Malte S. Kurz, and Stefan Mittnik

Risk Assessment and Spurious Seasonality

Risk Assessment and Spurious Seasonality[*]

Malte S. Kurz[†1] and Stefan Mittnik[‡1]

[1]Department of Statistics,
Ludwig-Maximilians-Universität München,
Akademiestr. 1, 80799 Munich, Germany

April 25, 2018[§]

Abstract

To determine the appropriate level of risk capital financial institutions are required to empirically estimate and predict specific risk measures. Although regulation commonly prescribes the forecasting horizon and the frequency with which risk assessments have to be reported, the scheme with which the underlying data are sampled typically remains unspecified. We show that, given assessment frequency and forecasting horizon, the choice of the sampling scheme can greatly affect the results of risk assessment procedures. Specifically, sequences of variance estimates are prone to exhibit spurious seasonality when the assessment frequency is higher than the sampling frequency of non-overlapping return data. We derive the autocorrelation function of such sequences for a general class of weak white noise processes and for a general class of variance estimators. To overcome the problem of spurious seasonality, we present a boundary-corrected exponentially-weighted moving-average version of the two-scales variance estimator introduced in the realized-volatility literature.

---

1   Introduction

Reliable estimation and prediction of the volatility of financial instruments is key to sound financial risk management. In practice, the return interval, the forecasting horizon and the assessment frequency are specified by regulation or management policies. Typically, however, the sampling frequency of the data underlying the empirical analysis remains unspecified. If the sampling frequency of the return data is more granular than the horizon for risk assessment, three strategies for estimating and forecasting risk measures are commonly adopted: (a) derive a risk estimate that matches the return interval specified (e.g., one-day volatility from daily return data) and then either use (square-root) scaling or derive model-based multi-step forecasts to obtain estimates for longer (e.g., monthly, annual) horizons; (b) temporally aggregate the underlying data so that they match the horizon for risk assessment, leadind to analyses with overlapping samples; or (c) temporally aggregate the data so that samples do not overlap.[1]

In this paper, we address consequences of assessing risk for horizons that exceed the assessment frequency of the risk estimates. This is, for example, the case when asset managers rebalance weekly or monthly but assess and report risk at a daily frequency. Similar situations arise in banking (Basel III) and insurance (Solvency II) regulation. According to the Basel Committee on Banking Supervision (BCBS 2016), in Basel III, banks have to estimate the ten-day-ahead expected shortfall (ES) on a daily basis; and BCBS (2016, §181 c) states that the ten-day ES estimates need to be derived without scaling from shorter horizons and allows using overlapping return data. Various studies have investigated possible consequences of using overlapping returns for risk estimation or, more general, for statistical inference.[2] What has not been studied are the consequences of the implicit overlap that arises when assessing risk measures at a higher frequency than the sampling frequency of the data.

In the following, we restrict our analysis to the return *variance*, since other risk measures, such as volatility, value-at-risk or expected shortfall, are directly or indirectly related to variance. Moreover, for the sake of simplicity, we assume that returns are recorded at a daily frequency – implying that the most granular sampling and assessment frequency is daily.[3] To illustrate the estimation strategies (a)–(c) outlined above, Figure 1 depicts possible specifications for return interval and data sampling schemes in h-day-ahead assessments. The two rows in each panel indicate the return data used for estimation on day t and $t + 1$, respectively. Panel (a) reflects the sampling scheme for risk estimation based on daily return data. In this case, to derive h-

---

[1]See Andersen and Bollerslev (1998) and Andersen et al. (1999) for a detailed analysis and discussion of the tradeoff between sampling frequency and forecast horizon. More recently Kole et al. (2017) studied the impact of temporal and portfolio aggregation on the quality of ten-day ahead VaR forecasts.

[2]See, for example, Bod et al. (2002), Daníelsson et al. (2016), Daníelsson and Zhou (2016), Hansen and Hodrick (1980), Hedegaard and Hodrick (2016), Kluitman and Franses (2002) and Mittnik (2011).

[3]Note, however, our results also apply to intraday analyses.
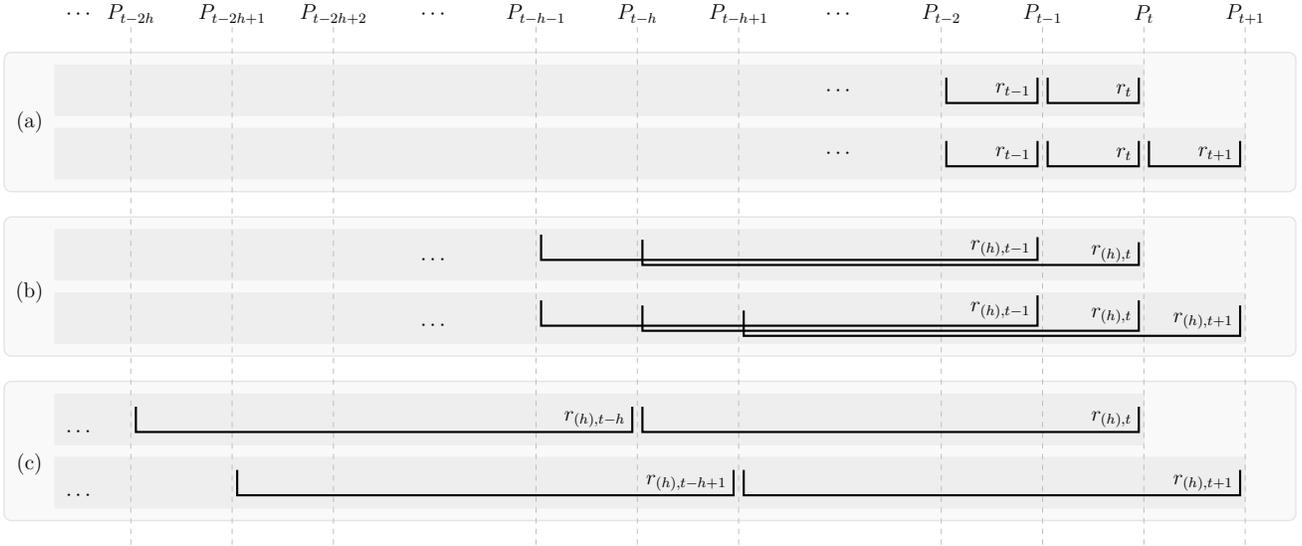
Figure 1: Illustration of different combinations of return intervals and sampling schemes for deriving h-day-ahead risk measures. Each panel consists of two rows: The first row sketches the data used for estimation at time t and the second row those at $t+1$. Panel (a) shows a scheme with daily sampling of daily returns. Here, risk estimates have to be scaled up to derive h-day-ahead risk estimates. Panel (b) illustrates the sampling scheme when using overlapping h-day returns. Panel (c) indicates the scheme when using non-overlapping h-day returns.

day-horizon estimates, one needs to either rely on a scaling rule that approximates h-day risk from one-day estimates or on some multi-step forecasting procedure. Panel (b) illustrate the sampling when estimating with overlapping h-day returns at times t and $t + 1$. Finally, Panel (c) shows the sampling scheme for returns when estimates are based on non-overlapping return intervals, revealing the implicit overlap when the assessment frequency is higher than the data sampling frequency. It is the latter scheme that is the main focus of this study.

In a recent study, Daníelsson and Zhou (2016) consider sampling strategies (a)–(c) for obtaining h-day-ahead risk estimates. They concentrate on a comparison of strategy (a) (square-root-of-time scaling) with strategy (b) (overlapping returns) with regards to bias and variance of risk estimates. Our work differs in two regards: First, we focus directly on strategies based on longer, namely, h-day return intervals. Given that BCBS (2016) explicitly rules out any risk assessment based on scaling, but also to avoid excessive clutter, we do not consider scaling strategies.[4] Second, we are not only interested in the accuracy of risk estimates (i.e., bias, variance, mean squared error etc.) at a given period, but also in the dynamic properties of risk estimates.

If data availability is not an issue, estimates based on non-overlapping returns are, from an econometric point of view, the preferred choice (Daníelsson and Zhou 2016). We need to be concerned, however, when assessing variances at a higher frequency (e.g., daily) from return

---

[4]For a discussion of the square-root-of-time scaling see Christoffersen et al. (1998), Daníelsson and Zigrand (2006), Diebold et al. (1997) and McNeil and Frey (2000). Scaling rules under other than multivariate normal processes and, especially, serially dependent observations are derived in Embrechts et al. (2005).

data that have longer return interval (e.g., weekly or monthly returns). We show that standard variance estimators, such as the moving-window sample variance or the exponentially-weighted moving average (EWMA) variance estimator (Riskmetrics 1996), tend to exhibit strong but spurious saw-tooth patterns. Clearly, risk managers who are obliged to assess risk more often (e.g., daily) than the horizon for risk-assessment implies (e.g., ten days in the Basel III or 259 days in the Solvency II framework) need to be aware of the fact that strong seasonal patterns may be induced. We demonstrate this phenomenon both empirically for real data and theoretically for well-behaved data-generating processes (DGPs), such as Gaussian white noise or GARCH(p,q) processes. We derive the theoretical autocorrelation function (ACF) for sequences of successive variance estimates for a broad class of DGPs and variance estimators. Moreover, we present variance estimators, based on overlapping h-day return intervals, that overcome the problem of spurious seasonality. Specifically, we introduce a boundary-corrected exponentially-weighted moving-average (EWMA) version of the two-scales estimator developed by Zhang et al. (2005). Our estimator does not suffer from spurious seasonality and performs best when compared to a range of alternative estimators.

The paper is organized as follows. In Section 2, using real data, we empirically illustrate and explain the presence of spurious seasonality in sequentially estimated variances. Section 3 defines the DGPs considered in this study, summarizes relevant results pertaining to stochastic processes and temporal aggregation, and derives quadratic-form representations for variance estimators. The theoretical ACF for sequences of daily estimated variances is derived in Section 4. Moreover, the phenomenon of spurious seasonality is illustrated and explained on theoretical grounds. Alternative variance estimators based on overlapping return intervals, but not suffering from spurious seasonality, are discussed in Section 5. Section 6 compares all variance estimators considered with respect to bias, variance, mean squared error (MSE) as well as their responsiveness to shocks in the data. Finally, Section 7 summarizes and concludes.

2 Spurious Seasonality in Variance Estimates from Temporally Aggregated, Non-Overlapping Returns

We are especially interested in the dynamic properties of sequential variance estimates. To illustrate our concern, we consider bi-weekly returns (i.e., returns over ten trading days) of the Dow Jones Industrial Average and look at two ways of displaying sequential variance estimates. First, we compute ten different bi-weekly return series, one for each of the ten trading days in the two-week window. For each of the ten return series we derive series of bi-weekly variance

estimates, using an EWMA variance estimator (Riskmetrics [1996])

$$\sigma^2_{(h),t,\lambda} = \frac{h}{\text{tr}(\boldsymbol{\mathcal{Q}}_{(h),\Delta,\lambda})} \frac{1-\lambda}{1-\lambda^\Delta} \sum_{\delta=0}^{\Delta-1} \lambda^\delta \big(r_{(h),t-h\delta} - \mu_{(h),t,\lambda}\big)^2, \tag{1}$$

where $r_{(h),t}$ is the h-day return at time t and $\mu_{(h),t,\lambda} = \frac{1-\lambda}{1-\lambda^\Delta} \sum_{\delta=0}^{\Delta-1} \lambda^\delta r_{(h),t-h\delta}$ is the EWMA estimator for the first moment.[5] We set $\lambda = 0.96$. The left graph in Figure 2 shows the ten different series of variance estimates, $(\sigma^2_{(10),10t+\tau,\lambda})_{t\in\mathbb{Z}}$, for $1 \leq \tau \leq 10$, each corresponding to a specific starting day. The right graph in Figure 2 is obtained by combining the ten bi-weekly variance estimates to a single, daily sequence. In other words, we appropriately connect the bi-weekly estimates, $(\sigma^2_{(10),t,\lambda})_{t\in\mathbb{Z}}$, obtained at a daily frequency and shown in the plot on the left. This means that the distance between two adjacent points of the sequence of variance estimates is always one day rather than ten days, as is the case with the plots on the left.
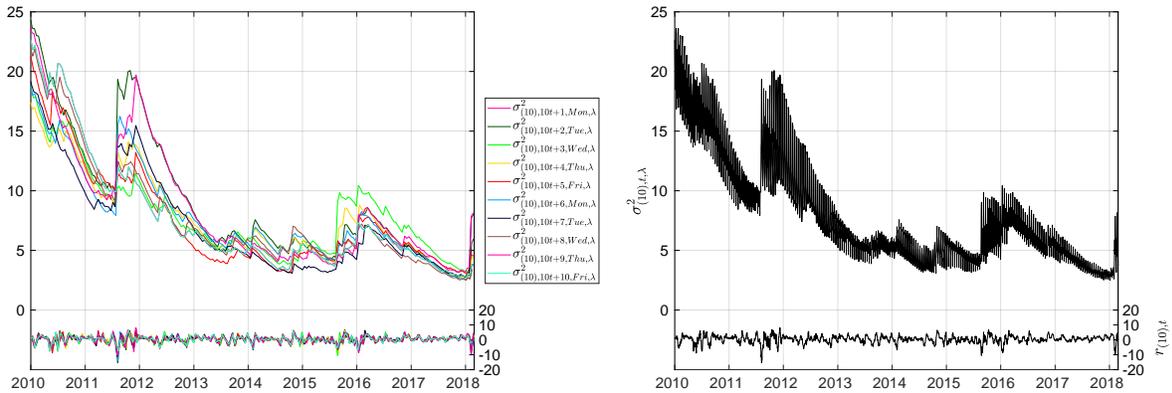
Figure 2: Estimated EWMA variances of the Dow Jones Industrial Average (DJIA) based on ten-day log-returns with a window length of 100 bi-weekly returns and an EWMA parameter of $\lambda = 0.96$. The first (last) estimates in both graphs are for 01-Jan-2010 (for 28-Feb-2018). The graph on the left shows at the top ten series of bi-weekly variance estimates, each corresponding to a specific weekday and start date, and the one on the right the daily series of bi-weekly variance estimates. The corresponding ten-day log-returns are plotted at the bottom of both graphs.

Each variance estimate shown in Figure 2 is based on non-overlapping ten-day returns. However, the assessment frequency of the estimates is higher than the sampling frequency of the underlying data set. As a consequence, there is a substantial overlap in data used for successive estimates.

In the following, we study the pronounced sawtooth pattern of the series of daily variance estimates shown on the right of Figure 2. Furthermore, we explain the reason for the slowly changing patterns in the ten variance series plotted on the left of Figure 2. To characterize the

---

[5] The multiplicative constant $\frac{h}{\text{tr}(\boldsymbol{\mathcal{Q}}_{(h),\Delta,\lambda})} = \big(1 - \frac{(1-\lambda)^2(1-\lambda^{2\Delta})}{(1-\lambda^\Delta)^2(1-\lambda^2)}\big)^{-1}$ is the bias-correction factor, see (3).

properties of estimated variance sequences we examine their autocorrelation function (ACF). The sample ACF of the daily series of estimates, $(\sigma^2_{(10),t,\lambda})_{t\in\mathbb{Z}}$, based on bi-weekly data, shown in Figure 3, displays a systematic periodic pattern, a feature we refer to as *spurious seasonality*. As will be shown below, this seasonal pattern is due to the sampling scheme for the data used
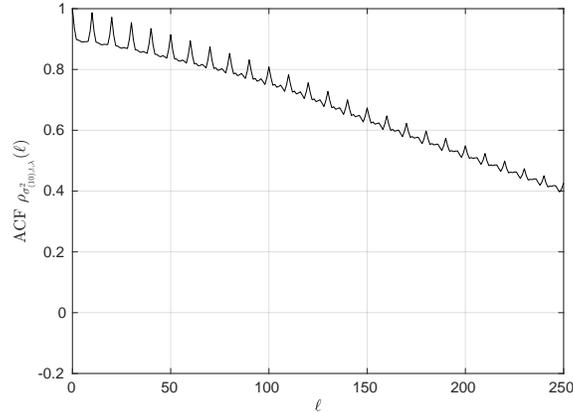
Figure 3: Sample autocorrelation function (ACF) for the daily series of bi-weekly EWMA variance estimates based on non-overlapping ten-day log-returns. The plot shows the sample ACF for the series of EWMA variance estimates for the Dow Jones Industrial Average (DJIA) from 01-Jan-2010 to 28-Feb-2018.

for variance estimation.

3 Some Prerequisites and Notation

In this section we introduce the two stochastic processes used in the analysis below, establish necessary notation, and briefly summarize relevant results on the temporal aggregation of stochastic processes. Finally, we introduce the (conditional) variance estimators that are the focus of this study.

