.18	1.82
$\hat{\Omega}$	1.65	1.62	1.67	14.57	13.02	6.30	0.75	0.69	1.99
Bias: $\mathbf{\Sigma}_{12}$	-1.57	-1.60	-1.68	13.98	14.33	13.67	-	-	-
Bias: $\mathbf{\Sigma}^{\epsilon}_{12}$	-2.48	-2.46	-2.23	2.88	3.80	5.91	-	-	-
RMSE: $\mathbf{\Sigma}_{12}$	21.10	20.96	23.02	58.30	55.28	46.44	-	-	-
RMSE: $\mathbf{\Sigma}^{\epsilon}_{12}$	19.88	19.78	22.40	45.07	42.39	37.08	-	-	-
$\mathbb{P}_{tr}$	13.4	12.7	8.80	52.4	53.0	51.0	-	-	-

Table 1: Estimates of the MMS noise-variance (scaled by 100) by the short run variance estimator $\hat{\omega}$ and the two long run variance estimators $\hat{\Omega}$ and $\hat{\Omega}_{BC}$ using the series for the leading asset after refresh time sampling (RTS), Hayashi-Yoshida sampling (HYS), or without synchronization. Moreover, the relative bias and RMSE (in percentages) of $\mathbf{\Sigma}_{12}$ shows the properties of a flat-top realized kernel estimate with $\gamma = 3/5$. The corresponding results for $\mathbf{\Sigma}^{\epsilon}_{12}$ is for the off-diagonal element of an element-wise flat-top realized kernel estimate with $\gamma = 3/5$, $ERK^{*}_{3/5}$, whose eigenvalues are truncated by $\epsilon_n = n^{-1/2}$ to ensure positive definiteness. $\mathbb{P}_{tr}$ denotes the percentage of binding *eig*envalue truncations, i.e. $\mathbb{P}_{tr} = \#(eig \leq 0)/1000\%$.

Impact of Eigenvalue Truncation

	DGP 1			**DGP 2**			**DGP 3**		
$\zeta_3 = 20$	$\mathbb{P}_{tr}$	M_1	M_2	$\mathbb{P}_{tr}$	M_1	M_2	$\mathbb{P}_{tr}$	M_1	M_2
$\epsilon = n^{-1/4}$	19.1	29.52	13.19	16.1	26.40	12.66	15.2	30.16	13.76
$\epsilon = n^{-1/3}$	19.1	21.54	10.80	16.1	19.97	10.44	15.2	23.19	11.82
$\epsilon = n^{-1/2}$	19.1	16.22	9.15	16.1	15.76	9.08	15.2	18.87	10.56
$\epsilon = n^{-2/3}$	19.1	15.34	9.04	16.1	15.05	9.11	15.2	18.22	10.49
$\epsilon = n^{-1}$	19.1	15.17	9.16	16.1	14.90	9.25	15.2	18.10	10.60
	DGP 1			**DGP 2**			**DGP 3**		
$\zeta_3 = 30$	$\mathbb{P}_{tr}$	M_1	M_2	$\mathbb{P}_{tr}$	M_1	M_2	$\mathbb{P}_{tr}$	M_1	M_2
$\epsilon = n^{-1/4}$	26.7	33.02	15.01	23.4	32.08	14.84	22.1	30.52	15.77
$\epsilon = n^{-1/3}$	26.7	24.39	12.34	23.4	23.78	12.32	22.1	24.32	13.29
$\epsilon = n^{-1/2}$	26.7	18.18	10.25	23.4	17.87	10.38	22.1	20.31	11.57
$\epsilon = n^{-2/3}$	26.7	16.98	10.04	23.4	16.74	10.20	22.1	19.59	11.50
$\epsilon = n^{-1}$	26.7	16.70	10.18	23.4	16.48	10.33	22.1	19.41	11.66

Table 2: Mean relative RMSE of the unique individual elements $\int_0^1 \Sigma_t^{ij} dt$, denoted M_1, and of four non-linear transformations (the correlations and betas using asset 1 as the base asset), denoted M_2, for the element-wise flat-top realized kernel for $\gamma = 3/5$, $ERK^*_{3/5}$. The simulations are performed using the non-synchronous configuration $\zeta = (5, 10, \zeta_3)'$, $\zeta_3 = (20, 30)'$ and noise-to-signal ratio $\psi^2 = 0.005$. $\mathbb{P}_{tr}$ denotes the percentage of binding *eig*envalue truncations, i.e. $\mathbb{P}_{tr} = \#(eig \leq 0)/1000\%$. Note that for $\zeta_3 = (20, 30)'$, $(n, \vartheta)' \approx \{(1170, 0.43), (780, 0.30)\}'$. All numbers are in percentages.

Estimation of Quadratic Covariation for $n \approx 1170$ and $\vartheta \approx 0.43$

		Relative Bias						Relative RMSE					
	$\mathbb{P}_{tr}$	$\mathbf{\Sigma}_{11}$	$\mathbf{\Sigma}_{12}$	$\mathbf{\Sigma}_{22}$	$\mathbf{\Sigma}_{13}$	$\mathbf{\Sigma}_{23}$	$\mathbf{\Sigma}_{33}$	$\mathbf{\Sigma}_{11}$	$\mathbf{\Sigma}_{12}$	$\mathbf{\Sigma}_{22}$	$\mathbf{\Sigma}_{13}$	$\mathbf{\Sigma}_{23}$	$\mathbf{\Sigma}_{33}$
DGP 1													
RC_{20}^{sub}	-	11.79	-8.62	14.01	-9.72	-8.97	16.63	27.96	27.78	29.18	28.15	28.23	30.95
RK	-	3.04	-1.88	4.13	-1.84	-1.45	1.86	20.94	21.02	21.14	21.17	21.06	20.67
$TSRK_1$	0.10	0.99	-1.12	1.68	-0.96	-0.83	0.08	17.02	16.96	16.98	16.98	16.73	16.26
$TSRK_2$	0.00	-0.10	-0.76	0.92	-0.85	-0.31	1.02	24.42	25.24	24.27	25.35	25.26	24.50
$PARC_1$	0.00	-0.71	-0.69	0.32	-0.71	-0.21	-0.37	19.62	20.20	19.69	20.34	20.32	19.65
$PARC_2$	0.10	-1.36	-0.81	-0.26	-0.92	-0.28	-0.18	24.75	25.59	24.59	25.61	25.65	24.77
$RK^*_{3/5}$	0.30	-0.50	-1.35	0.44	-1.25	-0.97	0.44	18.06	18.39	18.22	18.54	18.31	17.58
$RK^*_{2/5}$	0.40	-0.85	-1.56	0.34	-1.54	-1.07	0.46	20.22	20.94	20.60	20.94	20.84	20.27
$ERK^*_{3/5}$	19.1	1.66	-0.93	1.12	-1.94	-1.35	1.21	12.58	13.92	13.89	17.56	17.91	21.47
DGP 2													
RC_{20}^{sub}	-	12.91	-8.62	15.09	-9.76	-9.01	19.43	28.44	27.76	29.69	28.14	28.23	32.48
RK	-	3.32	-1.86	3.99	-1.86	-1.41	1.05	20.98	20.95	20.98	21.06	20.95	20.36
$TSRK_1$	0.20	1.11	-1.07	1.39	-0.98	-0.77	-2.14	17.09	16.91	16.78	16.80	16.54	15.85
$TSRK_2$	0.00	-0.04	-0.77	1.08	-0.89	-0.31	1.15	24.46	25.22	24.21	25.31	25.21	24.46
$PARC_1$	0.00	-0.72	-0.69	0.14	-0.74	-0.21	-1.48	19.63	20.20	19.62	20.28	20.24	19.45
$PARC_2$	0.00	-1.39	-0.82	-0.25	-0.96	-0.30	-0.41	24.78	25.57	24.52	25.60	25.60	24.66
$RK^*_{3/5}$	0.20	-0.46	-1.29	0.69	-1.27	-0.88	0.75	18.07	18.33	18.08	18.38	18.12	17.43
$RK^*_{2/5}$	0.30	-0.82	-1.51	0.60	-1.53	-0.98	0.67	20.19	20.87	20.42	20.76	20.68	19.90
$ERK^*_{3/5}$	16.1	1.40	-1.02	0.99	-1.86	-1.15	2.28	11.26	13.71	13.49	17.52	17.79	20.79
DGP 3													
RC_{20}^{sub}	-	11.73	-8.63	14.02	-9.60	-8.86	16.78	27.99	27.86	29.30	28.28	28.36	31.43
RK	-	2.31	-1.74	4.87	-1.89	-1.44	10.58	22.15	22.46	22.63	22.86	22.81	26.45
$TSRK_1$	0.10	0.79	-1.19	3.22	-1.10	-0.96	14.51	17.48	17.65	18.27	18.73	18.91	26.33
$TSRK_2$	0.40	-0.07	-0.66	1.20	-0.73	-0.10	2.63	24.48	25.28	24.42	25.64	25.59	25.86
$PARC_1$	0.00	-0.65	-0.64	1.25	-0.66	-0.11	7.23	19.77	20.31	19.98	21.04	21.06	23.16
$PARC_2$	1.30	-1.32	-0.70	0.01	-0.81	-0.13	1.29	24.78	25.61	24.73	25.91	25.93	26.00
$RK^*_{3/5}$	1.70	-0.46	-1.35	0.47	-1.45	-1.10	4.66	18.89	19.36	19.42	20.31	20.32	23.48
$RK^*_{2/5}$	7.10	-0.39	-1.55	0.46	-1.98	-1.30	1.31	22.13	22.13	21.84	22.72	22.71	24.57
$ERK^*_{3/5}$	15.2	1.53	-1.29	1.87	-1.85	-1.13	5.93	14.28	15.56	16.97	19.85	20.34	26.20

Table 3: Relative bias and RMSE of the individual elements of $\mathbf{\Sigma}_{ij} = \int_0^1 \mathbf{\Sigma}_t^{ij} dt$ for six competing estimators of quadratic covariation: Realized covariance with 20-minute sampling and subsampling, RC_{20}^{sub}, the realized kernel, RK, the two-scale realized kernel, $TSRK_j$, where $j = (1,2)'$ correspond to choosing $\tilde{\nu} = (1/3, 1/2)'$, pre-averaged realized covariance estimator with a jack-knife bias-correction, $PARC_j$, the flat-top realized kernel, RK^*_γ, with $\gamma = (2/5, 3/5)'$, and the element-wise flat-top realized kernel, $ERK^*_{3/5}$. The simulations are performed using the non-synchronous configuration $\zeta = (5, 10, 20)'$ and noise-to-signal ratio $\psi^2 = 0.005$. $\mathbb{P}_{tr}$ denotes the percentage of binding *eig*envalue truncations, i.e. $\mathbb{P}_{tr} = \#(eig \leq 0)/1000\%$. All numbers are in percentages.

