Regular and Modified Kernel-Based Estimators of Integrated Variance: The Case with Independent Noise*

OLE E. BARNDORFF-NIELSEN

Department of Mathematical Sciences, University of Aarhus

PETER REINHARD HANSEN♣

Department of Economics, Stanford University

ASGER LUNDE

Department of Marketing, Informatics and Statistics, Aarhus School of Business

NEIL SHEPHARD

Nuffield College, University of Oxford

This Draft: November 15, 2004

Abstract

We consider kernel-based estimators of integrated variances in the presence of independent market microstructure effects. We derive the bias and variance properties for all regular kernel-based estimators and derive a lower bound for their asymptotic variance. Further we show that the subsample-based estimator is closely related to a Bartlett-type kernel estimator. The small difference between the two estimators due to end effects, turns out to be key for the consistency of the subsampling estimator. This observation leads us to a modified class of kernel-based estimators, which are also consistent. We study the efficiency of our new kernel-based procedure. We show that optimal modified kernel-based estimator converges to the integrated variance at rate $m^{1/4}$, where m is the number of intraday returns.

JEL Classification: C13, C22.

Keywords: Quadratic Variation, Market microstructure noise, Integrated variance, Kernel-based realized variance, Realized variance, Realized volatility.

---

*We thank Yacine Ait-Sahalia, Ulrich Müller, and Joe Romano for valuable comments.

♣Corresponding author: Peter Reinhard Hansen, Stanford University, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305-6072, USA, Phone: 650-725-1869, Email: peter.hansen@stanford.edu

1. Introduction

In the last five years substantial improvements in our understanding of and ability to forecast financial volatility has been possible through the harnessing of high frequency financial return data. The key developments have been the use of estimators of quadratic variation, (e.g. Andersen, Bollerslev, Diebold & Labys (2003) and Barndorff-Nielsen & Shephard (2002)) and making sense of their properties when applied to 5 to 30 minute return data. A weakness with existing methods is their inability to deal with market microstructure effects whose effects are key when we use returns recorded over very short time intervals. Interesting recent innovations that improve our comprehension of this topic include Aït-Sahalia, Mykland & Zhang (2003), Bandi & Russell (2004), Hansen & Lunde (2004*a*, 2004*c*), and Zhang, Mykland & Aït-Sahalia (2004).

The problem of estimating the quadratic variation is, in some ways, similar to the estimation of the long-run variance in stationary time-series. So it is not surprising that the literature has studied estimation methods that are well-known from the literature on covariance estimation, including pre-whitening methods, likelihood-based estimators, and kernel estimators. For example, the popular realized variance (*RV*) is analogous to the sum-of-squares variance estimator. Because the *RV* is sensitive to market microstructure noise it is recommended to use sparse sampling in practice, and the optimal sampling frequency is derived in Bandi & Russell (2004) and Zhang et al. (2004). The moving average filter used by Andersen, Bollerslev, Diebold & Ebens (2001) and the autoregressive filter used by Bollen & Inder (2002), are estimators that use pre-whitening techniques, and Bandi & Russell (2004) analyze optimal sampling of pre-whiten series. Likelihood-like estimators include the maximum likelihood estimators of Aït-Sahalia et al. (2003) who use a homogeneous diffusion model framework and the GMM estimator of Oomen (2004*b*) who use a pure jump model. The subsample estimator of Zhang, Mykland & Aït-Sahalia (2002) stands out as the only existing non-parametric estimator that is consistent, and its analog for estimation of the long-run variance was introduced by Carlstein (1986).