3.1 Data Generating Stochastic Processes

We consider two data generating processes, the Gaussian white noise process and the GARCH(p,q) process. Both processes are so-called weak white noise processes.

Definition 1. *A stochastic process, $(x_t)_{t\in\mathbb{Z}}$, is called weak white noise process, if $\forall t, t_1, t_2 \in \mathbb{Z}$, $t_1 \neq t_2$:*

(i) $\mathbb{E}(x_t) = \mu$, with $|\mu| < \infty$,

(ii) $\mathrm{Var}(x_t) = \sigma^2$, with $0 < \sigma^2 < \infty$,

(iii) $\mathrm{Cov}(x_{t_1}, x_{t_2}) = 0$.

7

The Gaussian white noise process (Example 1) is the special case of independent and identically distributed (i.i.d.) random variables with normal distribution.

Example 1. *A stochastic process, $(x_t)_{t\in\mathbb{Z}}$, is called Gaussian white noise process, if $(x_t)_{t\in\mathbb{Z}}$ is a white noise process and $x_t \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$.*

As a second case we consider the generalized autoregressive conditional heteroskedasticity (GARCH) process (Example 2) introduced by Engle (1982) and Bollerslev (1986), a model class that is widely used in academic research and in practice in order to model the volatility of financial returns.

Example 2. *Let $(\epsilon_t)_{t\in\mathbb{Z}}$ be a sequence of i.i.d. random variables and let $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$. Further, let $\alpha_0 > 0$, $\alpha_1, \ldots, \alpha_q \geq 0$ and $\beta_1, \ldots, \beta_p \geq 0$ and assume $\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i < 1$, such that the process is weakly stationary.[6] Then, a GARCH(p,q) process, $(x_t)_{t\in\mathbb{Z}}$, with strictly positive volatility process, $(\sigma_t)_{t\in\mathbb{Z}}$, is defined by*

$$x_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i x_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2.$$

3.2 Temporal Aggregation of Returns and Stochastic Processes

Let $(P_t)_{t\in\mathbb{Z}}$ denote the process of daily prices of an asset, $(r_t)_{t\in\mathbb{Z}}$ with $r_t = \ln(P_t) - \ln(P_{t-1})$ the process of daily log-returns, and let vector $\boldsymbol{r}_{t,\delta} := [r_{t-\delta+1}, r_{t-\delta+2}, \ldots, r_{t-1}, r_t]'$ collect the δ daily returns from day $t - \delta + 1$ up to and including day t. h-day returns, $h > 1$, are then given by

$$r_{(h),t} = \ln(P_t) - \ln(P_{t-h}) = \sum_{j=0}^{h-1} r_{t-j} = \boldsymbol{1}_h' \boldsymbol{r}_{t,h},$$

where $\boldsymbol{1}_h$ is an $h \times 1$ column vector of ones. We call h the aggregation horizon.

In the following, we will always assume that the process of daily log-returns, $(r_t)_{t\in\mathbb{Z}}$, is generated by a weak white noise process (Definition 1) and, in some instances, consider the Gaussian white noise (Example 1) and the GARCH(p,q) process (Example 2) as special cases.

If we assume that the daily log-return series, $(r_t)_{t\in\mathbb{Z}}$, is a Gaussian white noise process with $\mathbb{E}(r_t) = \mu = 0$ and variance $\mathbb{E}(r_t^2) = \sigma^2 < \infty$, the temporally aggregated series, $(r_{(h),th})_{t\in\mathbb{Z}}$, where the sampling frequency coincides with the aggregation horizon, is again a Gaussian white noise process but with variance $\mathbb{E}(r_{(h),th}^2) = h\sigma^2 < \infty$. The situation changes, however, when the sampling frequency is lower than the aggregation horizon. This would, for example, be the case

---

[6]These restrictions on the parameter space guarantee a positive conditional variance σ_t^2 in the case of normally distributed innovations (Bollerslev 1986). Weaker necessary and sufficient conditions for a positive conditional variance are given in Nelson and Cao (1992).

when sampling h-day returns, $h > 1$, on a, say, daily basis. Then, $(r_{(h),t})_{t \in \mathbb{Z}}$ turns out to be a non-invertible moving average process of order $h - 1$ (in short: MA($h-1$) process) (Hansen and Hodrick 1980), with parameters $\theta_j = 1$ for $1 \leq j \leq h-1$, i.e., $r_{(h),t} = \sum_{j=0}^{h-1} r_{t-j} = \sum_{j=1}^{h-1} \theta_j r_{t-j} + r_t$, where $(r_t)_{t \in \mathbb{Z}}$ is the weak white noise series of daily log-returns. The autocorrelation function (ACF) $\rho_{r_{(h),t}}(\ell)$ for the process $(r_{(h),t})_{t \in \mathbb{Z}}$ is given by (cf. Mittnik (1988))

$$\rho_{r_{(h),t}}(\ell) = \mathrm{Corr}(r_{(h),t}, r_{(h),t-\ell}) = \begin{cases} \frac{\sum_{j=0}^{h-1-|\ell|} \theta_j \theta_{j+|\ell|}}{\sum_{j=0}^{h-1} \theta_j^2} = \frac{h-|\ell|}{h} & , |\ell| \leq h-1, \\ 0 & , |\ell| \geq h, \end{cases}$$

where we set $\theta_0 = 1$ for notational simplicity. Similar results can also be obtained under some regularity conditions for the GARCH(p,q) process.[7]

3.3 Estimating Variances

Analogous to the vector of daily returns, let $\boldsymbol{r}_{(h),t,\Delta} = \left[r_{(h),t-h(\Delta-1)}, r_{(h),t-h(\Delta-2)}, \ldots, r_{(h),t-h}, r_{(h),t} \right]'$ be the Δ-period vector of non-overlapping h-day returns up to and including time t. Denoting the $\Delta \times \Delta$ identity matrix by $\boldsymbol{I}_\Delta$, we define the $h\Delta \times \Delta$ matrix $\boldsymbol{H} = \boldsymbol{I}_\Delta \otimes \boldsymbol{1}_h$, where $\otimes$ is the Kronecker product, so that $\boldsymbol{r}_{(h),t,\Delta} = \boldsymbol{H}' \boldsymbol{r}_{t,h\Delta}$.

The most common estimator for the dispersion of returns is the sample variance. Defining the idempotent matrix $\boldsymbol{D} \in \mathbb{R}^{\Delta \times \Delta}$, $\boldsymbol{D} = \boldsymbol{I}_\Delta - \frac{1}{\Delta} \boldsymbol{1}_\Delta \boldsymbol{1}_\Delta'$, the moving-window sample variance for non-overlapping h-day returns is given by

$$\sigma_{(h),t}^2 = \frac{1}{\Delta-1} \sum_{\delta=0}^{\Delta-1} (r_{(h),t-h\delta} - \mu_{(h),t})^2 = \frac{1}{\Delta-1} \boldsymbol{r}_{(h),t,\Delta}' \boldsymbol{D}' \boldsymbol{D} \boldsymbol{r}_{(h),t,\Delta} = \frac{1}{\Delta-1} \boldsymbol{r}_{t,h\Delta}' \boldsymbol{H} \boldsymbol{D} \boldsymbol{H}' \boldsymbol{r}_{t,h\Delta},$$

with $\mu_{(h),t} = \frac{1}{\Delta} \boldsymbol{1}_\Delta' \boldsymbol{r}_{(h),t,\Delta} = \frac{1}{\Delta} \boldsymbol{1}_\Delta' \boldsymbol{H}' \boldsymbol{r}_{t,h\Delta}$ being the sample mean.

Below, we only discuss moving-window-type estimators. We restrict ourselves to this kind of estimators because in practice estimation is always based on a finite amount of data, so that finite-sample properties are of a relevance. The generalization of the results to the increasing window case is straightforward.[8]

Many variance estimators can be written as quadratic forms of the daily return vector, $\boldsymbol{r}_{t,h\Delta}$, i.e., $\sigma_{(h),t}^2 = \boldsymbol{r}_{t,h\Delta}' \boldsymbol{Q} \boldsymbol{r}_{t,h\Delta}$, where $\boldsymbol{Q} \in \mathbb{R}^{h\Delta \times h\Delta}$ is a positive definite, symmetric matrix. Examples are the sample variance given above, but also the exponentially-weighted moving

---

[7]Temporal aggregation of GARCH processes has been investigated by Drost and Nijman (1993), and a survey of studies on temporal aggregation of various types of univariate and multivariate time series processes is provided by Silvestrini and Veredas (2008).

[8]Asymptotic properties of sample variances when data are generated by a GARCH process are derived in Horváth et al. (2006).

average (EWMA) variance estimator (Riskmetrics 1996).

If we assume a weak white noise process (Definition 1) for $(r_t)_{t \in \mathbb{Z}}$ with $\text{Var}(r_t) = \sigma^2$, we have $\mathbb{E}(\boldsymbol{r}'_{t,h\Delta} \boldsymbol{\mathcal{Q}} \boldsymbol{r}_{t,h\Delta}) = \sigma^2 \text{tr}(\boldsymbol{\mathcal{Q}})$, and the bias of the variance estimator, $\boldsymbol{r}'_{t,h\Delta} \boldsymbol{\mathcal{Q}} \boldsymbol{r}_{t,h\Delta}$, is

$$\text{Bias}(\boldsymbol{r}'_{t,h\Delta} \boldsymbol{\mathcal{Q}} \boldsymbol{r}_{t,h\Delta}) = \mathbb{E}(\boldsymbol{r}'_{t,h\Delta} \boldsymbol{\mathcal{Q}} \boldsymbol{r}_{t,h\Delta}) - \text{Var}(r_{(h),t}) = \sigma^2(\text{tr}(\boldsymbol{\mathcal{Q}}) - h). \tag{2}$$

Therefore, variance estimates of the form $\boldsymbol{r}'_{t,h\Delta} \boldsymbol{\mathcal{Q}} \boldsymbol{r}_{t,h\Delta}$, can be bias-corrected by multiplying with factor $\frac{h}{\text{tr}(\boldsymbol{\mathcal{Q}})}$, i.e., by using

$$\sigma^2_{(h),t} = \boldsymbol{r}'_{t,h\Delta} \boldsymbol{Q} \boldsymbol{r}_{t,h\Delta}, \tag{3}$$

as variance estimate with $\boldsymbol{Q} = \frac{h}{\text{tr}(\boldsymbol{\mathcal{Q}})} \boldsymbol{\mathcal{Q}}$. Throughout the paper, we will use the bias-corrected versions of the variance estimators but will, in general, only define $\boldsymbol{\mathcal{Q}}$. Quantities $\boldsymbol{Q}$ and $\boldsymbol{\mathcal{Q}}$ are always related by $\boldsymbol{Q} = \frac{h}{\text{tr}(\boldsymbol{\mathcal{Q}})} \boldsymbol{\mathcal{Q}}$.

Specifically, the sample variance for non-overlapping h-day returns is given by

$$\sigma^2_{(h),t} = \boldsymbol{r}'_{t,h\Delta} \boldsymbol{Q}_{(h),\Delta} \boldsymbol{r}_{t,h\Delta}, \tag{4}$$

with $\boldsymbol{\mathcal{Q}}_{(h),\Delta} = \frac{1}{\Delta} \boldsymbol{H} \boldsymbol{D} \boldsymbol{H}' = \frac{1}{\Delta}(\boldsymbol{I}_\Delta \otimes \mathbf{1}_h)(\boldsymbol{I}_\Delta - \frac{1}{\Delta} \mathbf{1}_\Delta \mathbf{1}'_\Delta)(\boldsymbol{I}_\Delta \otimes \mathbf{1}'_h)$, and the EWMA variance for non-overlapping h-day returns (1) by

$$\sigma^2_{(h),t,\lambda} = \frac{h}{\text{tr}(\boldsymbol{\mathcal{Q}}_{(h),\Delta,\lambda})} \frac{1-\lambda}{1-\lambda^\Delta} \sum_{\delta=0}^{\Delta-1} \lambda^\delta (r_{(h),t-h\delta} - \mu_{(h),t,\lambda})^2 = \boldsymbol{r}'_{t,h\Delta} \boldsymbol{Q}_{(h),\Delta,\lambda} \boldsymbol{r}_{t,h\Delta}, \tag{5}$$

with $\boldsymbol{\mathcal{Q}}_{(h),\Delta,\lambda} = \boldsymbol{H} \boldsymbol{E}' \boldsymbol{\Lambda} \boldsymbol{E} \boldsymbol{H}' = (\boldsymbol{I}_\Delta \otimes \mathbf{1}_h)(\boldsymbol{I}_\Delta - \boldsymbol{w} \mathbf{1}'_\Delta) \boldsymbol{\Lambda} (\boldsymbol{I}_\Delta - \mathbf{1}_\Delta \boldsymbol{w}')(\boldsymbol{I}_\Delta \otimes \mathbf{1}'_h)$ and $\lambda \in (0,1)$. Vector $\boldsymbol{w} \in \mathbb{R}^{\Delta \times 1}$ and matrices $\boldsymbol{\Lambda}, \boldsymbol{E} \in \mathbb{R}^{\Delta \times \Delta}$ are defined by $\boldsymbol{w} = \frac{1-\lambda}{1-\lambda^\Delta} \cdot [\lambda^{\Delta-1}, \lambda^{\Delta-2}, \ldots, \lambda^1, 1,]'$, $\boldsymbol{\Lambda} = \text{Diag}(\boldsymbol{w}) = (\boldsymbol{w} \mathbf{1}'_\Delta) \odot \boldsymbol{I}_\Delta$ and $\boldsymbol{E} = \boldsymbol{I}_\Delta - \mathbf{1}_\Delta \boldsymbol{w}'$, respectively, with $\odot$ denoting the Hadamard product.

4 Autocorrelation of Estimated Variances

4.1 Theoretical Derivation

Let matrices $\boldsymbol{K}, \boldsymbol{L} \in \mathbb{R}^{h\Delta+\ell \times h\Delta}$ be defined by $\boldsymbol{K} = [\mathbf{0}_{(h\Delta \times \ell)}, \boldsymbol{I}_{h\Delta}]'$ and $\boldsymbol{L} = [\boldsymbol{I}_{h\Delta}, \mathbf{0}_{(h\Delta \times \ell)}]'$, $\ell \geq 0$, where $\mathbf{0}_{(h\Delta \times \ell)}$ denotes an $h\Delta \times \ell$ matrix of zeros, so that $\boldsymbol{r}_{t,h\Delta} = \boldsymbol{K}' \boldsymbol{r}_{t,h\Delta+\ell}$ and $\boldsymbol{r}_{t-\ell,h\Delta} = \boldsymbol{L}' \boldsymbol{r}_{t,h\Delta+\ell}$. Variance estimators are then given by the quadratic-form

$$\sigma^2_{(h),t} = \boldsymbol{r}'_{t,h\Delta} \boldsymbol{Q} \boldsymbol{r}_{t,h\Delta} = \boldsymbol{r}'_{t,h\Delta+\ell} \boldsymbol{K} \boldsymbol{Q} \boldsymbol{K}' \boldsymbol{r}_{t,h\Delta+\ell}. \tag{6}$$

We obtain the sample variance specified in (4) for $\boldsymbol{Q} = \boldsymbol{Q}_{(h),\Delta}$ and the EWMA variance specified in (5) for $\boldsymbol{Q} = \boldsymbol{Q}_{(h),\Delta,\lambda}$. Similarly, the ($\ell$ days) lagged variance estimator is given by

$$\sigma^2_{(h),t-\ell} = \boldsymbol{r}'_{t-\ell,h\Delta}\boldsymbol{Q}\boldsymbol{r}_{t-\ell,h\Delta} = \boldsymbol{r}'_{t,h\Delta+\ell}\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}'\boldsymbol{r}_{t,h\Delta+\ell}. \tag{7}$$

Expressions (6) and (7) allow us to write the variance estimator, $\sigma^2_{(h),t}$, and its lagged version, $\sigma^2_{(h),t-\ell}$, as quadratic forms of the very same vector of daily returns, $\boldsymbol{r}_{t,h\Delta+\ell}$. The quadratic forms $\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}'$ and $\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}'$ turn out to be block-diagonal matrices, with $\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}' = \text{blkDiag}(\boldsymbol{0}_{(\ell\times\ell)}, \boldsymbol{Q})$ and $\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}' = \text{blkDiag}(\boldsymbol{Q}, \boldsymbol{0}_{(\ell\times\ell)})$.

Next, to further analyze the properties of estimated variances based on non-overlapping h-day returns assessed at a frequency higher than the aggregation horizon, we derive the ACF of the series of estimated variances, $(\sigma^2_{(h),t})_{t\in\mathbb{Z}}$. Theorem 1 states a well-known result about the covariance of two quadratic forms of the same multivariate normally distributed random vector. It follows directly from results in Magnus (1978) on moments of products of quadratic forms for multivariate normally distributed random variables.