Estimation of Quadratic Covariation for $n \approx 780$ and $\vartheta \approx 0.15$

		Relative Bias						Relative RMSE					
	$\mathbb{P}_{tr}$	$\mathbf{\Sigma}_{11}$	$\mathbf{\Sigma}_{12}$	$\mathbf{\Sigma}_{22}$	$\mathbf{\Sigma}_{13}$	$\mathbf{\Sigma}_{23}$	$\mathbf{\Sigma}_{33}$	$\mathbf{\Sigma}_{11}$	$\mathbf{\Sigma}_{12}$	$\mathbf{\Sigma}_{22}$	$\mathbf{\Sigma}_{13}$	$\mathbf{\Sigma}_{23}$	$\mathbf{\Sigma}_{33}$
DGP 1													
RC_{20}^{sub}	-	11.82	-8.10	14.05	-11.04	-10.70	16.48	27.97	27.72	29.15	28.42	28.57	30.91
RK	-	3.05	-1.93	3.98	-1.88	-1.45	1.73	23.39	23.58	23.71	23.72	23.91	23.32
$TSRK_1$	0.10	0.79	-1.26	1.66	-1.14	-0.83	0.02	19.27	19.29	19.30	19.17	19.42	18.67
$TSRK_2$	0.00	0.35	-0.69	0.99	-0.70	-0.17	1.12	27.26	28.19	27.47	28.34	28.54	27.55
$PARC_1$	0.00	-0.22	-0.59	0.59	-0.80	-0.28	-0.33	21.88	22.43	21.87	22.51	22.69	21.78
$PARC_2$	0.00	-0.67	-0.63	0.01	-0.85	-0.24	-0.02	27.44	28.33	27.43	28.53	28.63	27.59
$RK_{3/5}^*$	0.60	-0.66	-1.60	0.15	-1.49	-1.09	-0.29	20.76	21.24	20.90	21.09	21.41	20.28
$RK_{2/5}^*$	0.60	-0.96	-1.92	-0.13	-1.88	-1.40	0.11	23.35	24.15	23.80	23.94	24.40	23.27
$ERK_{3/5}^*$	27.1	2.38	0.24	2.69	-2.70	-2.61	2.85	11.28	10.48	11.53	19.50	19.74	28.45
DGP 2													
RC_{20}^{sub}	-	12.93	-8.11	15.11	-11.02	-10.69	19.34	28.45	27.72	29.67	28.41	28.55	32.50
RK	-	3.33	-1.93	4.20	-1.88	-1.49	0.99	23.49	23.64	23.80	23.70	23.91	23.13
$TSRK_1$	0.40	0.91	-1.31	1.67	-1.18	-0.90	-2.00	19.44	19.40	19.36	19.17	19.37	18.42
$TSRK_2$	0.00	0.44	-0.67	1.08	-0.70	-0.20	1.21	27.28	28.24	27.52	28.27	28.49	27.40
$PARC_1$	0.00	-0.18	-0.59	0.58	-0.80	-0.32	-1.50	21.93	22.48	21.89	22.47	22.64	21.65
$PARC_2$	0.00	-0.64	-0.60	0.05	-0.83	-0.26	-0.23	27.45	28.37	27.44	28.45	28.58	27.40
$RK_{3/5}^*$	0.60	-0.58	-1.64	0.12	-1.52	-1.16	0.62	20.90	21.35	20.96	21.10	21.40	20.43
$RK_{2/5}^*$	0.60	-0.87	-1.92	-0.13	-1.89	-1.48	0.31	23.43	24.26	23.87	23.92	24.39	23.23
$ERK_{3/5}^*$	24.0	2.04	0.07	2.01	-2.53	-2.45	4.45	10.27	10.13	10.59	19.62	19.77	28.04
DGP 3													
RC_{20}^{sub}	-	11.83	-8.06	14.09	-11.11	-10.76	16.37	28.04	27.75	29.24	28.64	28.80	31.36
RK	-	2.31	-1.89	3.33	-2.09	-1.52	10.37	24.43	24.80	24.57	25.38	25.50	28.75
$TSRK_1$	0.00	0.56	-1.25	1.64	-1.54	-1.12	13.69	19.70	19.94	19.88	20.59	21.01	27.36
$TSRK_2$	0.90	0.22	-0.77	1.03	-0.77	-0.04	3.11	27.34	28.13	27.24	28.93	28.99	29.31
$PARC_1$	0.00	-0.29	-0.63	0.67	-0.87	-0.24	7.95	22.05	22.54	21.93	23.25	23.39	25.39
$PARC_2$	1.20	-0.79	-0.74	0.02	-0.93	-0.14	1.95	27.51	28.28	27.25	29.13	29.09	29.30
$RK_{3/5}^*$	1.00	-0.74	-1.58	0.39	-1.93	-1.37	4.42	21.59	22.16	21.68	22.75	23.04	25.15
$RK_{2/5}^*$	4.20	-1.08	-1.98	0.30	-2.41	-1.64	1.09	24.62	25.32	25.00	25.95	26.11	27.48
$ERK_{3/5}^*$	22.0	2.08	-0.14	2.47	-2.73	-2.31	7.34	12.32	11.87	13.70	21.70	21.83	31.46

Table 4: Relative bias and RMSE of the individual elements of $\mathbf{\Sigma}_{ij} = \int_0^1 \mathbf{\Sigma}_t^{ij} dt$ for six competing estimators of quadratic covariation: Realized covariance with 20-minute sampling and subsampling, RC_{20}^{sub}, the realized kernel, RK, the two-scale realized kernel, $TSRK_j$, where $j = (1,2)'$ correspond to choosing $\tilde{\nu} = (1/3, 1/2)'$, pre-averaged realized covariance estimator with a jack-knife bias-correction, $PARC_j$, the flat-top realized kernel, RK_γ^*, with $\gamma = (2/5, 3/5)'$, and the element-wise flat-top realized kernel, $ERK_{3/5}^*$. The simulations are performed using the non-synchronous configuration $\zeta = (3,3,30)'$ and noise-to-signal ratio $\psi^2 = 0.005$. $\mathbb{P}_{tr}$ denotes the percentage of binding *eig*envalue truncations, i.e. $\mathbb{P}_{tr} = \#(eig \leq 0)/1000\%$. All numbers are in percentages.

Estimation of Non-linear Transformation

	Relative Bias								Relative RMSE							
	$\zeta = (5, 10, 20)'$				$\zeta = (3, 3, 30)'$				$\zeta = (5, 10, 20)'$				$\zeta = (3, 3, 30)'$			
	ρ_{12}	ρ_{13}	β_{12}	β_{13}	ρ_{12}	ρ_{13}	β_{12}	β_{13}	ρ_{12}	ρ_{13}	β_{12}	β_{13}	ρ_{12}	ρ_{13}	β_{12}	β_{13}
DGP 1																
RC_{20}^{sub}	-20.09	-21.99	-19.11	-20.11	-19.66	-23.12	-18.68	-21.34	21.06	22.96	20.83	21.93	20.65	24.05	20.41	23.03
RK	-5.49	-4.38	-4.78	-4.74	-5.53	-4.41	-4.85	-4.79	6.53	5.67	8.65	8.66	6.90	5.99	9.46	9.58
$TSRK_1$	-2.39	-1.51	-1.81	-1.68	-2.45	-1.55	-1.74	-1.60	4.72	4.00	7.94	7.64	5.27	4.26	8.81	8.50
$TSRK_2$	-1.40	-1.51	-0.56	-0.63	-1.61	-1.69	-0.91	-0.91	4.34	4.38	8.96	8.95	4.95	4.92	10.14	10.18
$PARC_1$	-0.57	-0.27	0.17	0.13	-0.88	-0.65	-0.19	-0.40	3.30	3.17	7.52	7.49	3.88	3.63	8.31	8.33
$PARC_2$	-0.18	-0.30	0.72	0.64	-0.52	-0.74	0.24	-0.01	3.85	3.79	9.11	9.15	4.37	4.43	10.06	10.12
$RK^*_{3/5}$	-1.33	-1.30	-0.62	-0.54	-1.40	-1.10	-0.70	-0.55	4.14	4.05	7.88	7.79	4.85	4.32	8.96	8.80
$RK^*_{2/5}$	-1.37	-1.45	-0.52	-0.50	-1.50	-1.59	-0.74	-0.65	4.57	4.30	8.71	8.53	5.21	4.84	10.03	9.71
$ERK^*_{3/5}$	-2.30	-3.29	-2.43	-3.56	-2.12	-4.92	-1.89	-5.05	6.07	8.38	9.08	13.06	4.82	10.12	6.18	16.32
DGP 2																
RC_{20}^{sub}	-20.88	-23.37	-19.94	-20.97	-20.46	-24.44	-19.51	-22.12	21.85	24.32	21.58	22.71	21.44	25.36	21.18	23.74
RK	-5.53	-4.15	-5.01	-5.01	-5.76	-4.19	-5.11	-5.05	6.59	5.51	8.78	8.78	7.12	5.84	9.63	9.62
$TSRK_1$	-2.26	-0.46	-1.86	-1.79	-2.57	-0.63	-1.90	-1.75	4.62	3.60	7.97	7.55	5.40	3.89	8.94	8.28
$TSRK_2$	-1.51	-1.65	-0.60	-0.72	-1.68	-1.78	-0.97	-1.00	4.42	4.45	9.00	8.96	5.04	5.00	10.24	10.09
$PARC_1$	-0.48	0.27	0.18	0.12	-0.90	-0.07	-0.24	-0.44	3.27	3.04	7.49	7.31	3.93	3.42	8.35	8.19
$PARC_2$	-0.18	-0.23	0.75	0.62	-0.53	-0.63	0.24	-0.01	3.84	3.71	9.11	9.06	4.43	4.36	10.14	10.01
$RK^*_{3/5}$	-1.41	-1.50	-0.59	-0.59	-1.46	-1.62	-0.81	-0.65	4.19	3.99	7.91	7.66	4.97	4.34	9.08	8.61
$RK^*_{2/5}$	-1.48	-1.57	-0.49	-0.50	-1.56	-1.75	-0.84	-0.75	4.64	4.31	8.74	8.39	5.32	4.81	10.13	9.50
$ERK^*_{3/5}$	-2.26	-3.72	-2.38	-3.35	-1.86	-5.47	-1.81	-4.62	5.97	8.67	8.74	12.95	4.45	10.60	5.78	16.49
DGP 3																
RC_{20}^{sub}	-20.08	-21.91	-19.08	-19.96	-19.64	-23.15	-18.64	-21.41	21.07	22.91	20.84	21.92	20.63	24.13	20.39	23.27
RK	-5.40	-7.98	-3.97	-4.12	-4.88	-8.18	-4.10	-4.36	6.63	9.02	8.66	9.34	6.49	9.48	9.37	10.44
$TSRK_1$	-3.11	-7.87	-1.72	-1.70	-2.36	-7.86	-1.56	-1.82	5.22	9.67	8.21	9.67	5.00	9.76	8.61	10.46
$TSRK_2$	-1.42	-2.10	-0.47	-0.49	-1.65	-2.61	-0.80	-0.86	4.56	5.32	9.13	10.05	4.94	6.00	10.13	11.47
$PARC_1$	-0.99	-3.84	0.20	0.13	-0.93	-4.59	-0.14	-0.41	3.63	5.92	7.85	9.13	3.88	6.81	8.40	9.95
$PARC_2$	-0.21	-0.87	0.80	0.74	-0.58	-1.67	0.29	0.03	4.06	4.74	9.27	10.27	4.34	5.55	10.03	11.42
$RK^*_{3/5}$	-1.35	-3.29	-0.67	-0.82	-1.50	-3.59	-0.61	-0.96	4.41	6.32	8.30	9.84	4.72	6.80	9.07	10.78
$RK^*_{2/5}$	-1.59	-2.23	-0.80	-1.22	-1.74	-2.36	-0.65	-1.10	5.21	6.11	9.60	11.00	5.29	6.67	10.24	11.99
$ERK^*_{3/5}$	-2.84	-5.19	-2.50	-3.21	-2.17	-6.88	-1.92	-4.69	6.55	9.69	10.73	15.28	5.61	11.89	7.37	18.53