Our focus will be on kernel-based estimators. This literature was started by Zhou (1996) who proposed a particular kernel estimator, which only incorporates the first-order autocovariance. This suffices for unbiasness under "independence noise" where the population value of higher-order autocovariances are zero. Hansen & Lunde (2004*a*, 2004*c*) primarily use kernel-based estimators to characterize properties of market microstructure noise. Hansen & Lunde (2004*a*) use the estimator of Zhou (1996) to construct a test for "independent noise" and provide empirical evidence of time-dependence in the noise when return data are sampled at ultra high frequencies, such as every few ticks. Hansen & Lunde (2004*c*) analyze the properties of realized variance under general assumptions about the noise and derive a particular unbiased kernel estimator, that can be used to uncover the time-dependence in the noise. Thus, the existing literature on kernel estimators has either focused on that based on the first-order autocovariance, see Zhou (1996), or used particular unbiased kernels to analyze and characterize features of market microstructure noise, see Hansen & Lunde (2004*a*, 2004*c*).

In this paper we provide the first systematic study of kernel-based estimators of the integrated variance in the presence of market microstructure noise. We derive the optimal kernel-based estimator under an assumption that the noise is without memory and independent of the efficient price, an assumption which is empirically reasonable at moderate time scales such as 1-minute returns in highly liquid markets. Even though second and higher-order autocovariance are known to be zero under this assumption, we show that it pays off to estimate these. This makes it possible to derive kernel-based estimators that are far more precise than is that of Zhou (1996). However, we also show that there does not exist a consistent regular kernel-based estimator, so there is a limit to the precision of regular kernel-based estimators. Interestingly, we show that the consistent subsampling estimator of integrated variance by Zhang et al. (2004) is closely related to a particular kernel-based estimator. Importantly, it turns out that the difference between regular kernel estimators and the subsampling estimator, generated by *end effects*, is crucial for the consistency of the subsampling estimator. This observation allows us to propose a modified kernel-based estimator which is consistent. We study the efficiency of the new class of estimators and find its rate of convergence to be the optimal rate, $m^{1/4}$, where m is the number of intraday returns, see Gloter & Jacod (2001*a*, 2001b). So this rate is as good as the rate that can be obtained by a maximum likelihood estimator under more restrictive distributional assumptions for the noise.

In Section 2 we detail our assumptions about the noise, efficient price process and sampling scheme. In Section 3 we detail one of our main contributions, a systematic analysis of the properties of regular kernels. In Section 4 we related subsampling estimators to Bartlett-style regular kernels, and we see the difference is due to end conditions. In Section 5 we introduce the new modified kernel estimator and study its properties. In Section 6 we draw some conclusions. A lengthy Appendix provides the proofs of the results given in the paper.

2. Assumptions

2.1. *Price Process and Noise*

Without loss of generality we assume that the observed price process is given by

$$p(t) = p^*(t) + u(t), \qquad t \in [0, T], \tag{1}$$

where we label p^* as the efficient price process and u as the noise process. We assume that the efficient price is given from the simple diffusion model, $dp^*(t) = \sigma(t)dw(t)$, where $w(t)$ is a standard Wiener process that is independent of $\{\sigma^2(t)\}_{t=0}^{T}$, and we make the following assumptions about the noise process.

(**N**) The noise process u has mean zero, variance $\omega^2 \equiv E[u^2(t)] < \infty$, and kurtosis $\kappa \equiv E[u^4(t)]/\omega^4 < \infty$. Moreover, $u(s) \perp\!\!\!\perp p^*(t)$ for all $s, t \in [0, T]$ and $u(s) \perp\!\!\!\perp u(t)$ for all $s \neq t$.

There is plenty of empirical evidence against (**N**) when prices are sampled at ultra-high frequencies, such as every few ticks, see Hansen & Lunde (2004*a*, 2004*c*) who show that u is neither time-independent nor independent of p^*. On the other hand, Hansen & Lunde (2004*a*) also note that there is little evidence against (the implications of) (**N**) when prices are sampled at more moderate frequencies such as every 15 ticks. Because the analysis become much more complicated if u is time-dependent, all our results are derive using (**N**). So our results may not apply to tick-by-tick data. The advantage of our strategy is that it will produce a clear cut analysis of the core issues of kernel-based estimators.