Theorem 1. *Let* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ *be symmetric matrices and* $\boldsymbol{X}$ *be multivariate normally distributed* $n \times 1$ *vector with* $\boldsymbol{\mu} = \mathbb{E}(\boldsymbol{X})$ *and* $\boldsymbol{\Sigma} = \mathbb{E}((\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})') = \mathbb{E}(\boldsymbol{X}\boldsymbol{X}') - \boldsymbol{\mu}\boldsymbol{\mu}'$. *For the quadratic forms* $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ *and* $\boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}$ *we have*

$$\text{Cov}(\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}, \boldsymbol{X}'\boldsymbol{B}\boldsymbol{X}) = 2\text{tr}(\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\mu}.$$

Proof. This follows directly from Lemma 6.2 in Magnus (1978). $\qquad\square$

The following corollary to Theorem 1 establishes the autocovariance function of the variances given by the quadratic forms (3) when the daily log-returns, $(r_t)_{t\in\mathbb{Z}}$, follow a Gaussian white noise process.[9]

Corollary 1. *Let* $(r_t)_{t\in\mathbb{Z}}$ *be a Gaussian white noise process (Example 1) with* $\mathbb{E}(r_t) = 0$ *and variance* $Var(r_t) = \sigma^2$ *and consider variance estimates of the form* $\sigma^2_{(h),t} = \boldsymbol{r}'_{t,h\Delta}\boldsymbol{Q}\boldsymbol{r}_{t,h\Delta}$ *(Eq. (3)). Then, the autocovariance of the series* $(\sigma^2_{(h),t})_{t\in\mathbb{Z}}$, *for* $\ell \geq 0$, *is given by*

$$\gamma_{\sigma^2_{(h),t}}(\ell) = \text{Cov}(\sigma^2_{(h),t}, \sigma^2_{(h),t-\ell}) = 2\sigma^4\,\text{tr}(\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}'\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}').$$

Proof. See Appendix A.1.

---

[9]For the sake of simplicity, we assume a Gaussian white noise process with zero mean. In case of $\mathbb{E}(r_t) = \mu \neq 0$, we have $\gamma_{\sigma^2_{(h),t}}(\ell) = 2\sigma^4\,\text{tr}(\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}'\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}') + 4\mu^2\sigma^2\boldsymbol{1}'_{h\Delta+\ell}\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}'\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}'\boldsymbol{1}_{h\Delta+\ell}$.

Note that for $\ell > h\Delta$

$$\gamma_{\sigma^2_{(h),t}}(\ell) = 2\sigma^4 \operatorname{tr}(\boldsymbol{KQK'LQL'}) = 2\sigma^4 \operatorname{tr}(\boldsymbol{0}_{(h\Delta \times h\Delta)} \boldsymbol{Q} \, \boldsymbol{0}_{(h\Delta \times h\Delta)} \boldsymbol{Q}) = 0,$$

and, by definition, the ACF for $\ell \geq 0$ is given by $\rho_{\sigma^2_{(h),t}}(\ell) = \gamma_{\sigma^2_{(h),t}}(\ell)/\gamma_{\sigma^2_{(h),t}}(0)$.

In the following, we extend Theorem 1 to a more general class of weak white noise processes which contains many zero-mean weak white noise processes – especially, the Gaussian white noise process with $\mu = 0$ and GARCH(p,q) processes.

Theorem 2. *Let $(x_t)_{t \in \mathbb{Z}}$ be a stochastic process with $\mathbb{E}(|x_t|^i) < \infty$, for $t \in \mathbb{Z}$ and $i \leq 4$. For $t_1, t_2, t_3, t_4 \in \mathbb{Z}$ with $\forall i, j \in \{t_1, t_2, t_3, t_4\}$, $i \neq j$, we assume*

$$\mathbb{E}(x_{t_1}) = 0, \tag{8}$$

$$\mathbb{E}(x_{t_1} x_{t_2} x_{t_3} x_{t_4}) = 0, \tag{9}$$

$$\mathbb{E}(x_{t_1}^2 x_{t_2} x_{t_3}) = 0, \tag{10}$$

$$\mathbb{E}(x_{t_1}^3 x_{t_2}) = 0. \tag{11}$$

Let $\boldsymbol{X} = [x_1, \ldots, x_n]'$ and define $\boldsymbol{X}^{2\odot} = \boldsymbol{X} \odot \boldsymbol{X} = [x_1^2, \ldots, x_n^2]'$. Furthermore, define vector $\boldsymbol{\mu}_{\boldsymbol{X}^{2\odot}} \in \mathbb{R}^{n \times 1}$ and matrices $\boldsymbol{\Sigma_X}, \boldsymbol{\Sigma}_{\boldsymbol{X}^{2\odot}} \in \mathbb{R}^{n \times n}$ by

$$\boldsymbol{\Sigma_X} = \mathbb{E}(\boldsymbol{XX'}), \quad \boldsymbol{\mu}_{\boldsymbol{X}^{2\odot}} = \mathbb{E}(\boldsymbol{X}^{2\odot}) \quad and \quad \boldsymbol{\Sigma}_{\boldsymbol{X}^{2\odot}} = \mathbb{E}(\boldsymbol{X}^{2\odot} \boldsymbol{X}^{2\odot'}) - \boldsymbol{\mu}_{\boldsymbol{X}^{2\odot}} \boldsymbol{\mu}'_{\boldsymbol{X}^{2\odot}},$$

respectively. Then, for symmetric matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$, we have

$$\operatorname{Cov}(\boldsymbol{X'AX}, \boldsymbol{X'BX}) = \operatorname{tr}(\boldsymbol{C}(\boldsymbol{\Sigma}_{\boldsymbol{X}^{2\odot}} + \boldsymbol{\mu}_{\boldsymbol{X}^{2\odot}} \boldsymbol{\mu}'_{\boldsymbol{X}^{2\odot}})) - \operatorname{tr}(\boldsymbol{A\Sigma_X})\operatorname{tr}(\boldsymbol{B\Sigma_X}),$$

where $\boldsymbol{C} = \boldsymbol{ab'} + 2\boldsymbol{A} \odot \boldsymbol{B} \odot (\boldsymbol{1}_n \boldsymbol{1}'_n - \boldsymbol{I}_n)$, with $\boldsymbol{a} = \operatorname{diag}(\boldsymbol{A}) = (\boldsymbol{A} \odot \boldsymbol{I}_n)\boldsymbol{1}_n$ and $\boldsymbol{b} = \operatorname{diag}(\boldsymbol{B}) = (\boldsymbol{B} \odot \boldsymbol{I}_n)\boldsymbol{1}_n$.

Proof. See Appendix A.2.

Again, a corollary to Theorem 2 establishes the autocovariance function of the quadratic-form variance estimator when the daily log-return process, $(r_t)_{t \in \mathbb{Z}}$, is a weak white noise process (Definition 1) satisfying the moment conditions (8)-(11).

Corollary 2. *Let $(r_t)_{t \in \mathbb{Z}}$ be a weak white noise process fulfilling the moment conditions (8)-(11). Moreover, let $\sigma^2 = \operatorname{Var}(r_t) = \mathbb{E}(r_t^2)$ and $\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell} = \boldsymbol{r}_{t,h\Delta+\ell} \odot \boldsymbol{r}_{t,h\Delta+\ell}$, and define vector $\boldsymbol{\mu}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}} \in \mathbb{R}^{h\Delta+\ell \times 1}$ and matrix $\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}} \in \mathbb{R}^{h\Delta+\ell \times h\Delta+\ell}$ by*

$$\boldsymbol{\mu}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}} = \mathbb{E}(\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}) \qquad and \qquad \boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}} = \mathbb{E}(\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell} \boldsymbol{r}^{2\odot'}_{t,h\Delta+\ell}) - \boldsymbol{\mu}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}} \boldsymbol{\mu}'_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}},$$

respectively. Then, considering variance estimates of the form $\sigma^2_{(h),t} = \boldsymbol{r}'_{t,h\Delta} \boldsymbol{Q} \boldsymbol{r}_{t,h\Delta}$, the autocovariance of the series $(\sigma^2_{(h),t})_{t\in\mathbb{Z}}$, for $\ell \geq 0$, is given by

$$\gamma_{\sigma^2_{(h),t}}(\ell) = \operatorname{tr}\big(\boldsymbol{C}\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}}\big) + 2\sigma^4\big(\operatorname{tr}(\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}'\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}') - \boldsymbol{a}'\boldsymbol{b}\big),$$

with $\boldsymbol{C} = \boldsymbol{a}\boldsymbol{b}' + 2(\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}') \odot (\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}') \odot (\boldsymbol{1}_{h\Delta+\ell}\boldsymbol{1}'_{h\Delta+\ell} - \boldsymbol{I}_{h\Delta+\ell})$, where $\boldsymbol{a} = \operatorname{diag}(\boldsymbol{K}\boldsymbol{Q}\boldsymbol{K}')$ and $\boldsymbol{b} = \operatorname{diag}(\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}')$.

Proof. See Appendix A.3.

Remark 1. *The following processes satisfy the conditions of Corollary 2 and especially the moment conditions* (8)-(11)*:*

(i) *For $\mu = 0$, the Gaussian white noise process (Example 1) clearly fulfills all conditions.*

(ii) *Let $(x_t)_{t\in\mathbb{Z}}$ be a GARCH(p,q) process as defined in Example 2. For the innovations, $(\epsilon_t)_{t\in\mathbb{Z}}$, we assume a sequence of i.i.d. random variables being symmetrically distributed such that odd moments are zero. We further assume that the first four moments of $(x_t)_{t\in\mathbb{Z}}$ exist.*[10] *In Appendix A.4 we show that under these conditions the GARCH(p,q) satisfies all conditions such that Corollary 2 holds.*

(iii) *Under some regularity conditions, even more general classes of GARCH processes satisfy the conditions of Corollary 2. For discussions on families of GARCH processes and conditions on stationarity and the existence of moments see He and Teräsvirta (1999b) and Ling and McAleer (2002b).*

(iv) *Due to the moment conditions* (8)-(11)*, the autocovariance in Corollary 2 only depends on the variance $\sigma^2 = \operatorname{Var}(r_t)$, $\boldsymbol{\mu}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}}$, and the variance-covariance matrix of the vector of squares, $\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}}$. If the daily returns are, for example, asymmetrically distributed or follow a GARCH process with leverage, the moment conditions have to be weakened and additional terms, like unconditional skewness, are necessary to compute autocovariances. Note that these moments are often not available in closed form (He et al. 2008).*

The functional form of the relevant unconditional moments for different GARCH processes have been derived in He and Teräsvirta (1999a) and Karanasos (1999).

---

[10] Conditions for the existence of moments can be found in He and Teräsvirta (1999b) and Bollerslev (1986) for the GARCH(1,1) case and in Ling and McAleer (2002a) for GARCH(p,q).

4.2 Illustration

We illustrate the theoretical results of the previous section by presenting plots of the theoretical ACF and those obtained from a simulation study. All illustrations in this section are for a GARCH(1,1) data generating process and the EWMA variance estimator (5).[11]

Let $(r_t)_{t\in\mathbb{Z}}$ be generated by the GARCH(1,1) process

$$r_t = \sigma_t\epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{12}$$

with $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ and parameter vector, $[\alpha_0, \alpha_1, \beta_1]' := [0.01, 0.05, 0.94]'$, parameter values that are typical for daily stock returns. As estimator for the h-day variance at time t we use the EWMA estimator (5) with $\lambda = 0.96$. In view of the Basel III rules (BCBS 2016), we chose $h = 10$ as aggregation horizon, i.e., we consider a bi-weekly target horizon as, for example, in Kole et al. (2017).

As for the ten-day return series itself, we can obtain ten different series of estimates for the variance if we synchronize assessment and sampling frequencies to be equal to the aggregation horizon of $h = 10$ days, namely, $(\sigma_{(h),ht+\tau,\lambda}^2)_{t\in\mathbb{Z}}$, for $1 \leq \tau \leq h$.[12] As window length for the rolling-window estimates we choose $\Delta = 100$, giving rise to $h\Delta = 1000$ daily observations which corresponds to roughly four years of return data. At any point in time, each estimate is based on Δ non-overlapping h-day returns or $h\Delta$ daily returns. Two consecutive estimates, $\sigma_{(h),t,\lambda}^2$ and $\sigma_{(h),t+1,\lambda}^2$ (or $\sigma_{(h),t-1,\lambda}^2$), have $h\Delta - 2$ daily return observations in common. Our simulations are based on the GARCH(1,1) process defined in (12) with a sample size of $8 \times 250 = 2000$ trading days or about eight calendar years.

The left graph in Figure 4 shows the ten different series of variance estimates, $(\sigma_{(h),th+\tau,\lambda}^2)_{t\in\mathbb{Z}}$, $1 \leq \tau \leq h$, obtained when assessment and sampling frequencies are synchronized. In each of the ten plots, two consecutive points, $(\sigma_{(h),th+\tau,\lambda}^2)_{t\in\mathbb{Z}}$, $1 \leq \tau \leq h$, have distance $h = 10$. The graph on the right shows the sequence of daily EWMA variance estimates, $(\sigma_{(h),t,\lambda}^2)_{t\in\mathbb{Z}}$, based on non-overlapping h-day returns.

The plots for the simulations in Figure 4 are constructed as those for the DJIA returns in Figure 2. The daily series of EWMA variances estimates (right graph in Figure 4) fluctuates in a highly regular fashion, mimicking a strong seasonal pattern. From a risk management perspective, such strong fluctuations are bound to have detrimental implications as they induce volatile risk capital charges and risk mitigation activities.

---

[11] Appendix B.1 presents plots for a Gaussian white noise data generating process and variance being estimated by the sample variance (4).

[12] By synchronization of the assessment and sampling frequency we mean that the value of $\tau \in \{1,\ldots,h\}$ is the same for the series of h-day returns, $(r_{(h),ht+\tau,\lambda})_{t\in\mathbb{Z}}$, and the series of (assessed) variance estimates, $(\sigma_{(h),ht+\tau,\lambda}^2)_{t\in\mathbb{Z}}$. That is, both series are sampled on the same equidistant grid where we observe an h-day return and estimate the variance on every h-th day.
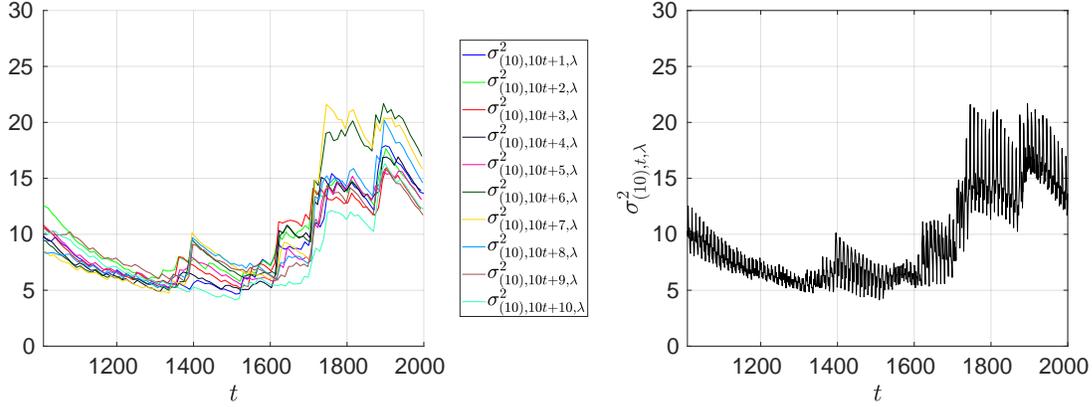
Figure 4: Time series of EWMA variance estimates (5), $\sigma^2_{(h),t,\lambda}$, for simulated daily return series from GARCH(1,1) process (12). The plot on the left shows the estimates $(\sigma^2_{(10),10t+\tau,\lambda})_{t\in\mathbb{Z}}$, for $1 \le \tau \le 10$. The right plot shows the series $(\sigma^2_{(10),t,\lambda})_{t\in\mathbb{Z}}$. Both plots are based on bi-weekly ($h = 10$) returns and estimation window $\Delta = 100$.