Table 5: Relative bias and RMSE of four non-linear transformations for six competing estimators of quadratic covariation: Realized covariance with 20-minute sampling and subsampling, RC_{20}^{sub}, the realized kernel, RK, the two-scale realized kernel, $TSRK_j$, where $j = (1, 2)'$ correspond to choosing $\tilde{\nu} = (1/3, 1/2)'$, pre-averaged realized covariance estimator with a jack-knife bias-correction, $PARC_j$, the flat-top realized kernel, RK^*_γ, with $\gamma = (2/5, 3/5)'$, and the element-wise flat-top realized kernel, $ERK^*_{3/5}$. The simulations are performed using the non-synchronous configurations $\zeta = (5, 10, 20)'$ and $\zeta = (3, 3, 30)'$. All numbers are in percentages.

Summary Statistics

	$\bar{n}$	$\vartheta(q)$	ψ_1^2	ψ_2^2	ψ_3^2	ACF(1)	ACF(2)	ACF(3)	ACF(4)	ACF(5)
IBM	6560	0.14	0.09	0.41	1.29	(-0.01,-0.22)	(0.09,-0.05)	(0.06,-0.05)	(0.05,-0.05)	(0.05,-0.05)
XOM	10380	0.09	0.09	0.43	1.28	(0.06,-0.21)	(0.12,-0.06)	(0.08,-0.08)	(0.05,-0.06)	(0.04,-0.05)
INTC	5173	0.19	0.11	0.64	1.34	(0.01,-0.42)	(0.05,-0.03)	(0.04,-0.04)	(0.03,-0.04)	(0.03,-0.03)
MSFT	5277	0.19	0.10	0.62	1.37	(0.03,-0.38)	(0.06,-0.05)	(0.04,-0.05)	(0.04,-0.04)	(0.04,-0.04)
SPY	9573	0.10	0.08	0.39	1.20	(0.13,-0.19)	(0.07,-0.02)	(0.04,-0.04)	(0.03,-0.04)	(0.02,-0.03)
TSS	908	1.00	0.12	0.30	1.32	(0.05,-0.35)	(0.14,-0.12)	(0.16,-0.12)	(0.13,-0.12)	(0.11,-0.11)

Table 6: Summary statistics for six stocks over the 251 trading days in 2007. Here, $\bar{n}$ denotes the average number of observations, $\vartheta(q)$ denotes the average fraction of kept data for asset q, ψ_j^2, $j = 1, 2, 3$, corresponds to the average noise-to-signal ratio using $\hat{\omega}$, $\hat{\Omega}_{BC}$, and $\hat{\Omega}$, respectively, to estimate the noise-variance and $RC^{sub}_{20,q}(\boldsymbol{p}, 1)$ to proxy the integrated quarticity. All noise-to-signal ratios have all been scaled by 100. ACF(j), $j = 1, \ldots, 5$ shows the estimated 97.5% and 2.5% quantiles of the autocorrelation function.

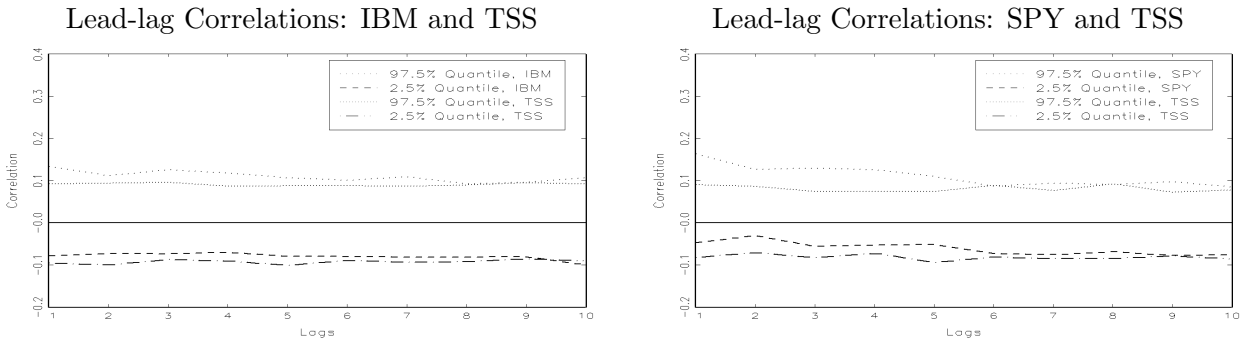

Figure 2: 97.5% and 2.5% quantiles of the first ten lags of the cross-autocorrelation function for the two pairs (IBM, TSS) and (SPY, TSS). Both assets have been used as the base asset. When TSS is the base asset, the 97.5% quantile is (dotted, frequent) and the 2.5% quantile is (dashed, dotted). When (IBM, SPY) is the base asset, the 97.5% quantile is (dotted, sparse) and the 2.5% quantile is (dashed).

Correlation Analysis

	$RK^*_{3/5} \setminus RK^*_{2/5}$						Bias: Open-to-close Covariance $\setminus RC^{sub}_{20}$					
	IBM	XOM	INTC	MSFT	SPY	TSS	IBM	XOM	INTC	MSFT	SPY	TSS
IBM		0.34	0.38	0.38	0.55	0.24		-1.92	-0.92	-1.50	-1.89	-3.37
XOM	0.33		0.33	0.34	0.62	0.22	12.54		-3.57	-4.14	-6.19	-1.46
INTC	0.37	0.33		0.40	0.57	0.23	18.68	15.43		0.32	0.80	-2.11
MSFT	0.38	0.34	0.39		0.57	0.25	9.53	13.27	14.66		-0.20	-4.40
SPY	0.55	0.62	0.57	0.57		0.40	9.53	12.38	10.31	6.39		-5.18
TSS	0.23	0.22	0.23	0.24	0.39		2.44	14.21	8.03	10.52	12.14	
	Bias: $RK \setminus PARC_1$						Bias: $TSRK_1 \setminus TSRK_2$					
	IBM	XOM	INTC	MSFT	SPY	TSS	IBM	XOM	INTC	MSFT	SPY	TSS
IBM		1.90	1.09	1.18	1.72	1.45		1.28	0.85	0.12	0.32	0.10
XOM	0.66		1.28	0.67	2.06	2.43	-0.34		0.28	-0.46	-0.26	0.44
INTC	0.20	0.08		0.89	1.45	1.75	-0.56	-0.12		0.21	0.72	0.45
MSFT	-0.27	-0.41	-0.14		1.18	1.76	-0.25	0.18	-0.13		-0.51	-0.09
SPY	-0.07	-0.14	0.21	-0.53		1.78	-0.26	0.26	-0.45	0.13		0.45
TSS	-0.56	-0.11	-0.16	-0.68	-0.53		-0.91	-0.55	-0.57	-0.85	-1.27	
	RMSE: $RK \setminus PARC_1$						RMSE: $TSRK_1 \setminus TSRK_2$					
	IBM	XOM	INTC	MSFT	SPY	TSS	IBM	XOM	INTC	MSFT	SPY	TSS
IBM		7.86	7.72	7.54	4.89	8.79		7.96	8.31	8.00	6.66	8.64
XOM	5.54		7.73	6.62	5.43	8.83	2.85		8.05	7.87	5.60	8.03
INTC	5.37	5.52		6.35	4.83	7.55	3.21	3.18		7.98	5.70	8.64
MSFT	5.56	5.08	5.13		5.42	8.37	3.13	3.14	3.19		5.75	8.76
SPY	4.55	3.84	3.64	3.80		5.61	2.76	2.35	2.55	2.61		7.91
TSS	5.88	5.37	6.09	5.92	5.55		3.22	3.47	3.32	3.23	3.09	

Table 7: Panel 1 shows the average correlation matrix for the flat-top realized kernel, RK^*_γ, with $\gamma = (2/5, 3/5)'$. Panels 2-4 show the estimated bias of the open-to-close covariance estimator, the 20-minute subsampled realized covariance estimator, RC^{sub}_{20}, the realized kernel, RK, the two-scale realized kernel, $TSRK_j$, where $j = (1, 2)'$ correspond to $\tilde{\nu} = (1/3, 1/2)'$, and the pre-averaged realized covariance estimator with a jack-knife bias-correction, $PARC_1$. Panels 5-6 show the estimated RMSE of the last four (consistent) estimators. The bias and RMSE's have been computed using $RK^*_{3/5}$ as a proxy for the true ex-post covariance matrix and both have been scaled by 100.