Equation (1) may be viewed as a (Beveridge-Nelson type) decomposition, where p^* and u represent the persistent component and transitory component, respectively. So the volatility of $p(t+s) - p(t)$ is well approximated by that of $p^*(t+s) - p^*(t)$ when s is large. Thus, the volatility of p^* is the appropriate object of interest, even for the reader who is exclusively interested in the volatility of p (whether p is autocorrelated or not).

Without loss of generality we consider the unit interval of time, $[0, 1]$, that is divided into m sub-intervals $t_{i,m} - t_{i-1,m}$, $i = 1, \ldots, m$, ($t_{0,m} = 0$ and $t_{m,m} = 1$). The innovations in p^*, p, and u over each of the sub-intervals are defined by, for $i = 1, 2, \ldots, m$,

$$y^*_{i,m} \equiv p^*(t_{i,m}) - p^*(t_{i-1,m}), \quad y_{i,m} \equiv p(t_{i,m}) - p(t_{i-1,m}), \quad e_{i,m} \equiv u(t_{i,m}) - u(t_{i-1,m}).$$

We will refer to $y^*_{i,m}$ and $y_{i,m}$ as intraday returns, and we note that $y_{i,m} = y^*_{i,m} + e_{i,m}$.

We define the integrated variance

$$IV \equiv \int_0^1 \sigma^2(s)ds,$$

which is the object we would like to estimate. Our assumptions about the efficient price implies that $IV = \sum_{i=1}^m \sigma_{i,m}^2$, where $\sigma_{i,m}^2 \equiv \mathrm{var}(y_{i,m}^*)$, [1] $i = 1, \ldots, m$. In fact we have that $y_{1,m}^*, \ldots, y_{m,m}^*$ are independent and Gaussian distributed, $y_i^* \sim N(0, \sigma_{i,m}^2)$, (conditionally on $\{\sigma^2(s)\}_{s=0}^1$). Throughout we make the following assumptions about the volatility path.

(**V**) The volatility is (pathwise) continuous on [0, 1], strictly positive, and satisfies

$$m^{-1/2} \sum_{i=1}^m |\sigma^r(s_{i,m}) - \sigma^r(\tilde{s}_{i,m})| = o(1),$$

for some $r > 0$ (equivalently for all $r > 0$)[2] where $s_{i,m}$ and $\tilde{s}_{i,m}$ are arbitrary points in the interval $[t_{i-1,m}, t_{i,m}]$, $i = 1, \ldots, m$.

2.2. *Sampling Scheme*

We make the following assumption about the sampling times, $t_{0,m}, t_{1,m}, \ldots, t_{m,m}$, where we use $\lceil a \rceil$ to denote the smallest integer greater than or equal to a.

(**T**) It holds that $\sup_{s\in[0,1]} |t_{\lceil sm \rceil,m} - \tau(s)| = o(m^{-1})$, where τ is continuous and differentiable function, $\tau(0) = 0$ and $\tau(1) = 1$, and $0 < \tau'(s) < \infty$ for all $s \in [0, 1]$.

The special case where the price observations are equidistant in time, corresponds to $t_{i,m} = i/m$, in which case $\tau(s) = s$ and $\tau'(s) = 1$. Mykland & Zhang (2005) use a similar framework for sampling times, see also Barndorff-Nielsen & Shephard (2005). Given (**T**) we have the following result that corresponds to Assumption A.v in Mykland & Zhang (2005).

Lemma 1 *Given* (**T**) *it holds that*

$$\lim_{m\to\infty} \sup_{1\le i\le m} \left| \frac{t_{i,m} - t_{i-1,m}}{1/m} - \tau'(\tfrac{i}{m}) \right| = 0.$$

[1] All population moments are made conditional on the stochastic volatility process, $\{\sigma^2(s)\}_{s=0}^1$, which defines our object of interest. To simplify notation we use the convention $E(\cdot) \equiv E(\cdot|\{\sigma^2(s)\}_{s=0}^1)$, and similar for $\mathrm{var}(\cdot)$, and $\mathrm{cov}(\cdot)$.