To derive the autocovariances of the series of variance estimates, $(\sigma^2_{(h),t,\lambda})_{t\in\mathbb{Z}}$, assuming a GARCH(1,1) process for the daily returns, $(r_t)_{t\in\mathbb{Z}}$, we use the fact that the variance-covariance matrix of the vector of squared returns, $\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell} = [r^2_{t-h\Delta-\ell+1}, \ldots, r^2_t]'$, is given by the symmetric Toeplitz matrix (cf. He and Teräsvirta (1999a) or Karanasos (1999))

$$\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell},[i,i+j]} = \begin{cases} \gamma_{r^2_t}(0) & , 1 \le i \le h\Delta + \ell, j = 0, \\ \frac{\gamma_{r^2_t}(1)}{(\alpha_1+\beta_1)^{1-j}} & , 1 \le i \le h\Delta + \ell - 1, \, 1 \le j \le h\Delta + \ell - i, \end{cases}$$

where $\gamma_{r^2_t}(0)$ and $\gamma_{r^2_t}(1)$ denote the variance and first-order autocovariance, respectively, of the squared returns from a GARCH(1,1) process. Thus, if daily returns follow a GARCH(1,1) process, the autocorrelation of the process of daily exponentially-weighted moving-average variances (5), $(\sigma^2_{(h),t,\lambda})_{t\in\mathbb{Z}}$, based on non-overlapping h-day returns, can be computed via Corollary 2 by plugging in $\boldsymbol{Q}_{(h),\Delta,\lambda}$, $\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}}$ and $\text{Var}(r_t) = \alpha_0/(1 - \alpha_1 - \beta_1)$ into the formula for the autocovariance and scaling via $\rho_{\sigma^2_{(h),t,\lambda}}(\ell) = \gamma_{\sigma^2_{(h),t,\lambda}}(\ell)/\gamma_{\sigma^2_{(h),t,\lambda}}(0)$.

The ACF with estimates based on daily return samples of size $h\Delta = 1000$ is presented on the left in Figure 5. The graph shows the effect for the aggregation horizons $h = 5, 10, 20$, amounting to quasi-weekly, quasi-bi-weekly and quasi-monthly return periods. It demonstrates that the ACF of EWMA variances, $\rho_{\sigma^2_{(h),t,\lambda}}(\ell)$, based on non-overlapping h-day returns, is highly cyclical and slowly decaying. The (spurious) seasonality that is present in the sample ACF of estimated variances for the DJIA data (Figure 3) is compatible with the (spurious) seasonality in the theoretical ACF in Figure 5. The right graph in Figure 5 further illustrates the interaction between aggregation horizon, h, and the window length, Δ. The aggregation horizon is bi-weekly ($h = 10$) and the window length, Δ, assumes values 25, 50 and 100, i.e., roughly one, two and

15
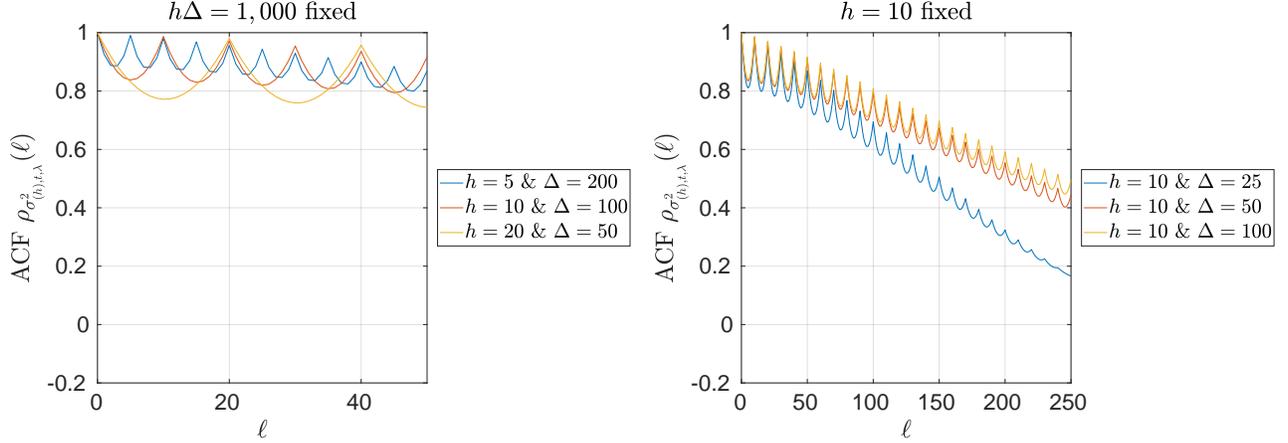
Figure 5: The ACF of EWMA variances (5), $\sigma^2_{(h),t,\lambda}$, for daily returns from GARCH(1,1) process (12). For the left plot we use a fixed number of daily returns to derive the EWMA variances. The right plot depicts the ACF of EWMA variances based on bi-weekly ($h = 10$) returns and estimation windows $\Delta = 25, 50, 100$.

four calendar years of daily return, respectively.

The formula for the autocovariance $\gamma_{\sigma^2_{(h),t,\lambda}}(\ell)$ given in Corollary 2 and for the ACF are rather handy. But it offers little insight into where the spurious seasonality exactly comes from or how the amplitude of the periodic spurious seasonality in the ACF depends on the variance estimator and the data generating process. Appendix B.2 expresses the ACF as the sum of three components, which provide more insight and show that the term $(\boldsymbol{KQK'}) \odot (\boldsymbol{LQL'})$ is crucially responsible for the spurious seasonality in the ACF.

The left graph in Figure 4 shows a pronounced periodicity, though it is not always the same observation within the h-day periods that assumes the highest or lowest value. In other words, the order statistics of the different estimates within an h-day period fluctuate, but do so rather slowly. Therefore, if the focus is on bi-weekly risk estimation but assessment occurs at a daily frequency, then, by construction, the ordering of the ten different variance estimates in a two-week period gradually changes over time. The color-coded series of EWMA variances on the left in Figure 4 make clear that, at some point in time, any particular color may be on top (or bottom) and that there is a high probability that this will also hold for the following h-day period.

To get further insights into why the order statistics gradually change over time, we take a look at the ACF of the first difference of the estimated variances,

$$\gamma_{\sigma^2_{(h),t}-\sigma^2_{(h),t-1}}(\ell) = 2\gamma_{\sigma^2_{(h),t}}(\ell) - \gamma_{\sigma^2_{(h),t}}(\ell+1) - \gamma_{\sigma^2_{(h),t}}(|\ell-1|),$$

$\ell \geq 0$. The ACF of the first differences of the estimated variances, $\rho_{\sigma^2_{(h),t,\lambda}-\sigma^2_{(h),t-1,\lambda}}(\ell)$, is plotted in Figure 6, where we used the same settings as for the ACF of the EWMA variances shown in Figure 5. The plots in Figure 6 demonstrate that the series of first differences is highly

16
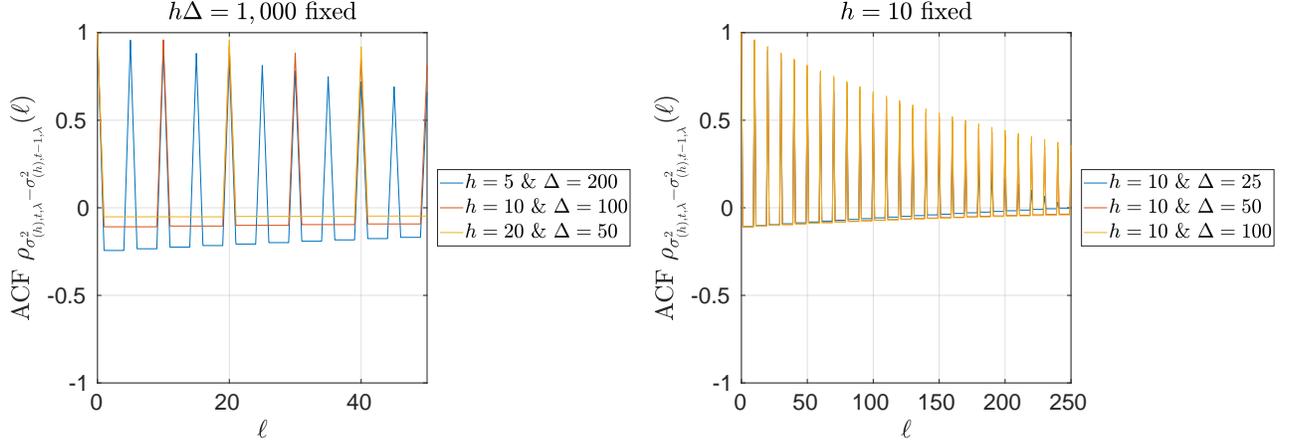
Figure 6: The ACF of the first difference of EWMA variances (5), $\sigma^2_{(h),t,\lambda}$, for daily returns from GARCH(1,1) process (12). For the left plot we use a fixed number of daily returns to derive the EWMA variances. The right plot depicts the ACF of the first difference of EWMA variances based on bi-weekly ($h = 10$) returns and estimation windows $\Delta = 25, 50, 100$.

autocorrelated for lags being multiples of the aggregation horizon, h. This means, the change of the estimate from one day to another is highly autocorrelated with that of h days ago. At lags that are not multiples of h, the autocorrelations of the first-order differences are quite small and slightly negative. This behavior explains the slowly changing ordering of the h different variance series in the left graph of Figure 4.

5 The Case of Overlapping Aggregated Returns

The previous section showed that daily variance estimates based on non-overlapping h-day returns suffer from spurious seasonality. In cases where the aggregation horizon is fixed, the only alternative is to synchronize both assessment and sampling at a daily frequency, i.e., to use overlapping h-day returns for daily risk estimations. In the following we consider several variance estimators for overlapping returns that avoid spurious seasonality.

The simplest variance estimator based on overlapping h-day returns is to apply standard formulas for variance estimators to sets of overlapping return observations. The quadratic-form representations for the sample variance for overlapping h-day returns is given by

$$\breve{\sigma}^2_{(h),t} = \frac{h}{\text{tr}(\breve{\boldsymbol{\mathcal{Q}}}_{(h),\Delta})} \frac{1}{h(\Delta - 1) + 1} \sum_{\tau=0}^{h(\Delta-1)} \left(r_{(h),t-\tau} - \breve{\mu}_{(h),t}\right)^2 = \boldsymbol{r}'_{t,h\Delta} \breve{\boldsymbol{Q}}_{(h),\Delta} \boldsymbol{r}_{t,h\Delta}, \qquad (13)$$

with $\breve{\boldsymbol{\mathcal{Q}}}_{(h),\Delta} = \frac{1}{h(\Delta-1)+1} \sum_{j=0}^{h-1} \boldsymbol{S}'_j \boldsymbol{S}_j$, $\boldsymbol{S}_0 = \boldsymbol{H}' - \mathbf{1}_\Delta \mathbf{1}'_\Delta \boldsymbol{J}'$ and $\boldsymbol{S}_j = \boldsymbol{H}'_j - \boldsymbol{a}\mathbf{1}'_\Delta \boldsymbol{J}'$, for $1 \leq j \leq h-1$,

17

where $\boldsymbol{J} = \boldsymbol{H} + \sum_{j=1}^{h-1} \boldsymbol{H}_j$, $\boldsymbol{a} = \begin{bmatrix} \mathbf{1}'_{\Delta-1} & 0 \end{bmatrix}'$ and

$$\boldsymbol{H}_j = \begin{bmatrix} \mathbf{0}_{(h-j\times 1)} & \mathbf{0}_{(h-j\times\Delta-1)} \\ \mathbf{0}_{(h(\Delta-1)\times 1)} & \boldsymbol{I}_{\Delta-1} \otimes \mathbf{1}_h \\ \mathbf{0}_{(j\times 1)} & \mathbf{0}_{(j\times\Delta-1)} \end{bmatrix},$$

and the EWMA variance for overlapping h-day returns by

$$\check{\sigma}^2_{(h),t,\lambda} = \frac{h}{\mathrm{tr}(\check{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,\lambda})} \frac{1-\lambda^{\frac{1}{h}}}{1-\lambda^{\frac{h(\Delta-1)+1}{h}}} \sum_{\tau=0}^{h(\Delta-1)} \lambda^{\frac{\tau}{h}} \big(r_{(h),t-\tau} - \check{\mu}_{(h),t,\lambda} \big)^2 = \boldsymbol{r}'_{t,h\Delta} \check{\boldsymbol{Q}}_{(h),\Delta,\lambda} \boldsymbol{r}_{t,h\Delta}, \qquad (14)$$

with $\check{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,\lambda} = \sum_{j=0}^{h-1} \lambda^{\frac{j}{h}} \tilde{\boldsymbol{S}}'_j \boldsymbol{\Gamma} \tilde{\boldsymbol{S}}_j$, $\tilde{\boldsymbol{S}}_0 = \boldsymbol{H}' - \mathbf{1}_\Delta \boldsymbol{\gamma}' \boldsymbol{G}'$ and $\tilde{\boldsymbol{S}}_j = \boldsymbol{H}'_j - \boldsymbol{a}\boldsymbol{\gamma}'\boldsymbol{G}'$, for $1 \leq j \leq h-1$, with $\boldsymbol{G} = \boldsymbol{H} + \sum_{j=1}^{h-1} \lambda^{\frac{j}{h}} \boldsymbol{H}_j$ and, for $\lambda \in (0,1)$, $\boldsymbol{\Gamma} = \mathrm{Diag}(\boldsymbol{\gamma}) = (\boldsymbol{\gamma}\mathbf{1}'_\Delta) \odot \boldsymbol{I}_\Delta$, where $\boldsymbol{\gamma} = \frac{1-\lambda^{1/h}}{1-\lambda^{(h(\Delta-1)+1)/h}} \big[\lambda^{\Delta-1}, \lambda^{\Delta-2}, \dots, \lambda^1, 1, \big]'$. Graphical evidence (not shown here) indicates that these variance estimators do not suffer from spurious seasonality.[13]

Another approach to avoid spurious seasonality is to simply take the average of the last h sample variances based on non-overlapping h-day returns. In the (ultra-)high-frequency context, this type of post-averaging of subsampling-based variance estimates has been proposed in Zhang et al. (2005) to overcome problems arising from microstructure noise. It is referred to as two-scales realized volatility and provides a consistent estimator of integrated volatility under the assumption of additive white noise. In the following, we show that the two-scales estimator has the potential to solve the spurious seasonality problem in variance estimation.[14]

For our setting, we obtain the two-scales sample variance[15]

$$\bar{\sigma}^2_{(h),t} = \frac{1}{h} \boldsymbol{r}'_{t,h\Delta} \boldsymbol{Q}_{(h),\Delta} \boldsymbol{r}_{t,h\Delta} + \frac{1}{h} \sum_{j=1}^{h-1} \boldsymbol{r}'_{t-j,h(\Delta-1)} \boldsymbol{Q}_{(h),\Delta-1} \boldsymbol{r}_{t-j,h(\Delta-1)} = \boldsymbol{r}'_{t,h\Delta} \bar{\boldsymbol{Q}}_{(h),\Delta} \boldsymbol{r}_{t,h\Delta}, \qquad (15)$$

---

[13]It turns out that there are two additional variance estimators based on overlapping returns that overcome spurious seasonality. In this section we report results for only one of the three variance estimators and refer to Appendix B.3 for graphical results and a comparison of all four EWMA variance estimators considered.

[14]In the high-frequency literature, several other variance estimators, such as the the multi-scales realized volatility (Zhang 2006) and the pre-averaging approach (Jacod et al. 2009), have been proposed. As the two-scales estimator, they can overcome spurious seasonality. We restrict ourselves, however, to the two-scales estimator of Zhang et al. (2005), since it is the simplest variance estimator handling the problem of spurious seasonality.