Covariance Analysis

	$RK^*_{3/5} \setminus TSRK_1$						Open-to-close Covariance $\setminus PARC_1$					
	IBM	XOM	INTC	MSFT	SPY	TSS	IBM	XOM	INTC	MSFT	SPY	TSS
IBM		0.58	0.63	0.55	0.52	0.43		0.57	0.61	0.54	0.52	0.44
XOM	0.57		0.71	0.63	0.73	0.53	0.61		0.68	0.60	0.72	0.55
INTC	0.63	0.70		0.70	0.65	0.53	0.75	0.77		0.68	0.64	0.54
MSFT	0.55	0.62	0.69		0.57	0.47	0.57	0.67	0.77		0.56	0.49
SPY	0.52	0.72	0.65	0.56		0.52	0.49	0.76	0.68	0.57		0.53
TSS	0.44	0.53	0.53	0.48	0.53		0.38	0.65	0.55	0.55	0.58	
	Diagonal						Diagonal - diag($RK^*_{3/5}$)					
	IBM	XOM	INTC	MSFT	SPY	TSS	IBM	XOM	INTC	MSFT	SPY	TSS
$TSRK_1$	1.30	1.82	1.87	1.43	0.62	2.88	0.35	2.43	3.65	2.34	0.62	-1.82
$PARC_1$	1.22	1.70	1.74	1.33	0.61	2.57	-7.83	-9.92	-9.61	-7.25	-0.56	-32.29
OTOC	1.14	1.59	1.60	1.27	0.65	2.00	-16.18	-21.18	-24.19	-13.54	2.78	-89.06

Table 8: Panel 1-2 shows the average estimated off-diagonal elements of the quadratic covariance matrix for the flat-top realized kernel with $\gamma = 3/5$, $RK^*_{3/5}$, the two-scale realized kernel and pre-averaged realized covariance estimator with $\tilde{\nu} = 1/3$, $TSRK_1$ and $PARC_1$, respectively, and an open-to-close covariance estimator (OTOC). Panels 3-4 show the average estimated diagonal elements and their bias using $RK^*_{3/5}$ as a proxy for the true ex-post covariance matrix. The bias has been scaled by 100.

Realized Beta Regressions

	β_0	β_1	R^2	Wald	AR_1	AR_2	AR_3
RC^{sub}_{20}	0.4304^{**} (0.0423)	0.5848^{**} (0.0525)	0.4398	51.70^{**}	4.1118^{*}	6.1973^{*}	12.033^{**}
RK	0.1337^{**} (0.0343)	0.8583^{**} (0.0388)	0.8834	7.649^{**}	0.0167	5.2590	5.2806
$TSRK_1$	-0.0337^{**} (0.0098)	1.0861^{**} (0.0115)	0.9725	35.72^{**}	22.98^{**}	29.94^{**}	35.36^{**}
$TSRK_2$	0.2173^{**} (0.0398)	0.7322^{**} (0.0431)	0.8154	19.82^{**}	1.3207	8.9466^{*}	9.0691^{*}
$PARC_1$	0.0089 (0.0324)	0.9820^{**} (0.0362)	0.8977	0.4579	3.6595	4.2035	5.0445

Table 9: Estimates of the constant, β_0, and slope, β_1, from univariate regressions of the market beta estimates from the flat-top realized kernel with $\gamma = 3/5$ on equivalent estimates from the 20-minute subsampled realized covariance estimator, RC^{sub}_{20}, the realized kernel, RK, the two-scale realized kernel, $TSRK_j$, where $j = (1, 2)'$ correspond to $\tilde{\nu} = (1/3, 1/2)'$, and the pre-averaged realized covariance estimator with a jack-knife bias-correction, $PARC_1$. Wald is a test of the joint hypothesis $\beta_0 = 0$ and $\beta_1 = 1$. Both inference and hypothesis testing is based on Andrews (1991) HAC standard errors. AR_j, $j = 1, 2, 3$ is an LM test (with j lags) of the null hypothesis of no serial correlation.

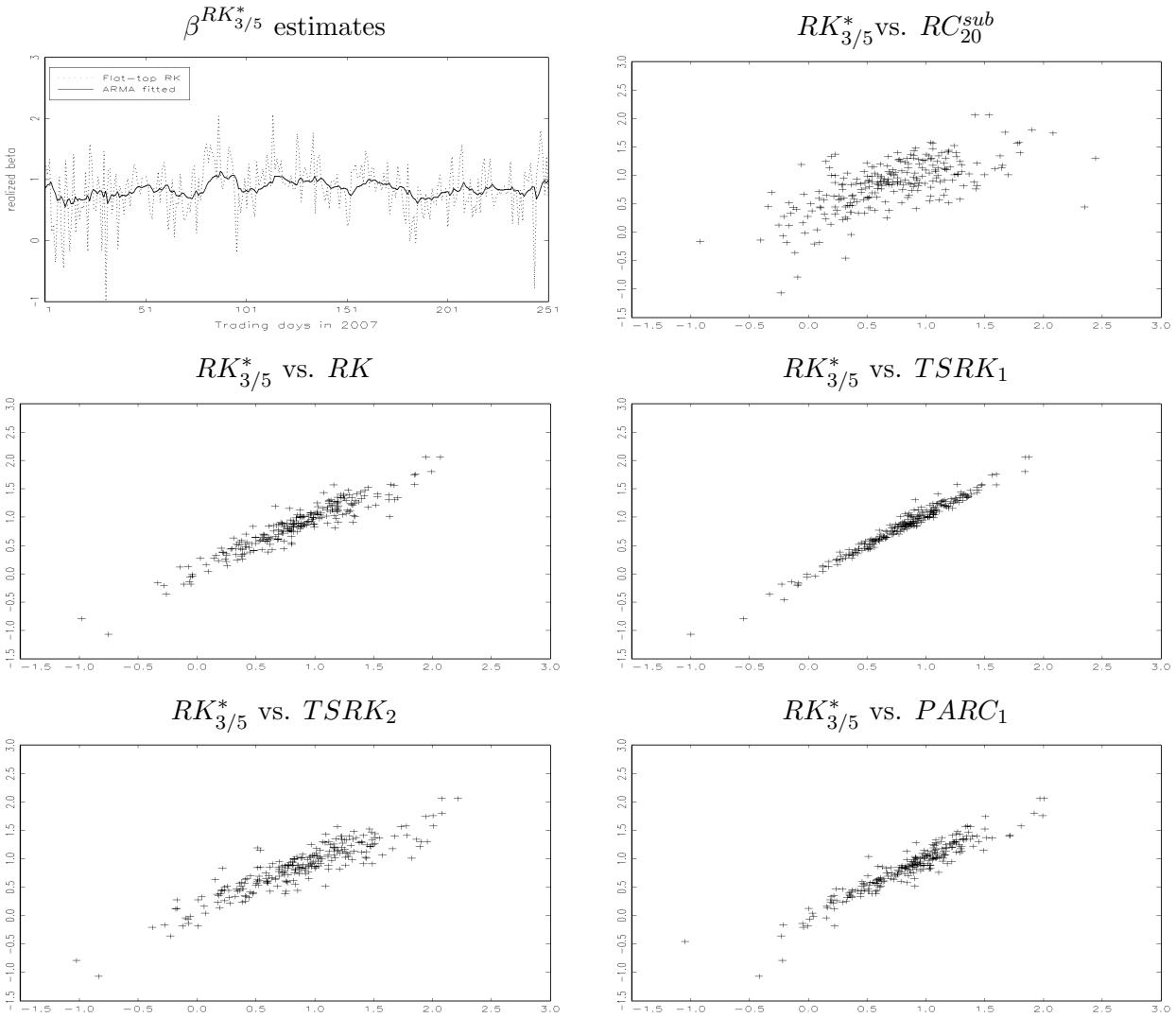

Figure 3: Panel 1 shows the estimated time series of market betas for TSS using the flat-top realized kernel with $\gamma = 3/5$, $RK^*_{3/5}$, and a smoothed series using an ARMA(1,1) filter. Panels 2-6 shows scatter plots of the market beta estimates using $RK^*_{3/5}$ (on the y-axis) against equivalents from the 20-minute subsampled realized covariance estimator, RC^{sub}_{20}, the realized kernel, RK, the two-scale realized kernel, $TSRK_j$, where $j = (1,2)'$ correspond to $\tilde{\nu} = (1/3, 1/2)'$, and the pre-averaged realized covariance estimator with a jack-knife bias-correction, $PARC_1$.

References

Abadir, K. M. & Magnus, J. R. (2005), *Matrix Algebra*, Cambridge University Press.

Aït-Sahalia, Y., Fan, J. & Xiu, D. (2010), 'High frequency covariance estimates with noisy and asynchronous financial data', *Journal of the American Statistical Association* **105**, 1504–1517.

Aït-Sahalia, Y., Mykland, P. A. & Zhang, L. (2011), 'Ultra high frequency volatility estimation with dependent microstructure noise', *Journal of Econometrics* **161**(1), 160–175.

Andersen, T., Bollerslev, T., Diebold, F. & Wu, G. (2006), Realized beta: Persistence and predictability, *in* T. Fomby & D. Terrell, eds, 'Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series', Vol. 20, Elsevier Science.

Andersen, T. G. & Benzoni, L. (2012), Stochastic volatility, *in* R. A. Meyers, ed., 'Encyclopedia of Complexity and Systems Science', Springer-Verlag. forthcoming.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modeling and forecasting realized volatility', *Econometrica* **71**, 579–625.

Andrews, D. W. (1991), 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica* **59**(3), 817–858.

Bandi, F. M. & Russell, J. R. (2006), Full-information transaction costs. Working paper, Graduate School of Business, The University of Chicago.

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

Barndorff-Nielsen, O. E., Hansen, P., Lunde, A. & Shephard, N. (2009), 'Realized kernels in practice: Trades and quotes', *Econometrics Journal* **12**(3), C1–C32.

Barndorff-Nielsen, O. E., Hansen, P., Lunde, A. & 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.

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

Chiriac, R. & Voev, V. (2011), 'Modelling and forecasting multivariate realized volatility', *Journal of Applied Econometrics* **28**, 922–947.

Christensen, K., Kinnebrock, S. & Podolskij, M. (2010), 'Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data', *Journal of Econometrics* **159**, 116–133.

Dahlhaus, R. (2009), 'Local inference for locally stationary time series based on the empirical spectral measure', *Journal of Econometrics* **151**, 101–112.

Dahlhaus, R. & Polonik, W. (2009), 'Empirical spectral processes for locally stationary time series', *Bernoulli* **15**, 1–39.

Davidson, J. (2002), *Stochastic Limit Theory*, Oxford University Press.

de Jong, F. & Nijman, T. (1997), 'High frequency analysis of lead-lag relationships in financial markets', *Journal of Empirical Finance* **4**, 259–277.

Diebold, F. X. & Strasser, G. (2012), 'On the correlation structure of microstructure noise: A financial economics perspective', *Review of Economic Studies* . forthcoming.

Epps, T. (1979), 'Comovements in stock prices in the very short run', *Journal of the American Statistical Association* **74**, 291–298.

Ferson, W. E. & Harvey, C. R. (1991), 'The variation of economic risk premiums', *Journal of Political Economy* **99**(2), 385–415.

Glosten, L. R. & Milgrom, P. R. (1985), 'Bid, ask and transaction prices in a specialist market with heterogeneously informed traders', *Journal of Financial Economics* **14**, 71–100.

Griffin, J. E. & Oomen, R. C. A. (2011), 'Covariance measurement in the presence of non-synchronous trading and market microstructure noise', *Journal of Econometrics* **160**(1), 58–68.

Hansen, P. R., Large, J. & Lunde, A. (2008), 'Moving averaged-based estimators of integrated variance', *Econometric Reviews* **27**, 79–111.

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

Hayashi, T. & Yoshida, N. (2005), 'On covariance estimation of non-synchronously observed diffusion processes', *Bernoulli* **11**, 359–379.

Ikeda, S. (2011), A bias-corrected rate-optimal estimator of the covariation matrix of multiple security returns with dependent and endogenous microstructure effect. Unpublished manuscript.

Ikeda, S. S. (2013), Two scale realized kernels: A univariate case. Unpublished manuscript.

Jacod, J. (2009), Statistics and high frequency data. SEMSTAT Seminar. Technical report, Univ. de Paris-6.

Jacod, J. & Todorov, V. (2009), 'Testing for common arrival of jumps in discretely-observed multidimensional processes', *Annals of Statistics* **37**, 1792–1838.

Jagannathan, R. & Wang, Z. (1996), 'The conditional CAPM and the cross-section of expected returns', *Journal of Finance* **51**(1), 3–53.

Kalnina, I. (2011), 'Subsampling high frequency data', *Journal of Econometrics* **161**(2), 262–283.

Kalnina, I. & Linton, O. (2008), 'Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error', *Journal of Econometrics* **147**(1), 47–59.

Large, J. (2007), 'Accounting for the Epps effect: Realized covariation, cointegration and common factors'. Unpublished manuscript, Oxford-Man Institute, University of Oxford.

Laurent, S., Rombouts, J. & Violante, F. (2012), 'Consistent ranking of multivariate volatility models', *Journal of Econometrics* . forthcoming.

Lunde, A., Shephard, N. & Sheppard, K. (2011), Econometric analysis of vast covariance matrices using composite realized kernels. Unpublished Manuscript, University of Oxford.

Mancini, C. & Gobbi, F. (2012), 'Identifying the Brownian covariation from the co-jumps given discrete observations', *Econometric Theory* **28**(2), 249–273.

Martens, M. (2004), Estimating unbiased and precise realized covariances. Econometric Institute, Erasmus University Rotterdam.

Mykland, P. A. (2010), 'A Gaussian calculus for incerence from high frequency data', *Annals of finance* **8**, 235–258.

Phillips, P. C. & Yu, J. (2008), Information loss in volatility measurement with flat price trading. Unpublished Manuscript, Cowles foundation for research in economics, Yale University.

Podolskij, M. & Vetter, M. (2010), 'Understanding limit theorems for semimartingales: A short survey', *Statistica Nederlandica* **64**, 329–351.

Politis, D. N. (2011), 'Higher-order accurate, positive semidefinite estimation of large-sample covariance and spectral density matrices', *Econometric Theory* **27**, 703–744.

Roll, R. (1984), 'A simple implicit measure of the effective bid-ask spread in an efficient market', *Journal of Finance* **39**(4), 1127–1140.

Shephard, N. & Xiu, D. (2013), Econometric analysis of multivariate realised QML: Estimation of the covariation of equity prices under asynchronous trading. Unpublished manuscript, University of Oxford.

Ubukata, M. & Oya, K. (2009), 'Estimation and testing for dependence in market microstructure noise', *Journal of Financial Econometrics* **7**, 106–151.

Varneskov, R. T. (2013), Estimating the quadratic variation spectrum of noisy asset prices using generalized flat-top realized kernels. Unpublished Manuscript, Aarhus University.

Varneskov, R. T. & Voev, V. (2013), 'The role of realized ex-post covariance measures and dynamic model choice on the quality of covariance forecasts', *Journal of Empirical Finance* **20**, 83–95.

Voev, V. & Lunde, A. (2007), 'Integrated covariance estimation using high-frequency data in the presence of noise', *Journal of Financial Econometrics* **5**, 68–104.

Yang, S. Y. (2007), 'Maximal moment inequality for partial sums of strong mixing sequences and application', *Acta Mathematica Sinica* **English Series 23**(6), 1013–1024.

Zhang, L. (2011), 'Estimating covariation: Epps effect and microstructure noise', *Journal of Econometrics* **160**, 33–47.

A Additional Theory

A.1 Matrix Concepts

The exhibition in the main part of the paper requires the definition of some matrix-valued concepts.

Definition A.1. *Let $\boldsymbol{X} \in \mathbb{R}^{d\times k}$ be a mixed Gaussian distributed random variable for ex-ante random variables $\boldsymbol{A} \in \mathbb{R}^{d\times k}$ and covariance matrix $\boldsymbol{B} \in \mathbb{R}^{dk\times dk}$, denoted by $\boldsymbol{X} \overset{d_s(X)}{\to} MN(\boldsymbol{A}, \boldsymbol{B})$. Then, $\boldsymbol{X}$ is characterized by the following properties: (1) $vec(\boldsymbol{X})$ follows a $dk \times 1$-variate mixed gaussian distribution. (2) $\mathbb{E}[vec(\boldsymbol{X})|\boldsymbol{X}] = vec(\boldsymbol{A})$. (3) For any four vectors $\boldsymbol{a}, \boldsymbol{c} \in \mathbb{R}^d$ and $\boldsymbol{b}, \boldsymbol{d} \in \mathbb{R}^k$*

$$\mathrm{Cov}(\boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}, \boldsymbol{c}'\boldsymbol{X}\boldsymbol{d}|\boldsymbol{X}) = \boldsymbol{v}_{ab}'\boldsymbol{B}\boldsymbol{v}_{cd}, \qquad \boldsymbol{v}_{ab} = vec([\boldsymbol{a}\boldsymbol{b}' + \boldsymbol{b}\boldsymbol{a}']/2) \in \mathbb{R}^{dk\times 1}.$$

Definition A.2. *The symmetrizer matrix of (Abadir & Magnus 2005, Section 11.2) is defined as $\boldsymbol{N}_d = \frac{1}{2}(\boldsymbol{I}_{d^2\times d^2} + \boldsymbol{K}_{d^2})$, $\boldsymbol{K}_{d^2} = \sum_{i=1}^d \sum_{j=1}^d \boldsymbol{x}_i\boldsymbol{x}_j' \otimes \boldsymbol{x}_j\boldsymbol{x}_i'$ where $\boldsymbol{x}_i$ is the i-th d-dimensional elementary vector.*