[15]Use of $\boldsymbol{\mathcal{Q}}_{(h),\Delta}$ instead of $\boldsymbol{Q}_{(h),\Delta}$ in (15) yields the biased "second-best" two-scales variance estimator of Zhang et al. (2005). The biased-corrected version (15), $\bar{\sigma}^2_{(h),t}$, is obtained by using $\boldsymbol{Q}_{(h),\Delta}$.

with $\bar{\boldsymbol{Q}}_{(h),\Delta} = \frac{1}{h}\boldsymbol{Q}_{(h),\Delta} + \frac{1}{h}\sum_{j=1}^{h-1}\boldsymbol{T}_j(\boldsymbol{Q}_{(h),\Delta-1})$, where for symmetric matrices $\boldsymbol{Q} \in \mathbb{R}^{h(\Delta-1)\times h(\Delta-1)}$

$$\boldsymbol{T}_j(\boldsymbol{Q}) = \begin{bmatrix} \boldsymbol{0}_{(h-j\times h-j)} & \boldsymbol{0}_{(h-j\times h(\Delta-1))} & \boldsymbol{0}_{(h-j\times j)} \\ \boldsymbol{0}_{(h(\Delta-1)\times h-j)} & \boldsymbol{Q} & \boldsymbol{0}_{(h(\Delta-1)\times j)} \\ \boldsymbol{0}_{(j\times h-j)} & \boldsymbol{0}_{(j\times h(\Delta-1))} & \boldsymbol{0}_{(j\times j)} \end{bmatrix}.$$

For the two-scales EWMA variance we have

$$\bar{\sigma}^2_{(h),t,\lambda} = \frac{1-\lambda^{\frac{1}{h}}}{1-\lambda}\left(\boldsymbol{r}'_{t,h\Delta}\boldsymbol{Q}_{(h),\Delta,\lambda}\boldsymbol{r}_{t,h\Delta} + \sum_{j=1}^{h-1}\lambda^{\frac{j}{h}}\boldsymbol{r}'_{t-j,h(\Delta-1)}\boldsymbol{Q}_{(h),\Delta-1,\lambda}\boldsymbol{r}_{t-j,h(\Delta-1)}\right) = \boldsymbol{r}'_{t,h\Delta}\bar{\boldsymbol{Q}}_{(h),\Delta,\lambda}\boldsymbol{r}_{t,h\Delta}, \tag{16}$$

with $\bar{\boldsymbol{Q}}_{(h),\Delta,\lambda} = \frac{1-\lambda^{\frac{1}{h}}}{1-\lambda}\boldsymbol{Q}_{(h),\Delta,\lambda} + \frac{1-\lambda^{\frac{1}{h}}}{1-\lambda}\sum_{j=1}^{h-1}\lambda^{\frac{j}{h}}\boldsymbol{T}_j(\boldsymbol{Q}_{(h),\Delta-1,\lambda})$.

The variance estimators studied so far can be written as quadratic forms and, thus, straightforwardly visualized in form of heatmaps as shown in Figure 7. Each pixel in a heatmap corresponds to an entry in matrix $\boldsymbol{Q}$ in quadratic form $\boldsymbol{r}'_{t,h\Delta}\boldsymbol{Q}\boldsymbol{r}_{t,h\Delta}$ in (3). The entries in $\boldsymbol{Q}$ can be interpreted as weights of cross-products $r_{t-x}r_{t-y}$, whose magnitude is indicated by the color scale in Figure 7. It is evident that the subsampling-based post-average EWMA variance estimator (16), shown in the lower left plot in Figure 7 as well as the EWMA variance based on overlapping h-day returns (top right) given by (14) have boundary problems. The weights assigned to the most recent squared return observations, r^2_{t-j}, $0 \leq j \leq h-1$, are much lower than the weight of the lagged squared return, r^2_{t-h}.

To eliminate such undesirable boundary effects, we propose modifications to the two-scales estimator of Zhang et al. (2005) and the EWMA variant. The two-scales sample variance with boundary-correction is given by

$$\tilde{\sigma}^2_{(h),t} = \boldsymbol{r}'_{t,h\Delta}\tilde{\boldsymbol{Q}}_{(h),\Delta}\boldsymbol{r}_{t,h\Delta}, \tag{17}$$

with $\tilde{\boldsymbol{Q}}_{(h),\Delta} = \frac{h}{\text{tr}(\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta})}\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta}$ and $\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta} \in \mathbb{R}^{h\Delta\times h\Delta}$ being a symmetric Toeplitz matrix with the j-th (off-)diagonal element given by

$$\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,[i,i+j]} = \begin{cases} \frac{1}{\Delta}\left(1-\frac{1}{\Delta}\right)\left(1-\frac{j\Delta}{h\Delta-j}\right) & ,1 \leq i \leq h\Delta,\ 0 \leq j \leq \min\{h\Delta-i,h-1\}, \\ -\frac{1}{\Delta^2} & ,1 \leq i \leq h(\Delta-1),\ h \leq j \leq h\Delta-i. \end{cases}$$

The two-scales EWMA variance with boundary-correction becomes

$$\tilde{\sigma}^2_{(h),t,\lambda} = \boldsymbol{r}'_{t,h\Delta}\tilde{\boldsymbol{Q}}_{(h),\Delta,\lambda}\boldsymbol{r}_{t,h\Delta}, \tag{18}$$

where $\tilde{\boldsymbol{Q}}_{(h),\Delta,\lambda} = \frac{h}{\text{tr}(\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,\lambda})}\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,\lambda}$, with symmetric matrix $\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,\lambda} \in \mathbb{R}^{h\Delta\times h\Delta}$ being defined by
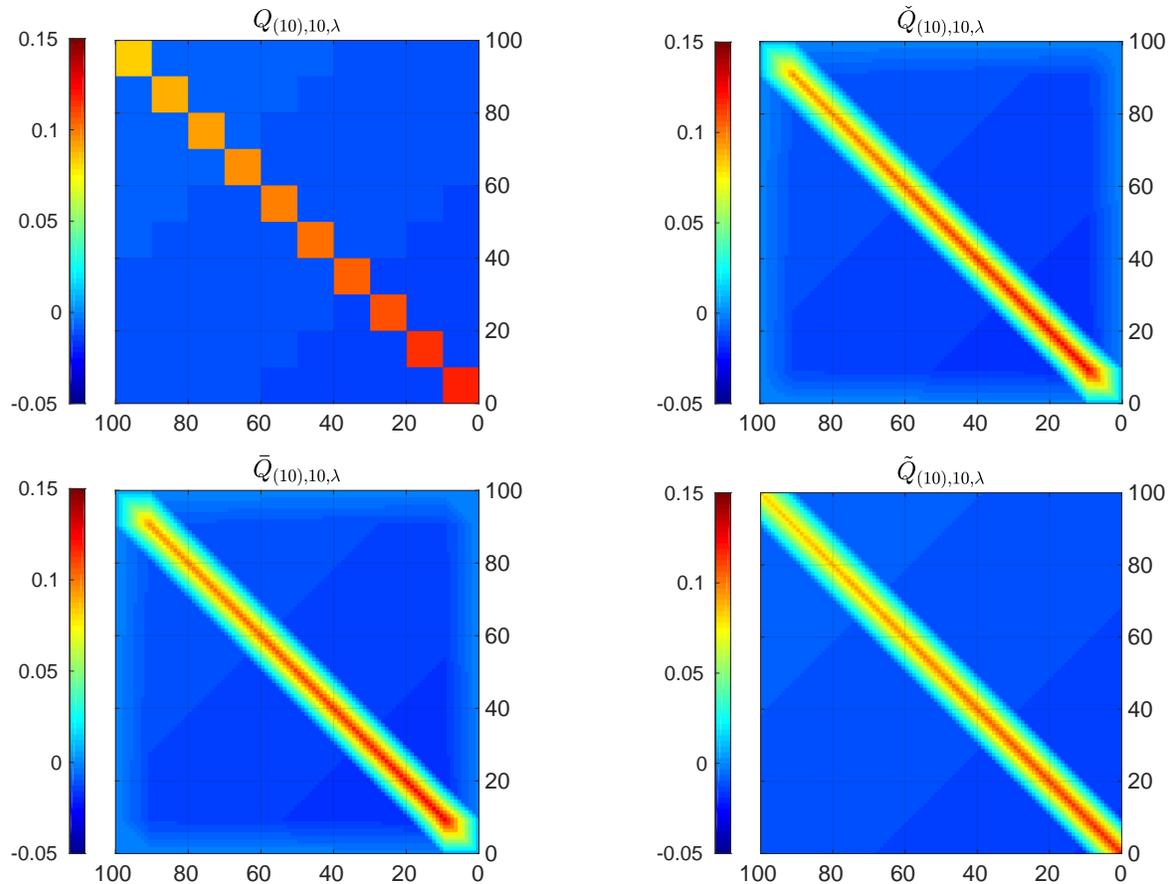
Figure 7: Heatmaps of the quadratic-form matrices $\boldsymbol{Q}$ for different EWMA variance estimators of the form (3). For this illustration, we set $h = 10$, $\Delta = 10$ and $\lambda = 0.96$. The pixels correspond to the entries of $\boldsymbol{Q}$ and reflect the weights of cross-products $r_{t-x}r_{t-y}$. The magnitude of a weight is indicated by the color scale. The upper plots show the standard estimator based on non-overlapping h-day returns (left; Eq. (5)) and that based on overlapping h-day returns (right; Eq. (14)). The lower plots show subsampling-based post-average EWMA variance estimators. The left plot corresponds to the two-scales variance (16) and the right to the boundary corrected version (18).

$\tilde{\boldsymbol{\mathcal{Q}}}_{(h),\Delta,\lambda} = \boldsymbol{\Psi} - \boldsymbol{\Xi}$, where $\boldsymbol{\Psi}, \boldsymbol{\Xi} \in \mathbb{R}^{h\Delta \times h\Delta}$ are symmetric matrices with entries

$$
\boldsymbol{\Psi}_{[i,i+j]} = \begin{cases} \frac{(h-j)(1-\lambda^{\frac{1}{h}})}{1-\lambda^{\frac{h\Delta-j}{h}}} \lambda^{\frac{h\Delta-i}{h}} & , 1 \leq i \leq h\Delta,\ 0 \leq j \leq \min\{h\Delta - i, h-1\}, \\ 0 & , 1 \leq i \leq h(\Delta - 1),\ h \leq j \leq h\Delta - i, \end{cases}
$$

and, for $1 \leq i \leq h\Delta$, $0 \leq j \leq h\Delta - i$,

$$
\boldsymbol{\Xi}_{[i,i+j]} = \lambda^{\frac{2(h\Delta-i-j)}{h}+\delta} \frac{(1-\lambda)^2 (1-\lambda^{\frac{2}{h}}) \left[(h-k)(1-\lambda^{2(\Delta-\delta)}) + k\lambda(1-\lambda^{2(\Delta-\delta-1)}) \right]}{(1-\lambda^{\Delta})^2 (1-\lambda^2)(1-\lambda^{\frac{2(h\Delta-j)}{h}})},
$$

with $\delta = \lfloor \frac{j}{h} \rfloor$, $k = j - h\lfloor \frac{j}{h} \rfloor = j - h\delta$ and $\lambda \in (0,1)$. The heatmap of the weighting scheme for the two-scales EWMA variance estimator with boundary-correction (18), associated with the
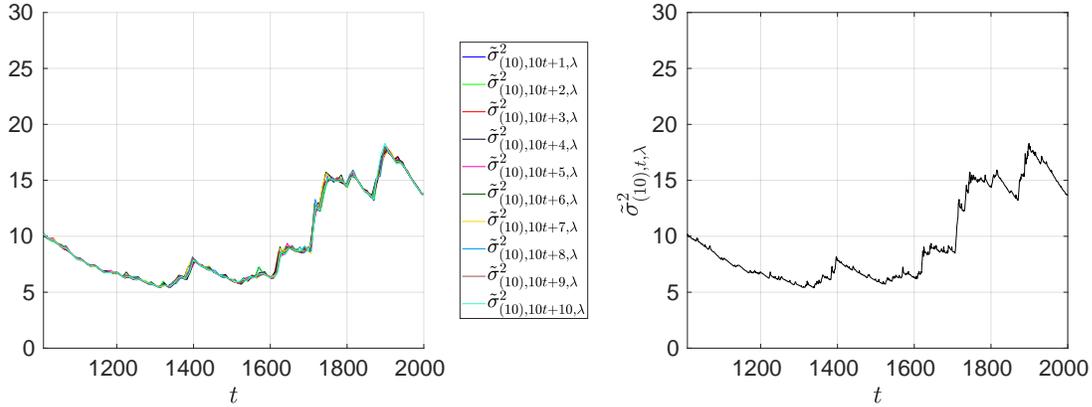
Figure 8: Time series of subsampling-based post-average EWMA variance estimates with boundary correction (18), $\tilde{\sigma}^2_{(h),t,\lambda}$, for simulated daily return series from GARCH(1,1) process (12). The plot on the left shows the estimates $(\tilde{\sigma}^2_{(10),10t+\tau,\lambda})_{t\in\mathbb{Z}}$, for $1 \leq \tau \leq 10$. The right plot shows the series $(\tilde{\sigma}^2_{(10),t,\lambda})_{t\in\mathbb{Z}}$. Both plots are based on bi-weekly ($h = 10$) returns and estimation window $\Delta = 100$.

quadratic-form matrix $\tilde{\boldsymbol{Q}}_{(h),\Delta,\lambda}$, is depicted in the lower right of Figure 7.

If we apply the subsampling-based post-average EWMA variance estimator with boundary correction (18) to the same simulated GARCH(1,1) series used in Figure 4, we obtain the series of variance estimators plotted in Figure 8. The right panel in Figure 8 clearly shows the absence of spurious seasonality as compared to the right panel in Figure 4. The ten different series of variance estimates (left in Figure 8), where assessment and sampling frequencies are in sync and equal to the aggregation horizon, h, turn out to be much more stable; and long-memory effects in the order statistics, associated with h-day periods, are no longer present.

Note that the autocovariances for the daily subsampling-based post-average EWMA variance estimates (with boundary correction) can again be obtained from Corollary 2. The corresponding ACF of the estimated variances is shown in Figure 9. The data generating process and the combinations of aggregation horizons, h, and estimation window length, Δ, are the same, as in Figure 5. The periodicity in the ACF has been eliminated, and the functional form of the ACF seems reasonable for an EWMA type estimator of the variance.

In Figure 10 we repeat the DJIA analysis (Section 2) using the boundary-corrected two-scales variance estimates (18) instead. The construction of the plot is exactly as in Figure 2. We see that the problem of spurious seasonality is no longer present and that the ten different variance series, shown on the left in Figure 10, do no longer slowly change their position.

21
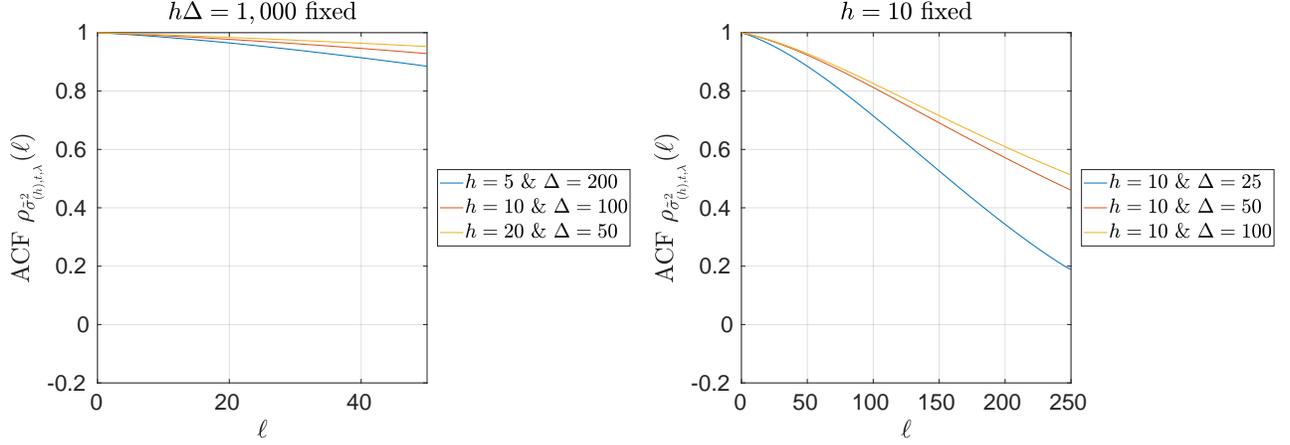
Figure 9: The ACF of subsampling-based post-average EWMA variances with boundary correction (18), for daily returns from GARCH(1,1) process (12). For the left plot we use a fixed number of daily returns to derive the EWMA variances. The right plot depicts the ACF of EWMA variances based on bi-weekly ($h = 10$) returns and estimation windows $\Delta = 25, 50, 100$.
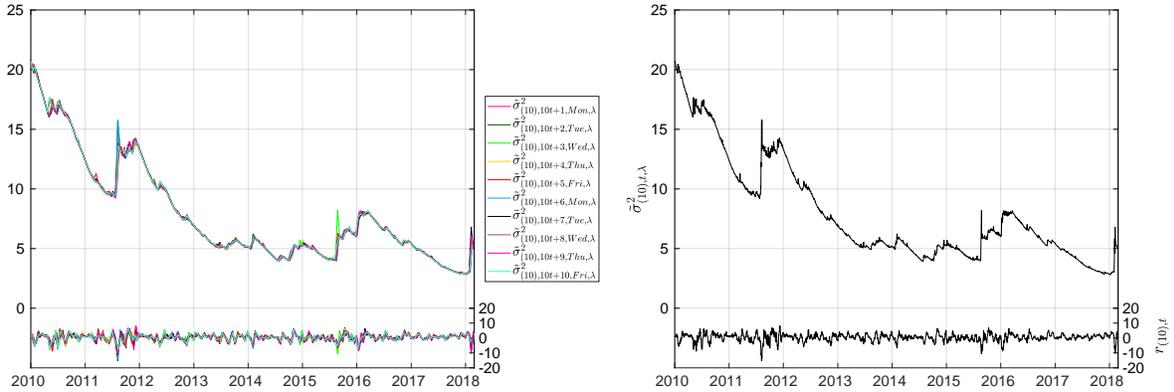
Figure 10: Estimated boundary-corrected two-scales EWMA variances of the Dow Jones Industrial Average (DJIA) based on ten-day log-returns with a window length of 100 bi-weekly returns and an EWMA parameter of $\lambda = 0.96$. The first (last) estimates in both graphs are for 01-Jan-2010 (for 28-Feb-2018). The graph on the left shows at the top ten series of bi-weekly variance estimates, each corresponding to a specific weekday and start date, and the one on the right the daily series of bi-weekly variance estimates. The corresponding ten-day log-returns are plotted at the bottom of both graphs.