A.2 Variance of the Element-wise Flat-top Realized Kernel

First, for the exact expression of $\boldsymbol{\mathcal{B}}_{a,b}(\lambda, \kappa_{a,b})$ in Corollary 1, let $2\boldsymbol{\mathcal{Z}} = (\tilde{\boldsymbol{\mathcal{Z}}}_{ab,a'bl'})_{a,b,a',b'\in 1\ldots,d}$ for $\boldsymbol{\mathcal{Z}} = (\boldsymbol{\mathcal{Q}}, \boldsymbol{\mathcal{N}}, \boldsymbol{\mathcal{C}})$ where

$$\tilde{\boldsymbol{\mathcal{Q}}}_{ab,a'b'} = \int_0^1 \left(\boldsymbol{\Sigma}_t^{a,a'}\boldsymbol{\Sigma}_t^{b,b'} + \boldsymbol{\Sigma}_t^{a,b'}\boldsymbol{\Sigma}_t^{b,a'}\right)\chi_2(t)\chi_1^{-1}(t)dt,$$

whose exact form follows from definition A.1, and $\tilde{\boldsymbol{\mathcal{N}}}_{ab,a'b'}$ and $\tilde{\boldsymbol{\mathcal{C}}}_{ab,a'b'}$ are defined with a similar element-wise decomposition of the different Kronecker products $\boldsymbol{\Omega}^{(uu)} \otimes \boldsymbol{\Omega}^{(uu)}$, $\boldsymbol{\Omega}_t^{(ee)} \otimes \boldsymbol{\Omega}_t^{(ee)}$, $\boldsymbol{N}_d\left(\boldsymbol{\Omega}^{(uu)} \otimes \int_0^1 \boldsymbol{\Omega}_t^{(ee)}\chi_1^{-1}(t)dt\right)$, $\boldsymbol{N}_d\left(\boldsymbol{\Sigma}_t \otimes \left(\boldsymbol{\Omega}^{(uu)} + \boldsymbol{\Omega}_t^{(ee)}\right)\right)$ and $\boldsymbol{N}_d\left(\boldsymbol{\Omega}_t^{(ep)} \otimes \boldsymbol{\Omega}_t^{(ep)}\right)$ of $\boldsymbol{\mathcal{N}}$ and $\boldsymbol{\mathcal{C}}$, respectively. The exact form of $\boldsymbol{\mathcal{B}}_{a,b}(\lambda, \kappa_{a,b})$ is then

$$\boldsymbol{\mathcal{B}}_{a,b}(\lambda, \kappa_{a,b}) = 2\kappa_{a,b}\lambda^{(00)}\tilde{\boldsymbol{\mathcal{Q}}}_{ab,a'b'} + 2\kappa_{a,b}^{-3}\lambda^{(22)}\tilde{\boldsymbol{\mathcal{N}}}_{ab,a'b'} + 4\kappa_{a,b}^{-1}\lambda^{(11)}\tilde{\boldsymbol{\mathcal{C}}}_{ab,a'b'}.$$

A.3 The Exact Form of $\boldsymbol{\mathcal{B}}_{\rho_{ab}}(\lambda, \kappa)$ and $\boldsymbol{\mathcal{B}}_{\beta_{ab}}(\lambda, \kappa)$

Let $\boldsymbol{\mathcal{B}}_{[ab]}(\lambda, \kappa)$ be the asymptotic covariance matrix of the 3×1 vector $(RK_{aa}^{\epsilon}, RK_{ab}^{\epsilon}, RK_{aa}^{\epsilon})'$ and $\boldsymbol{\Sigma}_{ab} = \int_0^1 \boldsymbol{\Sigma}_t^{a,b}dt$, then $\boldsymbol{\mathcal{B}}_{\rho_{ab}}(\lambda, \kappa) = \boldsymbol{f}_{ab}'\boldsymbol{\mathcal{B}}_{[ab]}(\lambda, \kappa)\boldsymbol{f}_{ab}$ and $\boldsymbol{\mathcal{B}}_{\beta_{ab}}(\lambda, \kappa) = \boldsymbol{g}_{ab}'\boldsymbol{\mathcal{B}}_{[ab]}(\lambda, \kappa)\boldsymbol{g}_{ab}$ where $\boldsymbol{f}_{ab}$ and $\boldsymbol{g}_{ab}$ are 3×1 gradient vectors defined as

$$\boldsymbol{f}_{ab} = \begin{pmatrix} -\boldsymbol{\Sigma}_{aa}^{-3/2}\boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1/2}/2 \\ \boldsymbol{\Sigma}_{aa}^{-1/2}\boldsymbol{\Sigma}_{bb}^{-1/2} \\ -\boldsymbol{\Sigma}_{aa}^{-1/2}\boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-3/2}/2 \end{pmatrix}, \qquad \boldsymbol{g}_{ab} = \begin{pmatrix} -\boldsymbol{\Sigma}_{aa}^{-2}\boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{aa}^{-1} \\ 0 \end{pmatrix}.$$

B Proofs

In the following, K, k, and ϵ denote generic constants where $K, k \in (0, \infty)$ and $\epsilon \in (0, 1)$ unless specified otherwise, and they may take different values in different places. Furthermore, (stochastic) orders are sometimes referring to scalars, vectors, and matrices. All convergence results are for $n \to \infty$.

B.1 Proofs of Lemmas 1-2 and Theorem 1

Following along the lines of Barndorff-Nielsen et al. (2011), the proofs are provided for the univariate case since the Cramér-Wold Theorem, e.g. (Davidson 2002, Theorem 25.6), will, then, deliver the multivariate result. (Varneskov 2013, Sections B.1 and C), provides a detailed proof of three similar results for the univariate case with equally spaced observations. Hence, the proof of the three results in the present setting is completed by providing the remaining arguments in the form of auxiliary lemmas. Let the boldface notation be dropped for the $d = 1$ case, and write the decomposition of $RK^*(p)$,

$$RK^*(p) = RK^*(p^*) + RK^*(U) + RK^*(p^*, U) + RK^*(U, p^*). \tag{B.1}$$

Lemma B.1 provides two approximations that will be used throughout without explicit reference, Lemma B.2 provides bounds on end-effects, Lemma B.3 provides a result for convergence of moments with random durations, Lemma B.4 on the bias of $RK^*(p^*, U) + RK^*(U, p^*)$, and, finally, Lemma B.5 provides a central limit theorem for products of orthogonal variables.

Definition B.1. *$b(h/H) = H\Delta k(h/H)$ is the sample analog of $k^{(1)}(h/H)$.*

Definition B.2. *For $h \in \mathbb{Z}$, denote $S_h^+ = \max(h, 0)$, and $S_h^- \min(h, 0)$, $S^{(2,h)} = \{1 + S_h^+, \ldots, n - 1 + S_h^-\}$, and $S^{(1,h)} = S^{(2,h)} \setminus \{1\}$. Further, denote $\mathbb{Z}_k = \{-k, \ldots, -1, 0, 1, \ldots, k\}$ for $k \in \mathbb{N}$ and $\mathbb{Z}_{k+1}^K = \mathbb{Z}_K \setminus \mathbb{Z}_k$ for $K - k \in \mathbb{N}$.*

Lemma B.1 (Jacod (2009), 6.23 and Ikeda (2011), Lemma 6)**.** *Under Assumptions 1-2, 4, and 6, let $\Delta\Upsilon_{t_i} = \int_{t_{i-1}}^{t_i} \Upsilon_t dt$, then for $i \geq 2$,*

$$\mathbb{E}\left[(\Delta t_i)^{-1/2} \left|\Delta p_{t_i}^* - \Sigma_{t_{i-1}}^{1/2} \Delta W_{t_i}\right|^s |\mathcal{H}_{t_{i-1}}\right] \leq K_s \mathbb{E}\left[(\Delta t_i)^{\min(1, s/2)} |\mathcal{H}_{t_{i-1}}\right],$$
$$\mathbb{E}\left[(\Delta t_i)^{-1/2} \left|\Delta\Upsilon_{t_i} - \Upsilon_{t_{i-1}} \Delta t_i\right|^s |\mathcal{H}_{t_{i-1}}\right] \leq K_s \mathbb{E}\left[(\Delta t_i)^{\min(1, s/2)} |\mathcal{H}_{t_{i-1}}\right].$$

Lemma B.2 (Jittered End-points)**.** *Under Assumptions 1-6,*

(a) $\Delta p_{t_1}^* + \Delta p_{t_n}^* \leq o_p\left(m^{1+\delta/2} n^{-1/2}\right)$.

(b) $(\Delta p_{t_1}^*)^2 + (\Delta p_{t_n}^*)^2 + 2\sum_{h=1}^{n-1} k\left(\frac{h}{H}\right)\left(\Delta p_{t_{h+1}}^* \Delta p_{t_1}^* + \Delta p_{t_n}^* \Delta p_{t_{n-h}}^*\right) \leq o_p\left(m^{2+\delta} n^{-1}\right) + o_p\left(H^{1/2} m^{1+\delta/2} n^{1/2\delta - 1}\right)$.

(c) $U_{t_0}^2 + U_{t_n}^2 - 2\sum_{h=1}^{n-1}\left(k\left(\frac{h}{H}\right) - k\left(\frac{h-1}{H}\right)\right)\left(U_{t_0} U_{t_h} + U_{t_n} U_{t_{n-h}}\right) = O_p\left(m^{-1}\right)$.

(d) $2\sum_{h=1}^{n-2} k\left(\frac{h}{H}\right)\left(U_{t_n} \Delta p_{n-h}^* - U_{t_0} \Delta p_{h+1}^*\right) + \frac{2}{H}\sum_{h=1}^{n-1} b\left(\frac{h}{H}\right)\left(\Delta p_n^* U_{t_{n-h}} - \Delta p_1^* U_{t_h}\right) + k(0)(U_{t_n} r_n^* - U_{t_0} r_1^*) + k((n-1)/H)(U_{t_n} r_1^* - U_{t_0} r_n^*) \leq o_p\left((H/m)^{1/2} n^{(\delta-1)/2}\right) + o_p\left(m H^{-1/2} n^{(\delta-1)/2}\right) + o_p(m^{(1+\delta)/2} n^{-1/2})$.

Proof. **(a)** Recall the definition of the jittered end-point returns $\Delta p^*_{t_1} = p^*_{t_m} - m^{-1}\sum_{i=1}^m p_{t_{i-1}}$ and $\Delta p^*_{t_n} = m^{-1}\sum_{i=1}^m p^*_{t_{N-m+1}} - p^*_{t_{n-1}}$. The result is derived for $\Delta p^*_{t_1}$, since the symmetric result for $\Delta p^*_{t_n}$ follows immediately. Using the telescoping sum property of returns, write

$$\begin{aligned}\Delta p^*_{t_1} = \frac{1}{m}\sum_{i=1}^m (p^*_{t_m} - p^*_{t_{i-1}}) = \frac{1}{m}\sum_{i=1}^m\sum_{j=1}^i \Delta p^*_{t_{m+1-j}} = \frac{1}{m}\sum_{i=1}^m O_p\left(i\sqrt{D_{n,i}/n}\right) \\ \leq \max_{i=1,\dots,m} O_p\left(m\sqrt{D_{n,i}/n}\right) = o_p\left(m^{1+\delta/2}n^{-1/2}\right)\end{aligned}$$

which provides the first result. **(b)** The probabilistic order of the first two terms follows by **(a)**. The order of the third term follows by calculating the mean and variance of, given $h > 0$, two conditionally independent variables and noticing $(\Delta p^*_{t_{h+1}})^2 \leq \max_{i=1,\dots,n} O_p(D_{n,i}/n) = o_p\left(n^{\delta-1}\right)$ for $h > 0$. The boundary terms for $h = n-1$ are of order $O_p(H^{-1/2})$ smaller than the first two terms. **(c)** follows by (Barndorff-Nielsen et al. 2011, Propositions A.1-A.2), since Assumptions 4-5 ensures $\sum_{h\in\mathbb{Z}}(|\int_0^1 \Omega_t^{(ee)}(h)dt| + |\Omega^{(uu)}(h)|) < \infty$. **(d)** Using the bounds in **(a)** and **(c)** along with $\Delta p^*_{t_{h+1}} \leq o_p\left(n^{(\delta-1)/2}\right)$ for $h > 1$ to describe the probabilistic orders of $U_{t_n}\Delta p^*_{n-h}$ and $\Delta p^*_n U_{t_{n-h}}$, the result follows by the same derivations as for (Varneskov 2013, Lemma C.3 (d)). □

Remark B.1. *Lemma B.2 (b), (c) and (d) correspond to the contribution from the end-points of $RK^*(p^*)$, $RK^*(U)$ and $RK^*(p^*,U) + RK^*(U,p^*)$, respectively. Algebraic manipulation of the powers shows that $\delta \in (0, 1-\nu)$ guarantees the $O_p\left(m^{-1}\right)$ term to provide the lower bound on jittering, and in conjunction with $\xi > 1/4$ that the $o_p\left(m^{2+\delta}n^{-1}\right)$ term in (b) provides the upper bound.*

Lemma B.3 (Moments with Random Duration)**.** *Let $f(t)$ be $\mathcal{H}_1$-measurable, bounded, càdlàg, independent of Δt, and $\sum_{i=1}^n f(t_{i-1})\Delta t_i \xrightarrow{\mathbb{P}} \int_0^1 f(t)dt$. Under Assumption 6, then $D^\beta_{n,i} = 1$ for $\beta = 0$, trivially, and $n^{-\alpha}\sum_{i=1}^n \mathbb{E}[f(t_{i-1})(\Delta t_i)^\beta|\mathcal{H}_1] = n^{1-\alpha-\beta}\int_0^1 f(t)\frac{\chi_\beta(t)}{\chi_1(t)}dt(1+o_p(1))$.*

Proof. Define $g(t_{i-1}) = f(t_{i-1})\mathbb{E}[D^\beta_{n,i}|\mathcal{H}_1]/\mathbb{E}[D_{n,i}|\mathcal{H}_1]$, then

$$n^{-\alpha}\sum_{i=1}^n \mathbb{E}[f(t_{i-1})(\Delta t_i)^\beta|\mathcal{H}_1] = n^{(1-\alpha-\beta)}\sum_{i=1}^n g(t_{i-1})\Delta t_i - n^{(1-\alpha-\beta)}\sum_{i=1}^n g(t_{i-1})\left(\Delta t_i - \mathbb{E}[\Delta t_i|\mathcal{H}_1]\right).$$

For the second term, $\sum_{i=1}^n g(t_{i-1})\left(\Delta t_i - \mathbb{E}[\Delta t_i|\mathcal{H}_1]\right) \leq O_p(1)\sum_{i=1}^n\left(\Delta t_i - \mathbb{E}[\Delta t_i|\mathcal{H}_1]\right)$ since $f(t)$ is bounded. Let $x_i = \Delta t_i - \mathbb{E}[\Delta t_i|\mathcal{H}_1]$, then $\sum_{i=1}^n \mathbb{E}\left[x_i|\mathcal{H}_{t_{i-1}}\right] = 0$ by the law of iterated expectations, and $\sum_{i=1}^n \mathbb{E}\left[|x_i|^2|\mathcal{H}_{t_{i-1}}\right] = O_p(n^{-1})$ by Assumption 6 (3). Hence, by (Jacod 2009, Lemma 4.1), $\sum_{i=1}^n x_i = o_p(1)$. The final result, then, follows by Lebesque integration in conjunction with the continuous mapping theorem. □

Remark B.2. *Lemma B.3 is similar to (Ikeda 2011, Lemma 2), but differs subtly with respect to its treatment of filtrations $\mathcal{H}_1$, respectively, $\mathcal{H}_t$. The difference arises as the second moment of the endogenous noise component in Assumption 4 is not progressively measurable with respect to $\mathcal{H}_t$.*

Lemma B.4 (Bias of $RK^*(p^*,U) + RK^*(U,p^*)$)**.** *Let Assumptions 1-6 hold. Apart from the endpoint in Lemma B.2 (d), the remainder of $RK^*(p^*,U) + RK^*(U,p^*)$ may be written as $B(r^*,U) = B_1(r^*,U) + B_2(r^*,U)$ where*

$$B_1(r^*,U) = \frac{2}{H}\sum_{h\in\mathbb{Z}_{cH-1}} b\left(\frac{|h|}{H}\right)\sum_{i\in S^{(1,h)}} r_i^* U_{t_{i-h}}, \quad B_2(r^*,U) = \frac{2}{H}\sum_{h\in\mathbb{Z}_{cH}^{n-1}} b\left(\frac{|h|}{H}\right)\sum_{i\in S^{(1,h)}} r_i^* U_{t_{i-h}}.$$

Then, $\mathbb{E}[B(r^*,U)|\mathcal{H}_1] = O_p\left(n^{1/2}H^{-1}\alpha_e(cH)\right)$.

Proof. The representation follows straightforwardly. As $\mathbb{E}[B(r^*,U)|\mathcal{H}_1] = \mathbb{E}[B_2(r^*,U)|\mathcal{H}_1]$, it follows

$$\begin{aligned}
\mathbb{E}[B_2(r^*,U)|\mathcal{H}_1] &= \frac{2}{H}\sum_{h\in\mathbb{Z}_{cH}^{n-1}} b\left(\frac{|h|}{H}\right)\sum_{i\in S^{(1,h)}} \theta(t_{i-h},h)\Upsilon_{t_{i-1}}\Sigma_{t_{i-1}}^{1/2}\mathbb{E}\left[(\Delta t_i)^{1/2}|\mathcal{H}_1\right](1+o_p(1)) \\
&= \frac{2}{H}\sum_{h\in\mathbb{Z}_{cH}^{n-1}} b\left(\frac{|h|}{H}\right)\sum_{i\in S^{(1,h)}} \theta_{t_{i-h}}(h)\Upsilon_{t_{i-1}}\Sigma_{t_{i-1}}^{1/2}\mathbb{E}\left[(\Delta t_i)^{1/2}|\mathcal{H}_1\right](1+o_p(1)) \\
&+ \frac{2}{H}\sum_{h\in\mathbb{Z}_{cH}^{n-1}} b\left(\frac{|h|}{H}\right)\sum_{i\in S^{(1,h)}} \left(\theta(t_{i-h},h) - \theta_{t_{i-h}}(h)\right)\Upsilon_{t_{i-1}}\Sigma_{t_{i-1}}^{1/2}\mathbb{E}\left[(\Delta t_i)^{1/2}|\mathcal{H}_1\right](1+o_p(1)).
\end{aligned}$$

Then, as

$$\sum_{i\in S^{(1,h)}} \theta_{t_{i-h}}(h)\Upsilon_{t_{i-1}}\Sigma_{t_{i-1}}^{1/2}\mathbb{E}\left[(\Delta t_i)^{1/2}|\mathcal{H}_1\right] = n^{1/2}\int_0^1 \Omega_t^{(ep)}(h)\chi_{1/2}(t)\chi_1^{-1}(t)dt(1+o_p(1)),$$

$$\sum_{i\in S^{(1,h)}} \left(\theta(t_{i-h},h) - \theta_{t_{i-h}}(h)\right)\Upsilon_{t_{i-1}}\Sigma_{t_{i-1}}^{1/2}\mathbb{E}\left[(\Delta t_i)^{1/2}|\mathcal{H}_1\right] \leq n^{-1/2}\sup_{t\in[0,1]}\left|\Upsilon_t\Sigma_t^{1/2}\chi_{1/2}(t)\right| k \leq n^{-1/2}K,$$

using Lemma B.3 and Assumption 4 (4), the final bound is established using $\sup_{h\in\mathbb{Z}_{cH}^{n-1}} b(|h|/H) \leq k$, $\sup_{t\in[0,1]}\Omega_t^{(ep)}(h) \leq O(\alpha_e(h))$, and (Varneskov 2013, Lemma C.4). □