6   Properties of the Variance Estimators

Apart from the presence or absence of spurious seasonality, properties such as bias, variance and mean-squared error (MSE) of variance estimators are typically of interest. Another concern is the question whether or not the volatility dynamics are captured in an adequate fashion (Engle and Patton 2001). This is reflected by the responsiveness of variance estimators with respect to shocks. Both issues are addressed next.
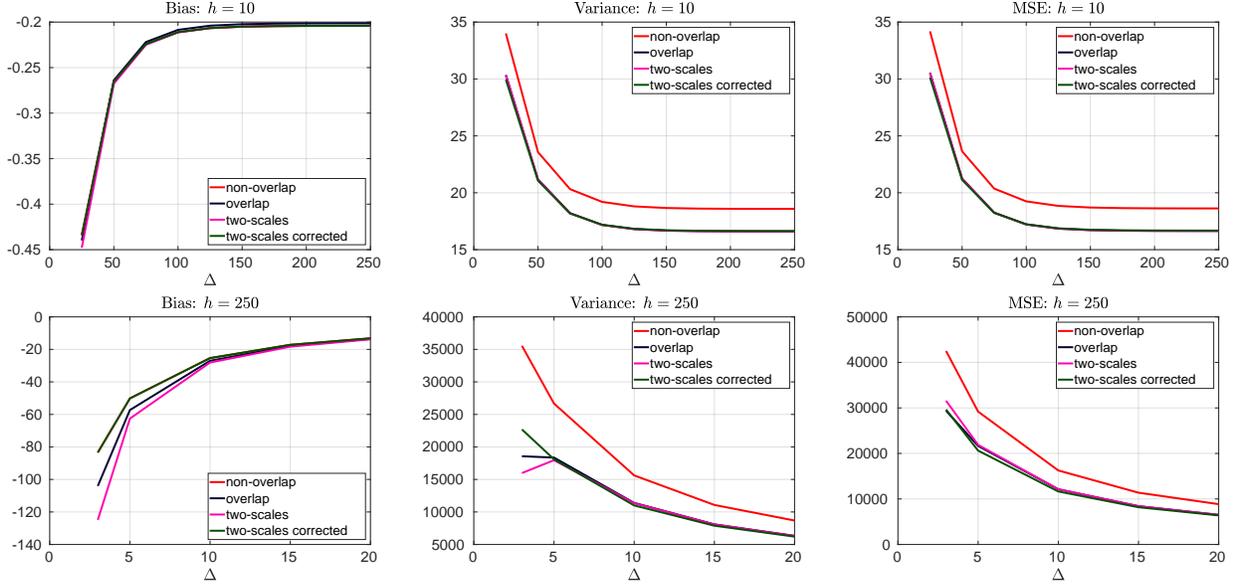
Figure 11: Bias, variance and mean squared error (MSE) of EWMA variance estimators (5), (14), (16), and (18). For the series of daily returns, $(r_t)_{t \in \mathbb{Z}}$, GARCH(1,1) process (12) is assumed. The upper (lower) panel shows results for aggregation horizon $h = 10$ ($h = 250$) for different window sizes, Δ. The results for bias, variance and MSE are shown in left, center and right panels, respectively.

6.1 Bias, Variance and MSE

We use the quadratic-form representation (3) for estimating the unconditional variance of a weak white noise process that satisfies the moment conditions (8)-(11) to derive the bias, variance and mean squared error (MSE) of the variance estimators. Denoting the unconditional variance of process $(r_t)_{t \in \mathbb{Z}}$ by $\sigma^2 = \text{Var}(r_t)$, the bias of $\sigma^2_{(h),t}$ is given by (2) as $\text{Bias}(\sigma^2_{(h),t}) = \sigma^2(\text{tr}(\boldsymbol{Q}) - h)$. From Corollary 2 we obtain

$$\text{Var}(\sigma^2_{(h),t}) = \gamma_{\sigma^2_{(h),t}}(0) = \text{tr}\big(\boldsymbol{C}\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}}\big) + \sigma^4\big(\text{tr}(\boldsymbol{Q}\boldsymbol{Q}) - \boldsymbol{q}'\boldsymbol{q}\big),$$

with $\boldsymbol{C} = \boldsymbol{q}\boldsymbol{q}' + 2\boldsymbol{Q}^{2\odot} \odot (\mathbf{1}_{h\Delta}\mathbf{1}'_{h\Delta} - \boldsymbol{I}_{h\Delta})$ and $\boldsymbol{q} = \text{diag}(\boldsymbol{Q}) = (\boldsymbol{Q} \odot \boldsymbol{I}_{h\Delta})\mathbf{1}_{h\Delta}$.

In the following, we assume $(r_t)_{t \in \mathbb{Z}}$ follows the GARCH(1,1) process (12). As estimators for the variance we analyze the different EWMA variance estimators $\sigma^2_{(h),t,\lambda}$ (non-overlapping h-day returns, (5)), $\breve{\sigma}^2_{(h),t,\lambda}$ (overlapping h-day returns, (14)), $\bar{\sigma}^2_{(h),t,\lambda}$ (two-scales, (16)), and $\tilde{\sigma}^2_{(h),t,\lambda}$ (corrected two-scales, (18)). The top panel in Figure 11 shows the bias, variance and MSE for $h = 10$ and window sizes, Δ, ranging from 25 to 250. For a low aggregation horizon ($h = 10$) all estimators have a similar bias. This is in line with the findings of Bod et al. (2002). With respect to variance and MSE, the three estimators based on overlapping returns produce smaller values. The results differ, however, when the aggregation level increases to $h = 250$ (bottom panel in Figure 11). The standard overlapping estimator, $\breve{\sigma}^2_{(h),t}$, and the two-scales estimator, $\bar{\sigma}^2_{(h),t}$, produce the highest absolute bias. In terms of the MSE, the corrected two-scales estimator

23
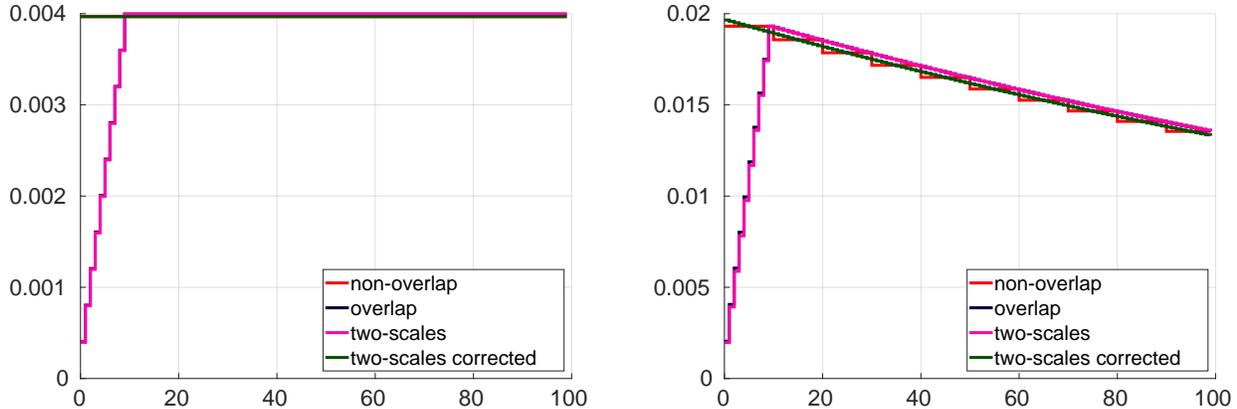
Figure 12: Values of the diagonal elements of the quadratic-form matrices for variance estimators of the form (3), $\boldsymbol{r}'_{t,h\Delta}\boldsymbol{Q}\boldsymbol{r}_{t,h\Delta}$. The horizontal axis indicates the diagonal position in matrix $\boldsymbol{Q}$. The diagonal entries reflect the weights assigned to the squared daily returns, r^2_{t-x}. The left plot shows the entries for variance estimators (4), (13), (15), and (17); and the one on the right for EWMA variance estimators (5), (14), (16), and (18).

performs best.

6.2   Responsiveness to Shocks

Especially the EWMA variance estimator, which is not only used for estimating the unconditional but also the conditional variance, turns out to be more responsive to recent shocks. To illustrate and compare the responsiveness of the estimators we focus on the last 100 diagonal elements of the quadratic-form matrices $\boldsymbol{Q}$. These elements correspond to the weights the respective variance estimators assign to the squared daily returns $r^2_t, r^2_{t-1}, \ldots, r^2_{t-99}$. The left (right) graph in Figure 12 plots the weights of the different sample (EWMA) variance estimators. As becomes evident, the corrected two-scales estimator does not suffer from the boundary issue as do the standard two-scales estimator and the standard variance estimators based on overlapping returns. Furthermore, in the EWMA case, the corrected two-scales estimator allocates the weights more smoothly to past squared shocks than the estimators based on non-overlapping returns.

In summary, the corrected two-scales estimator does not suffer from spurious seasonality and dominates other overlapping-return estimators in terms of bias, variance and MSE as well as the responsiveness with respect to recent shocks. A shortcoming of the corrected two-scales estimator is the fact that, in contrast to the other estimators discussed, it cannot be directly expressed as an estimator based on (non-)overlapping h-day returns. But, as the other estimators, it has a quadratic-form representation in terms of the daily return vector $\boldsymbol{r}_{t,h\Delta}$, i.e., $\boldsymbol{r}'_{t,h\Delta}\boldsymbol{Q}\boldsymbol{r}_{t,h\Delta}$ (3).

7 Concluding Remarks

We have investigated the phenomenon of spurious seasonality in sequentially estimated variances. It arises when the assessment frequency is higher than the sampling frequency of the (non-overlapping) return data used for estimation. The phenomenon, which, to our knowledge, has not yet been addressed in the literature, is attributable to an implicit overlap in the return data used for estimation. To provide a better understanding of this phenomenon, we have analyzed the properties of series of variance estimates in terms of their theoretical autocorrelation functions, considering a large class of data generating processes and various alternative variance estimators. We have shown ways how to overcome the problem of spurious seasonality, introducing an EWMA-based estimator and a boundary correction for the two-scales estimator of Zhang et al. (2005).

In our analysis, we have focused exclusively on variance estimation. However, the phenomenon of spurious seasonality also translates directly to other risk measures, such as value-at-risk or expected shortfall, which are widely used in order to determine the capital requirements of financial institutions. As a consequence, capital charges based on such risk estimates will be subject to spurious seasonality. Risk managers and regulators need to be aware of that phenomenon and, more importantly, understand it in order to establish sound risk management practices. Our findings also provide an explanation for the variation in daily GARCH-parameter estimates derived from different non-overlapping monthly samples reported in Hedegaard and Hodrick (2016). Finally, although we have simplified our discussion by focussing on a daily data frequency, it should be understood that spurious seasonality also arises with other frequencies, such as in (ultra-)high-frequency realized-volatility analysis.

References

Andersen, T. G., and T. Bollerslev. 1998. "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts". *International Economic Review* 39 (4): 885–905.

Andersen, T. G., T. Bollerslev, and S. Lange. 1999. "Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon". *Journal of Empirical Finance* 6 (5): 457–477.

BCBS. 2016. "Minimum capital requirements for market risk". http://www.bis.org/bcbs/publ/d352.htm.

Bod, P., D. Blitz, P. H. Franses, and R. Kluitman. 2002. "An unbiased variance estimator for overlapping returns." *Applied Financial Economics* 12 (3): 155–158.

Bollerslev, T. 1986. "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics* 31 (3): 307–327.

Christoffersen, P., F. X. Diebold, and T. Schuermann. 1998. "Horizon problems and extreme events in financial risk management". *Economic Policy Review* 4 (3): 98–16.

Daníelsson, J., and C. Zhou. 2016. "Why risk is so hard to measure". DNB Working Paper, No. 494 / January 2016.

Daníelsson, J., and J.-P. Zigrand. 2006. "On time-scaling of risk and the square-root-of-time rule". *Journal of Banking & Finance* 30 (10): 2701–2713.

Daníelsson, J., K. R. James, M. Valenzuela, and I. Zer. 2016. "Model risk of risk models". *Journal of Financial Stability* 23: 79–91.

Diebold, F. X., A. Hickman, A. Inoue, and T. Schuermann. 1997. "Converting 1-day volatility to h-day volatility: scaling by h is worse than you think". *Wharton Financial Institutions Center, Working Paper 97-34.*

Drost, F. C., and T. E. Nijman. 1993. "Temporal Aggregation of GARCH Processes". *Econometrica* 61 (4): 909–927.

Embrechts, P., R. Kaufmann, and P. Patie. 2005. "Strategic Long-Term Financial Risks: Single Risk Factors". *Computational Optimization and Applications* 32 (1): 61–90.

Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation". *Econometrica* 50 (4): 987–1007.

Engle, R. F., and A. J. Patton. 2001. "What good is a volatility model?" *Quantitative Finance* 1 (2): 237–245.

Hansen, L. P., and R. J. Hodrick. 1980. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis". *Journal of Political Economy* 88 (5): 829–853.

He, C., and T. Teräsvirta. 1999a. "Fourth Moment Structure of the GARCH(p,q) Process". *Econometric Theory* 15 (6): 824–846.

— . 1999b. "Properties of moments of a family of GARCH processes". *Journal of Econometrics* 92 (1): 173–192.

He, C., A. Silvennoinen, and T. Teräsvirta. 2008. "Parameterizing Unconditional Skewness in Models for Financial Time Series". *Journal of Financial Econometrics* 6 (2): 208–230.

Hedegaard, E., and R. J. Hodrick. 2016. "Estimating the risk-return trade-off with overlapping data inference". *Journal of Banking & Finance* 67: 135–145.

Horváth, L., P. Kokoszka, and R. Zitikis. 2006. "Sample and Implied Volatility in GARCH Models". *Journal of Financial Econometrics* 4 (4): 617–635.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter. 2009. "Microstructure noise in the continuous case: The pre-averaging approach". *Stochastic Processes and their Applications* 119 (7): 2249–2276.

Karanasos, M. 1999. "The second moment and the autocovariance function of the squared errors of the GARCH model". *Journal of Econometrics* 90 (1): 63–76.

Kluitman, R., and P. H. Franses. 2002. "Estimating volatility on overlapping returns when returns are autocorrelated." *Applied Mathematical Finance* 9 (3): 179–188.

Kole, E., T. Markwat, A. Opschoor, and D. van Dijk. 2017. "Forecasting Value-at-Risk under Temporal and Portfolio Aggregation". *Journal of Financial Econometrics* 15 (4): 649–677.

Ling, S., and M. McAleer. 2002a. "Necessary and Sufficient Moment Conditions for the GARCH(r,s) and Asymmetric Power GARCH(r,s) Models". *Econometric Theory* 18 (3): 722–729.

— . 2002b. "Stationarity and the existence of moments of a family of GARCH processes". *Journal of Econometrics* 106 (1): 109–117.

Magnus, J. R. 1978. "The moments of products of quadratic forms in normal variables". *Statistica Neerlandica* 32 (4): 201–210.

McNeil, A. J., and R. Frey. 2000. "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach". Special issue on Risk Management, *Journal of Empirical Finance* 7 (3–4): 271–300.

Mittnik, S. 1988. "Derivation of the theoretical autocovariance and autocorrelation function of autoregessive moving average processes". *Communications in Statistics - Theory and Methods* 17 (11): 3825–3831.

— . 2011. "Solvency II Calibrations: Where Curiosity Meets Spuriosity". Center for Quantitative Risk Analysis (CEQURA), Working Paper Number 04.

Nelson, D. B., and C. Q. Cao. 1992. "Inequality Constraints in the Univariate GARCH Model". *Journal of Business & Economic Statistics* 10 (2): 229–235.

Riskmetrics, T. 1996. *JP Morgan Technical Document*.

Silvestrini, A., and D. Veredas. 2008. "Temporal aggregation of univariate and multivariate time series models: A survey". *Journal of Economic Surveys* 22 (3): 458–497.

Zhang, L. 2006. "Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach". *Bernoulli* 12 (6): 1019–1043.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia. 2005. "A Tale of Two Time Scales". *Journal of the American Statistical Association* 100 (472): 1394–1411.

A Proofs

A.1 Proof of Corollary 1

Assume for the daily returns, $(r_t)_t$, a Gaussian white noise process (Example 1) with $\mathbb{E}(r_t) = 0$ and variances $\text{Var}(r_t) = \sigma^2$. Then $\mathbb{E}(\boldsymbol{r}_{t,h\Delta+\ell}) = \mathbf{0}_{(h\Delta+\ell\times 1)}$ and $\mathbb{E}(\boldsymbol{r}_{t,h\Delta+\ell}\boldsymbol{r}'_{t,h\Delta+\ell}) = \sigma^2 \boldsymbol{I}_{h\Delta+\ell}$, so that, due to the independence, the joint distribution of the vector $\boldsymbol{r}_{t,h\Delta+\ell}$ is a multivariate normal distribution with zero mean vector and variance-covariance matrix $\sigma^2 \boldsymbol{I}_{h\Delta+\ell}$. Using Theorem 1 it follows immediately

$$\text{Cov}(\sigma^2_{(h),t}, \sigma^2_{(h),t-\ell}) = 2\,\text{tr}(\boldsymbol{KQK}'\sigma^2\boldsymbol{I}_{h\Delta+\ell}\boldsymbol{LQL}'\sigma^2 I_{h\Delta+\ell}) = 2\sigma^4\,\text{tr}(\boldsymbol{KQK}'\boldsymbol{LQL}'). \qquad \square$$

A.2 Proof of Theorem 2

$$\begin{aligned}
\text{Cov}(\boldsymbol{X}'\boldsymbol{AX}, \boldsymbol{X}'\boldsymbol{BX}) &= \mathbb{E}(\boldsymbol{X}'\boldsymbol{AXX}'\boldsymbol{BX}) - \mathbb{E}(\boldsymbol{X}'\boldsymbol{AX})\mathbb{E}(\boldsymbol{X}'\boldsymbol{BX}) \\
&= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n a_{ij}b_{kl}\mathbb{E}(x_i x_j x_k x_l) - \text{tr}(\mathbb{E}(\boldsymbol{X}'\boldsymbol{AX}))\text{tr}(\mathbb{E}(\boldsymbol{X}'\boldsymbol{BX})) \\
&= \sum_{i=1}^n \sum_{j=1}^n a_{ii}b_{jj}\mathbb{E}(x_i^2 x_j^2) + \sum_{i=1}^n \sum_{j=1,j\neq i}^n (a_{ij}b_{ij} + a_{ij}b_{ji})\mathbb{E}(x_i^2 x_j^2) - \mathbb{E}(\text{tr}(\boldsymbol{X}'\boldsymbol{AX}))\mathbb{E}(\text{tr}(\boldsymbol{X}'\boldsymbol{BX})) \\
&= \sum_{i=1}^n \sum_{j=1}^n (a_{ii}b_{jj} + \mathbb{1}_{\{i\neq j\}}2a_{ij}b_{ij})\mathbb{E}(x_i^2 x_j^2) - \mathbb{E}(\text{tr}(\boldsymbol{X}'\boldsymbol{AX}))\mathbb{E}(\text{tr}(\boldsymbol{X}'\boldsymbol{BX})) \\
&= \mathbb{E}(\boldsymbol{X}^{2\odot'}\boldsymbol{C}\boldsymbol{X}^{2\odot}) - \mathbb{E}(\text{tr}(\boldsymbol{AXX}'))\mathbb{E}(\text{tr}(\boldsymbol{BXX}')) \\
&= \text{tr}(\boldsymbol{C}\mathbb{E}(\boldsymbol{X}^{2\odot}\boldsymbol{X}^{2\odot'})) - \text{tr}(\boldsymbol{A}\mathbb{E}(\boldsymbol{XX}'))\text{tr}(\boldsymbol{B}\mathbb{E}(\boldsymbol{XX}')) \\
&= \text{tr}(\boldsymbol{C}(\boldsymbol{\Sigma}_{\boldsymbol{X}^{2\odot}} + \boldsymbol{\mu}_{\boldsymbol{X}^{2\odot}}\boldsymbol{\mu}'_{\boldsymbol{X}^{2\odot}})) - \text{tr}(\boldsymbol{A}\boldsymbol{\Sigma}_{\boldsymbol{X}})\text{tr}(\boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{X}}) \\
&= \text{tr}(\boldsymbol{C}\boldsymbol{\Sigma}_{\boldsymbol{X}^{2\odot}}) + \boldsymbol{\mu}'_{\boldsymbol{X}^{2\odot}}\boldsymbol{C}\boldsymbol{\mu}_{\boldsymbol{X}^{2\odot}} - \text{tr}(\boldsymbol{A}\boldsymbol{\Sigma}_{\boldsymbol{X}})\text{tr}(\boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{X}}) \qquad \square
\end{aligned}$$

A.3 Proof of Corollary 2

From Theorem 2 we get

$$\gamma_{\sigma^2_{(h),t}}(\ell) = \text{tr}(\boldsymbol{C}(\boldsymbol{\Sigma}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}} + \boldsymbol{\mu}_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}}\boldsymbol{\mu}'_{\boldsymbol{r}^{2\odot}_{t,h\Delta+\ell}})) - \text{tr}(\boldsymbol{KQK}'\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}})\text{tr}(\boldsymbol{LQL}'\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}}),$$

with $\boldsymbol{C} = \boldsymbol{ab}' + 2(\boldsymbol{KQK}') \odot (\boldsymbol{LQL}') \odot (\mathbf{1}_{h\Delta+\ell}\mathbf{1}'_{h\Delta+\ell} - \boldsymbol{I}_{h\Delta+\ell})$, where $\boldsymbol{a} = \text{diag}(\boldsymbol{KQK}')$ and $\boldsymbol{b} = \text{diag}(\boldsymbol{LQL}')$. By assumption $(r_t)_{t\in\mathbb{Z}}$ follows a weak white noise process (Definition 1) with

zero mean, which directly implies

$$\boldsymbol{\mu}_{\boldsymbol{r}_{t,h\Delta+\ell}^{2\odot}} = \mathbb{E}(\boldsymbol{r}_{t,h\Delta+\ell}^{2\odot}) = \mathrm{Var}(r_t)\mathbf{1}_{h\Delta+\ell} = \sigma^2\mathbf{1}_{h\Delta+\ell}$$

and

$$\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}} = \mathbb{E}(\boldsymbol{r}_{t,h\Delta+\ell}\boldsymbol{r}'_{t,h\Delta+\ell}) = \mathrm{Var}(r_t)\boldsymbol{I}_{h\Delta+\ell} = \sigma^2\boldsymbol{I}_{h\Delta+\ell}.$$

Plugging in gives

$$\gamma_{\sigma^2_{(h),t}}(\ell) = \mathrm{tr}\big(\boldsymbol{C}(\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}^{2\odot}} + \sigma^4\mathbf{1}_{h\Delta+\ell}\mathbf{1}'_{h\Delta+\ell})\big) - \sigma^4\mathrm{tr}(\boldsymbol{KQK}')\mathrm{tr}(\boldsymbol{LQL}')$$
$$= \mathrm{tr}\big(\boldsymbol{C}\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}^{2\odot}}\big) + \sigma^4\big(\mathbf{1}'_{h\Delta+\ell}\boldsymbol{C}\mathbf{1}_{h\Delta+\ell} - \mathrm{tr}(\boldsymbol{Q})^2\big).$$

Furthermore, it holds with $\boldsymbol{A} := \boldsymbol{KQK}'$ and $\boldsymbol{B} := \boldsymbol{LQL}'$

$$\mathbf{1}'_{h\Delta+\ell}\boldsymbol{C}\mathbf{1}_{h\Delta+\ell} = \mathbf{1}'_{h\Delta+\ell}(\boldsymbol{ab}' + 2\boldsymbol{A}\odot\boldsymbol{B}\odot(\mathbf{1}_{h\Delta+\ell}\mathbf{1}'_{h\Delta+\ell} - \boldsymbol{I}_{h\Delta+\ell}))\mathbf{1}_{h\Delta+\ell}$$
$$= \underbrace{\mathbf{1}'_{h\Delta+\ell}\boldsymbol{a}}_{=\mathrm{tr}(\boldsymbol{A})}\underbrace{\boldsymbol{b}'\mathbf{1}_{h\Delta+\ell}}_{=\mathrm{tr}(\boldsymbol{B})} + 2\mathbf{1}'_{h\Delta+\ell}(\boldsymbol{A}\odot\boldsymbol{B}\odot\mathbf{1}_{h\Delta+\ell}\mathbf{1}'_{h\Delta+\ell})\mathbf{1}_{h\Delta+\ell} - 2\mathbf{1}'_{h\Delta+\ell}(\boldsymbol{A}\odot\boldsymbol{B}\odot\boldsymbol{I}_{h\Delta+\ell})\mathbf{1}_{h\Delta+\ell}$$
$$= \mathrm{tr}(\boldsymbol{Q})^2 + 2\mathbf{1}'_{h\Delta+\ell}(\boldsymbol{A}\odot\boldsymbol{B})\mathbf{1}_{h\Delta+\ell} - 2\boldsymbol{a}'\boldsymbol{b}$$
$$= \mathrm{tr}(\boldsymbol{Q})^2 + 2\mathrm{tr}(\boldsymbol{AB}) - 2\boldsymbol{a}'\boldsymbol{b}$$
$$= \mathrm{tr}(\boldsymbol{Q})^2 + 2\mathrm{tr}(\boldsymbol{KQK}'\boldsymbol{LQL}') - 2\boldsymbol{a}'\boldsymbol{b}.$$

Implying

$$\gamma_{\sigma^2_{(h),t}}(\ell) = \mathrm{tr}\big(\boldsymbol{C}\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}^{2\odot}}\big) + \sigma^4\big(\mathrm{tr}(\boldsymbol{Q})^2 + 2\mathrm{tr}(\boldsymbol{KQK}'\boldsymbol{LQL}') - 2\boldsymbol{a}'\boldsymbol{b} - \mathrm{tr}(\boldsymbol{Q})^2\big)$$
$$= \mathrm{tr}\big(\boldsymbol{C}\boldsymbol{\Sigma}_{\boldsymbol{r}_{t,h\Delta+\ell}^{2\odot}}\big) + 2\sigma^4\big(\mathrm{tr}(\boldsymbol{KQK}'\boldsymbol{LQL}') - \boldsymbol{a}'\boldsymbol{b}\big). \qquad\square$$

A.4   GARCH(p,q) Fulfills the Conditions of Corollary 2

Let $(r_t)_{t\in\mathbb{Z}}$ be a GARCH(p,q) as defined in Example 2. We further assume that the first four moments of r_t exist and are finite.[16]

It is well known that GARCH processes are weak white noise process (Definition 1), so it remains to show that the moment conditions (8)-(11) of Theorem 2 are satisfied. The first moment condition (8) is obviously fulfilled for the zero-mean process $(r_t)_{t\in\mathbb{Z}}$.

---

[16] Conditions for the existence of moments can be found in He and Teräsvirta (1999b) and Bollerslev (1986) for the GARCH(1,1) model and for the GARCH(p,q) model in Ling and McAleer (2002a). The functional form of the moments are given in He and Teräsvirta (1999a) and Karanasos (1999).

Let $\mathcal{I}_\tau := \{r_s : s \leq \tau\}$. W.l.o.g. assume $t_1 < t_2 < t_3 < t_4$, then it holds

$$\mathbb{E}(r_{t_1}r_{t_2}r_{t_3}r_{t_4}) = \mathbb{E}(\mathbb{E}(r_{t_1}r_{t_2}r_{t_3}r_{t_4}|\mathcal{I}_{t_4-1})) = \mathbb{E}(r_{t_1}r_{t_2}r_{t_3}\sigma_{t_4}\mathbb{E}(\epsilon_{t_4}|\mathcal{I}_{t_4-1})) = 0,$$

which shows that moment condition (9) holds for GARCH(p,q) processes with symmetric innovation distributions and existing and finite fourth moments. If $t_1 < t_2$, we get

$$\mathbb{E}(r_{t_1}^3 r_{t_2}) = \mathbb{E}(\mathbb{E}(r_{t_1}^3 r_{t_2}|\mathcal{I}_{t_2-1})) = \mathbb{E}(r_{t_1}^3 \sigma_{t_2}\mathbb{E}(\epsilon_{t_2}|\mathcal{I}_{t_2-1})) = 0$$

and if $t_1 > t_2$ it follows

$$\mathbb{E}(r_{t_1}^3 r_{t_2}) = \mathbb{E}(\mathbb{E}(r_{t_1}^3 r_{t_2}|\mathcal{I}_{t_1-1})) = \mathbb{E}(r_{t_2}\sigma_{t_1}^3 \mathbb{E}(\epsilon_{t_1}^3|\mathcal{I}_{t_1-1})) = 0,$$

which shows that moment condition (11) holds for GARCH(p,q) processes with symmetric innovation distributions and existing and finite fourth moments. Let $t_1 < \max\{t_2, t_3\}$ and w.l.o.g. $t_3 > t_2$ then it follows

$$\mathbb{E}(r_{t_1}^2 r_{t_2}r_{t_3}) = \mathbb{E}(\mathbb{E}(r_{t_1}^2 r_{t_2}r_{t_3}|\mathcal{I}_{t_3-1})) = \mathbb{E}(r_{t_1}^2 r_{t_2}\sigma_{t_3}\mathbb{E}(\epsilon_{t_3}|\mathcal{I}_{t_3-1})) = 0.$$

If $t_1 > \max\{t_2, t_3\}$, $\mathbb{E}(\epsilon_t^2) = \sigma^2$ and w.l.o.g. $t_1 = t$, $t_2 = t-1$ and $t_3 = t-2$ it holds

$$\mathbb{E}(r_{t_1}^2 r_{t_2}r_{t_3}) = \mathbb{E}(r_t^2 r_{t-1}r_{t-2}) = \mathbb{E}(\mathbb{E}(r_t^2 r_{t-1}r_{t-2}|\mathcal{I}_{t-1})) = \mathbb{E}(r_{t-1}r_{t-2}\mathbb{E}(\sigma_t^2\epsilon_t^2|\mathcal{I}_{t-1}))$$

$$= \mathbb{E}(r_{t-1}r_{t-2}\sigma_t^2\mathbb{E}(\epsilon_t^2)) = \sigma^2\mathbb{E}(r_{t-1}r_{t-2}(\alpha_0 + \sum_{i=1}^{q}\alpha_i r_{t-i}^2 + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2))$$

$$= \sigma^2\alpha_0\mathbb{E}(r_{t-1}r_{t-2}) + \sigma^2\sum_{j=1}^{p}\beta_j\mathbb{E}(r_{t-1}r_{t-2}\sigma_{t-j}^2) + \sigma^2\alpha_1\mathbb{E}(r_{t-1}^3 r_{t-2}) + \sigma^2\sum_{i=2}^{q}\alpha_i\mathbb{E}(r_{t-1}r_{t-2}r_{t-i}^2)$$

$$= 0,$$

which shows that moment condition (10) holds for GARCH(p,q) processes with symmetric innovation distributions and existing and finite fourth moments.