Lemma B.5 (CLT for Orthogonal Variables)**.** *Let Assumption 5 hold and suppose $\{x_t\}_{t\in[0,1]}$ is an $\mathcal{X}_1$-measurable, bounded random variable where $\mathcal{X}_t \subset \mathcal{F}_t$ is a σ-algebra on $(\mathcal{O},\mathcal{F},(\mathcal{F}_t),\mathbb{P})$ satisfying $\mathcal{X}_t \perp\!\!\!\perp \mathcal{G}_s$ $\forall(t,s)\in[0,1]^2$. Further, suppose $\exists(b_1,b_2)\in\mathbb{R}^2$ such that*

$$\sum_{h\in\mathbb{Z}_{n-1}} \frac{1}{n^{b_1}}\sum_{i\in S^{(1,h)}} x_{t_i}x_{t_{i-h}} \xrightarrow{\mathbb{P}} \sum_{h\in\mathbb{Z}}\int_0^1 c_t(h)dt = \Omega^{(xx)} \tag{B.2}$$

where $c_t(h)$ is $\mathcal{X}_1$-measurable $\forall h\in\mathbb{Z}$, $\mathbb{P}$-uniformly bounded $\forall(h,t)\in\mathbb{Z}\times[0,1]$ and $\Omega^{(xx)}\in(0,\infty)$ $\mathbb{P}$-almost surely. Define the realized kernel estimator

$$RK(f,x,u) = \frac{1}{n^{b_2}}\sum_{h\in\mathbb{Z}_{n-1}} f\left(\frac{h}{H}\right)\sum_{i\in S^{(1,h)}} x_{t_i}u_{t_{i-h}}$$

where $H \propto n^{\nu}$ and $f(x) : \mathbb{R} \to [-1,1]$ is a weight function, which is differentiable at all but a finite number of points and $f^{(jj)} = \int_{-\infty}^{\infty}[f^{(j)}(x)]^2 dx < \infty$ for $j = 0$ and $j = 1$ almost everywhere. Last, suppose $n^{2b_2-b_1}H^{-1}\mathbb{V}[RK(f,x,u)|\mathcal{H}_1] \xrightarrow{\mathbb{P}} \mathcal{V}(f,x,u)$ where $\mathcal{V}(f,x,u)$ is measurable with respect to $\mathcal{X}_1$ and $\mathcal{V}(f,x,u) \in (0,\infty)$ $\mathbb{P}$-almost surely. Then, if $2b_2 - b_1 \in (\nu + 1/(2+r_u), 1+\nu]$,

$$n^{(2b_2-b_1)/2}H^{-1/2}RK(f,x,u) \overset{d_s(\mathcal{X}_1)}{\to} MN\left(0, \mathcal{V}(f,x,u)\right).$$

Proof. First, $\mathbb{E}[RK(f,x,u)|\mathcal{X}_1] = 0$ is immediate. Next, rewrite $RK(f,x,u)$ as

$$RK(f,x,u) = \sum_{i \in S^{(1,0)}} u_{t_i}\bar{w}_{n,i}, \qquad \bar{w}_{n,i} = \frac{1}{n^{b_2}} \sum_{h \in S^{(1,0)}} f\left(\frac{h-i}{H}\right) x_{t_h}$$

and define the sequences $(\bar{K}_n, \bar{L}_n) \in \mathbb{R}^+ \times \mathbb{R}^+$ where $\bar{K}_n = O(n^{\bar{k}})$ and $\bar{L}_n = O(n^{\bar{l}})$ for $0 < \bar{k} < \bar{l} < 1$. The stable central limit theorem follows by the central limit theorem for weighted α-mixing processes from (Yang 2007, Theorem 3.1) $\mathcal{X}_1$-conditionally in conjunction with (Barndorff-Nielsen et al. 2008, Lemma 1), since: (1) $\bar{w}_n = \max_{i \in S^{(1,0)}} |\bar{w}_{n,i}| \leq O_p(\sum_{i \in S^{(1,0)}} \bar{w}_{n,i}^2) = O_p(\mathbb{V}[RK(f,x,u)|\mathcal{X}_1])$, (2) $n\bar{L}_n^{-1}\alpha_u(\bar{K}_n) = o(1)$, (3) $n\bar{K}_n\bar{L}_n^{-1}\bar{w}_n^2\mathbb{V}[RK(f,x,u)|\mathcal{X}_1]^{-1} = o_p(1)$, and (4) $\bar{L}_n \sum_{i \in S^{(1,0)}} \bar{w}_{n,i}^2 = o_p(1)$. (1) The last equality is immediate as $u_{t_i} \perp\!\!\!\perp \bar{w}_{n,j}$ $\forall i,j$ and Assumption 5 implies $\sum_{h \in \mathbb{Z}} |\Omega^{(uu)}(h)| < \infty$. The first inequality follows using $\sum_{i \in S^{(1,0)}} \bar{w}_{n,i}^2 = O_p(Hn^{-(2b_2-b_1)})$ and (B.2) to infer $\bar{w}_n = O_p(H^{1/2}n^{-(1+2b_2-b_1)/2})$. Then, $O_p(H^{1/2}n^{-(1+2b_2-b_1)/2}) \leq O_p(Hn^{-(2b_2-b_1)})$ when $2b_2 - b_1 \leq 1 + \nu$. (2) is satisfied by having $0 < (1-\bar{l})/(1+r_u+\epsilon) < \bar{k} < \bar{l} < 1$ for $r_u \in \mathbb{N}^+$. (3) Using the results in (1), $n\bar{K}_n\bar{L}_n^{-1}\bar{w}_n^2\mathbb{V}[RK(f,x,u)|\mathcal{X}_1]^{-1} = n\bar{K}_n\bar{L}_n^{-1}O_p(n^{-1}) \xrightarrow{\mathbb{P}} 0$. (4) As $\bar{L}_n \sum_{i \in S^{(1,0)}} \bar{w}_{n,i}^2 = O_p(n^{\bar{l}+\nu-(2b_2-b_1)})$, $2b_2 - b_1 > \bar{l} + \nu$ must hold. From (2), select $\bar{l}$ such that $\bar{l} > (1-\bar{l})/(1+r_u+\epsilon)$ or $\bar{l} > 1/(2+r_u+\epsilon)$. Choosing $\bar{l} = 1/(2+r_u)$ satisfies this condition, thus providing the lower bound on $2b_2 - b_1$. □

Remark B.3. *Lemma B.5 generalizes (Varneskov 2013, Lemma C.5) to allow for arbitrarily scaled convergence in* (B.2) *provided $2b_2 - b_1 \in (\nu + r_u/(2+r_u), 1+\nu]$. In Theorem 1, $\nu = 1/2$, implying that the scale coefficients for products of u_{t_i} and e_{t_j} in $RK^*(U)$ are $b_1 = 1$ and $b_2 = 2\nu = 1$, respectively, while they for products of u_{t_i} and Δp_{t_j} in $RK^*(p^*,U) + RK^*(U,p^*)$ are $b_1 = 1$ and $b_2 = 1/2 + \nu = 1$. The latter follows by $\Delta p_{t_i}^* = (1 + o_p(1))\sigma_{t_{i-1}}(\Delta t_i)^{1/2}\epsilon_{t_i}$, $\epsilon_{t_i} \sim i.i.d.N(0,1)$, and Lemma B.3.*

B.2 Proof of Theorem 2

Write $RK^{\epsilon}(\boldsymbol{p}) - RK^*(\boldsymbol{p}) = \boldsymbol{M}'(\hat{\boldsymbol{K}} - \boldsymbol{K})\boldsymbol{M} = \boldsymbol{M}'\boldsymbol{Z}\boldsymbol{M}$ where $\boldsymbol{Z}$ is a d-dimensional diagonal matrix with entries $z_q = \hat{k}_q - k_q = \max(0, \epsilon_n - k_q)$, $q = 1, \ldots, d$. Let σ_q be the q-th eigenvalue of $\int_0^1 \boldsymbol{\Sigma}_t dt$, then from Theorem 1, it follows $k_q = \sigma_q + n^{-1/4}\eta_q$ where $\eta_q \overset{d_s(\mathcal{H}_1)}{\sim} MN(0, \boldsymbol{\mathcal{B}}_{\sigma_q}(\kappa))$ and $\boldsymbol{M} = O_p(1)$. Thus, it suffices to prove that $z_q = o_p(n^{-1/4})$ $\forall q = 1, \ldots, d$. Let $\|\cdot\|_1 = \mathbb{E}[|\cdot|]$, then

$$\|z_q\|_1 \leq (\|\epsilon_n\|_1 + \|k_q\|_1)\|\mathbf{1}_{\{\epsilon_n - k_q > 0\}}\|_1$$

by the triangle inequality. As $\|\mathbf{1}_{\{\epsilon_n - k_q > 0\}}\|_1 = \mathbb{P}[\epsilon_n - k_q > 0]$, the Markov inequality may be use to write,

$$\mathbb{P}[\epsilon_n - k_q > 0] = \mathbb{P}[\epsilon_n - n^{-1/4}\eta_q > \sigma_q] \leq \frac{\epsilon_n - \mathbb{E}[n^{-1/4}\eta_q]}{\sigma_q} = O_p(\epsilon_n).$$

Hence, $\|z_q\|_1 \leq O_p(\epsilon_n)$, which implies $z_q \leq O_p(\epsilon_n)$. This provides the final result.

Research Papers 2007

2011-21: Bent Jesper Christensen, Olaf Posch and Michel van der Wel: Estimating Dynamic Equilibrium Models using Macro and Financial Data

2011-22: Antonis Papapantoleon, John Schoenmakers and David Skovmand: Efficient and accurate log-Lévi approximations to Lévi driven LIBOR models

2011-23: Torben G. Andersen, Dobrislav Dobrev and Ernst Schaumburg: A Functional Filtering and Neighborhood Truncation Approach to Integrated Quarticity Estimation

2011-24: Cristina Amado and Timo Teräsvirta: Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations

2011-25: Stephen T. Ziliak: Field Experiments in Economics: Comment on an article by Levitt and List

2011-26: Rasmus Tangsgaard Varneskov and Pierre Perron: Combining Long Memory and Level Shifts in Modeling and Forecasting of Persistent Time Series

2011-27: Anders Bredahl Kock and Timo Teräsvirta: Forecasting Macroeconomic Variables using Neural Network Models and Three Automated Model Selection Techniques

2011-28: Anders Bredahl Kock and Timo Teräsvirta: Forecasting performance of three automated modelling techniques during the economic crisis 2007-2009

2011-29: Yushu Li: Wavelet Based Outlier Correction for Power Controlled Turning Point Detection in Surveillance Systems

2011-30: Stefano Grassi and Tommaso Proietti: Stochastic trends and seasonality in economic time

2011-31: Rasmus Tangsgaard Varneskov: Generalized Flat-Top Realized Kernel Estimation of Ex-Post Variation of Asset

2011-32: Christian Bach: Conservatism in Corporate Valuation

2011-33: Adrian Pagan and Don Harding: Econometric Analysis and Prediction of Recurrent Events

2011-34: Lars Stentoft: American Option Pricing with Discrete and Continuous Time Models: An Empirical Comparison

2011-35: Rasmus Tangsgaard Varneskov: Flat-Top Realized Kernel Estimation of Quadratic Covariation with Non-Synchronous and Noisy Asset Prices