B   Additional Results and Figures

B.1   Gaussian White Noise Process with the Sample Variance

Figure 13 and Figure 14 in this section of the appendix are analogously to Figure 4 and Figure 5 but with GARCH(1,1) (12) being replaced by the Gaussian white noise as data generating process and the EWMA variance (5) being being substituted by the sample variance (4).
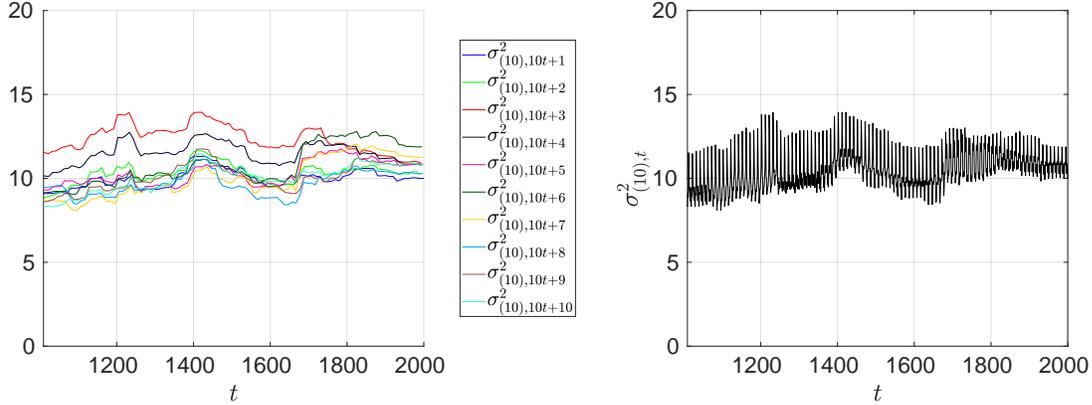
Figure 13: Time series of sample variance estimates (4), $\sigma^2_{(h),t}$, for simulated daily return series from the Gaussian white noise process with variance $\sigma^2 = 1$. The plot on the left shows the estimates $(\sigma^2_{(10),10t+\tau})_{t\in\mathbb{Z}}$, for $1 \le \tau \le 10$. The right plot shows the series $(\sigma^2_{(10),t})_{t\in\mathbb{Z}}$. Both plots are based on bi-weekly ($h = 10$) returns and estimation window $\Delta = 100$.
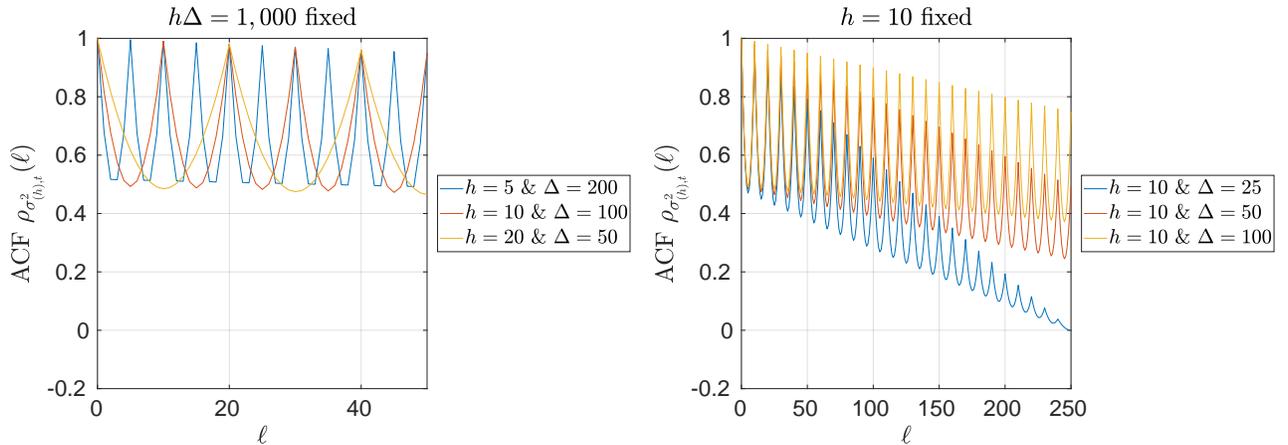
Figure 14: The ACF of sample variances (4), $\sigma^2_{(h),t}$, for daily returns from the Gaussian white noise process with $\sigma^2 = 1$. For the left plot we use a fixed number of daily returns to derive the sample variances. The right plot depicts the ACF of sample variances based on bi-weekly ($h = 10$) returns and estimation windows $\Delta = 25, 50, 100$.

B.2 The Functional Form and Amplitude of the Periodic Spurious Seasonality in the ACF

The autocovariance function of the quadratic-form variance estimator when the daily log-return process, $(r_t)_{t \in \mathbb{Z}}$, is a weak white noise process satisfying the moment conditions (8)-(11) is given in Corollary 2 for $\ell \geq 0$ by

$$\gamma_{\sigma^2_{(h),t}}(\ell) = \text{tr}\big(C\Sigma_{r^{2\odot}_{t,h\Delta+\ell}}\big) + 2\sigma^4\big(\text{tr}(KQK'LQL') - a'b\big),$$

with $C = ab' + 2(KQK') \odot (LQL') \odot (1_{h\Delta+\ell}1'_{h\Delta+\ell} - I_{h\Delta+\ell})$, where $a = \text{diag}(KQK')$ and $b = \text{diag}(LQL')$. The expression is nicely compact, but lacks intuition. It is not obvious, where the spurious seasonality is exactly coming from and how the amplitude of the periodic spurious seasonality in the ACF depends on the variance estimator and the data generating process. To provide more insight, we re-write the autocovariance as a sum of three components

$$\gamma_{\sigma^2_{(h),t}}(\ell) = s_1(Q) + s_2(Q) + s_3(Q),$$

with

$$s_1(Q) := b'\Sigma_{r^{2\odot}_{t,h\Delta+\ell}}a - 2(\sigma^4 + \gamma_{r^2_t}(0))a'b,$$
$$s_2(Q) := 2\text{tr}\big(((KQK') \odot (LQL'))\Sigma_{r^{2\odot}_{t,h\Delta+\ell}}\big),$$
$$s_3(Q) := 2\sigma^4 1'_{h\Delta+\ell}((KQK') \odot (LQL'))1_{h\Delta+\ell}.$$

The three terms depend on the variance estimator, defining Q, and the data generating process, which impacts σ^4, $\gamma_{r^2_t}(0)$ and $\Sigma_{r^{2\odot}_{t,h\Delta+\ell}}$. In the following we consider again the Gaussian white noise process with $\sigma^2 = 1$ and GARCH(1,1) process (12). As variance estimators, we study the sample variance and the EWMA variance based on non-overlapping h-day returns with $h = 10$. The window length is set to $\Delta = 100$.

Figure 15 shows the ACF of the estimated variances, $\rho_{\sigma^2_{(h),t}}(\ell)$, and the three components $s_i(\rho_{\sigma^2_{(h),t}}(\ell)) = s_i(Q_{(h),\Delta})/\gamma_{\sigma^2_{(h),t}}(0)$. Note that the components add up to the ACF, i.e., $\rho_{\sigma^2_{(h),t}}(\ell) = \sum_{i=1}^3 s_i(\rho_{\sigma^2_{(h),t}}(\ell))$. The two top rows correspond to the Gaussian white noise process with the sample variance in the first row and the EWMA variance in the second row. Accordingly, the third and fourth row show the results for the GARCH(1,1) process and the respective variance estimators. One can see that the first component $s_1(\rho_{\sigma^2_{(h),t}}(\ell))$ is not contributing to the periodic spurious seasonality effect and the functional form depends on the variance estimator and the data generating process. The second component, $s_2(\rho_{\sigma^2_{(h),t}}(\ell))$, is not periodic for the Gaussian white noise process and periodic for the GARCH(1,1) process. This is due to the fact that $\Sigma_{r^{2\odot}_{t,h\Delta+\ell}}$ is diagonal for Gaussian white noise but not for GARCH(1,1) processes since squared
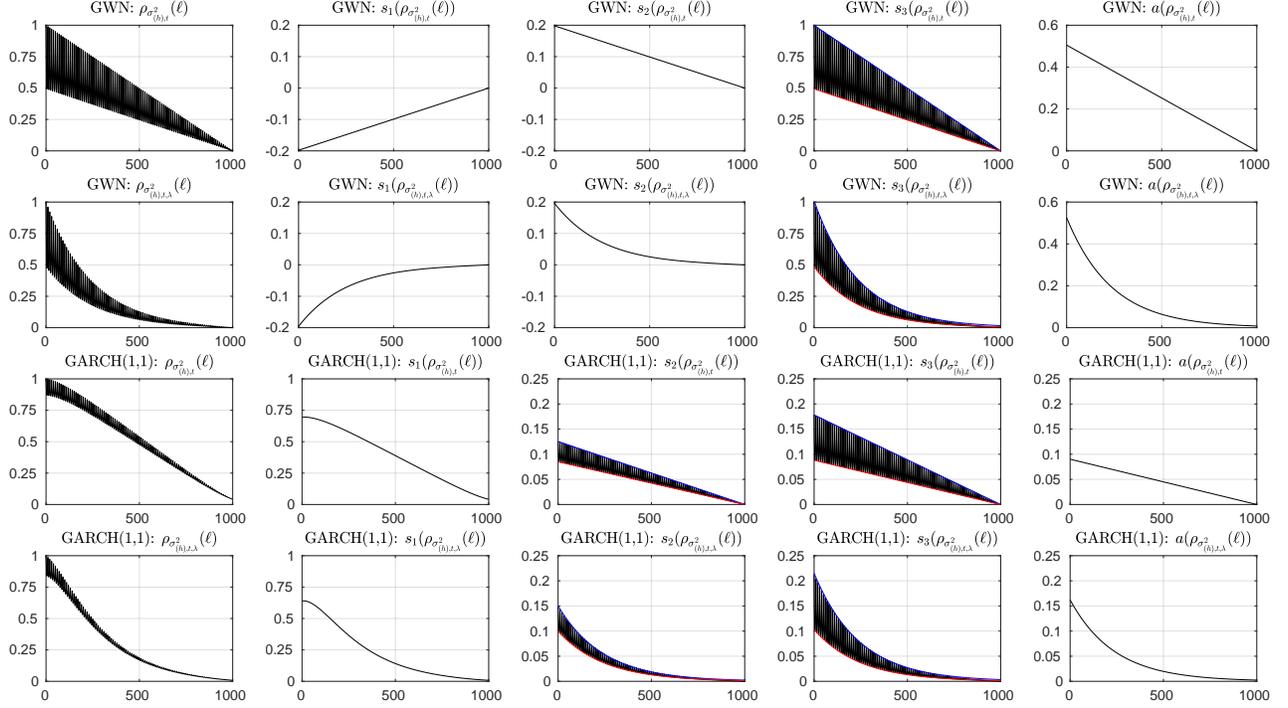
Figure 15: The ACF, its components $s_1(\cdot)$, $s_2(\cdot)$, $s_3(\cdot)$, and peak-to-peak amplitudes, $a(\cdot)$, of sample variances (first and third row), $\sigma^2_{(h),t}$, and EWMA variances (second and fourth row), $\sigma^2_{(h),t,\lambda}$, for different lags ℓ. The time-scale on the horizontal axis represents lags of ℓ days. In the first two rows, for the daily returns the Gaussian white noise process with $\sigma^2 = 1$ is assumed and in the third and fourth row GARCH(1,1) process (12). All plots are based on bi-weekly ($h = 10$) returns and estimation window $\Delta = 100$.

observations are autocorrelated. In case of the sample variance, the peak-to-peak amplitude in $s_2(\rho_{\sigma^2_{(h),t}}(\ell))$ is decreasing more slowly than for the EWMA variance estimator. The third term, $s_3(\rho_{\sigma^2_{(h),t}}(\ell))$, is periodic in all four cases and the functional form of the amplitude is comparable to those of the second component. The fifth column shows the contribution of the second and third component to the peak-to-peak amplitudes, denoted by $a(\rho_{\sigma^2_{(h),t}}(\ell))$.[17] One can see that the peak-to-peak amplitude of the sample variance is decreasing in ℓ, and it is larger for the Gaussian white noise process than for the GARCH(1,1) process. The peak-to-peak amplitudes for the EWMA variances are comparable to those of the sample variance for small values of ℓ, but the amplitudes are decreasing much faster in ℓ.

The crucial term in both periodic components, $s_2(\boldsymbol{Q})$ and $s_3(\boldsymbol{Q})$, is $(\boldsymbol{KQK'}) \odot (\boldsymbol{LQL'})$. The block-structure of $\boldsymbol{Q}_{(h),\Delta,\lambda}$ and $\boldsymbol{Q}_{(h),\Delta}$ (see, for example, the top-left plot in Figure 7) and the fact that $\boldsymbol{KQK'} = \text{blkDiag}(\boldsymbol{0}_{(\ell \times \ell)}, \boldsymbol{Q})$ and $\boldsymbol{LQL'} = \text{blkDiag}(\boldsymbol{Q}, \boldsymbol{0}_{(\ell \times \ell)})$ have a block-diagonal structure reveal how the periodicity of length h is generated, when different lags, ℓ, are considered and the Hadamard product of the matrices $\boldsymbol{KQK'}$ and $\boldsymbol{LQL'}$ is formed.

---

[17] The peak-to-peak amplitudes have been approximated by fitting linear functions (for the sample variance) and exponential functions (for the EWMA variance) through the peaks and taking the pointwise differences between the fitted functions. The fitted curves are shown in blue and red in the plots of $s_2(\cdot)$ and $s_3(\cdot)$.

B.3 Comparison of EWMA Variance Estimators

In this section of the appendix we present plots for the four different EWMA variance estimators: $\sigma^2_{(h),t,\lambda}$ (non-overlapping h-day returns, (5)), $\check{\sigma}^2_{(h),t,\lambda}$ (overlapping h-day returns, (14)), $\bar{\sigma}^2_{(h),t,\lambda}$ (two-scales, (16)), and $\tilde{\sigma}^2_{(h),t,\lambda}$ (corrected two-scales, (18)). In Figure 16, time series of variance estimates for simulated data from GARCH(1,1) process (12) are shown. The most left and right
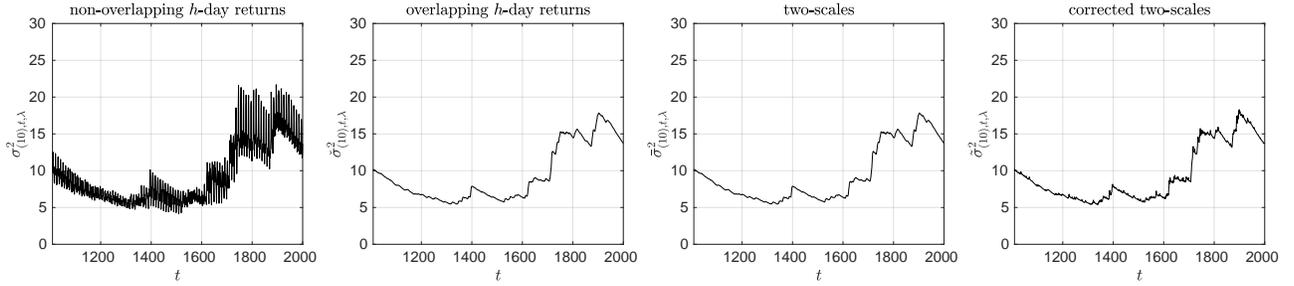
Figure 16: Time series of EWMA variance estimates (5), (14), (16), (18) for simulated daily return series from GARCH(1,1) process (12). The plots are based on bi-weekly ($h = 10$) returns and estimation window $\Delta = 100$.

plots, for $\sigma^2_{(h),t,\lambda}$ and $\tilde{\sigma}^2_{(h),t,\lambda}$, have already been shown on the right of Figure 4 and Figure 8, respectively. In Figure 17, the ACF for all four EWMA variances is plotted. Again, the most left and right plot, for $\sigma^2_{(h),t,\lambda}$ and $\tilde{\sigma}^2_{(h),t,\lambda}$, have already been shown on the right of Figure 5 and Figure 9, respectively.
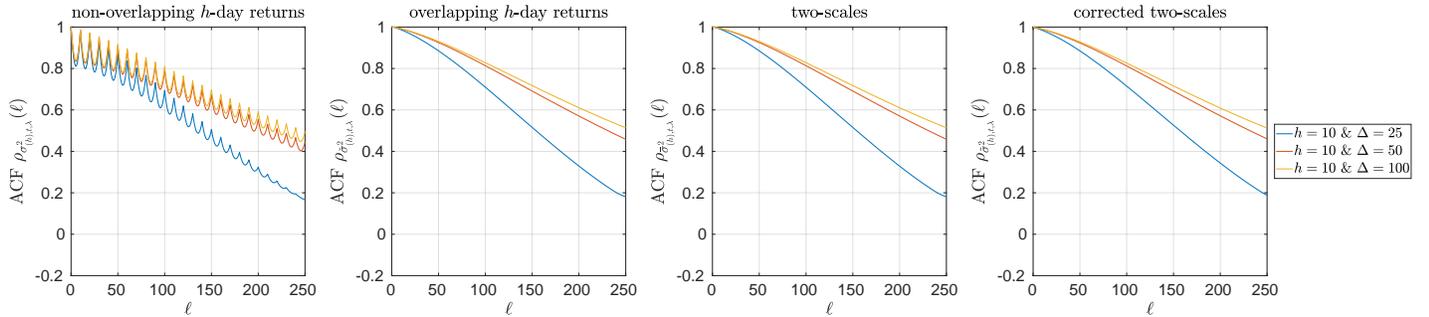
Figure 17: The ACF of EWMA variances (5), (14), (16), (18), for daily returns from GARCH(1,1) process (12). The ACF of EWMA variances is based on bi-weekly ($h = 10$) returns and estimation windows $\Delta = 25, 50, 100$.