[2] See Barndorff-Nielsen & Shephard (2003).

Also key for our analysis is the (time-deformed) integrated quarticity,

$$IQ \equiv \int_0^1 \tau'(s)\sigma^4(s)ds,$$

and it holds that $m\sum_{i=1}^m \sigma_{i,m}^4 = IQ + o(1)$, where $\sigma_{i,m}^4 \equiv (\sigma_{i,m}^2)^2$, see Lemma A.2 in the appendix.

An interesting sampling scheme is that where $\tau(s)$ is the solution to $\int_0^{\tau(s)} \sigma^2(r)dr = s \cdot IV$, such that $\sigma_{i,m}^2 = IV\big/m$ for all $i = 1, \ldots, m$. We refer to this as Business Time Sampling (BTS), see Oomen (2004*a*, 2004b). As noted by Hansen & Lunde (2004*a*), BTS minimize $IQ \equiv \int_0^1 \tau'(s)\sigma^4(s)ds = IV^2$, as the implicit function theorem shows that $\tau'(s) = IV\big/\sigma^2(s)$ under BTS.

(**T**′) Condition (**T**) holds with $\tau'(s) = IV\big/\sigma^2(s)$.

3. Properties of Regular Kernel-Based Estimators

We consider the family of *RV*-estimators $\{RV_{\mathbf{w}} : \mathbf{w} \in \mathbb{R}^m\}$ given by

$$RV_{\mathbf{w}} \equiv w_0\hat{\gamma}_0 + \sum_{h=1}^{m-1} w_h(\hat{\gamma}_{-h} + \hat{\gamma}_h), \qquad \text{where } \hat{\gamma}_{\pm h} \equiv \sum_{i=1}^{m-h} y_i y_{i+h} \text{ for } h = 0, \ldots, m-1,$$

and we call this the class of *regular kernels*. These types of statistics are familiar from the literature on covariance stationary processes, where they are used to estimate the long-run variances and covariances. Leading examples of this include Newey & West (1987) and Andrews (1991). This theory is not directly applicable here as our in-fill asymptotics is entirely different from the conventional setup. Further, the market microstructure noise in our problem will induce a particular autocovariance structure that we will use to characterize the kernels that provide good estimates of the *IV*.

Examples of kernel-based estimators for estimation of integrated variance from high-frequency data include those by Zhou (1996) ($\omega_h = 0$ for $h \geq 2$), Hansen & Lunde (2004*c*) ($\omega_h = (m+h)/m$ for $h \leq \lceil \rho m \rceil$ $0 \leq \rho < 1$), and Hansen & Lunde (2003, 2004*b*) (Bartlett kernel). Interestingly, we will show in Section 4 that the subsample-based estimator of Zhang et al. (2004) is almost identical to a Bartlett-type kernel estimator. However, the feature that makes the subsample estimator distinct from any kernel estimator turns out to be very informative about the estimation problem, and suggests a modified class of kernel estimators. We will spell this out in Section 5.

Since any kernel-based *RV* is a linear combination of $\hat{\gamma} \equiv (\hat{\gamma}_0, 2\hat{\gamma}_1, \ldots, 2\hat{\gamma}_{m-1})'$, we can study the properties of $RV_{\mathbf{w}}$ from the properties of $\hat{\gamma}$.

For any $m \times m$ matrix $\mathbf{A} = \{a_{ij}\}_{i,j=1}^m$ and any function, f, that is integrable on $[0, 1]$ we define the operator $f \mapsto \mathbf{I}(\mathbf{A}, f)$, which yields the $m \times m$ matrix with elements

$$\{\mathbf{I}(\mathbf{A}, f)\}_{i,j=1}^m \equiv \mathbf{A}_{ij} \int_0^1 \psi_{\iota_{ij}}(s) f(s) ds, \qquad \text{where } \iota_{ij} \equiv \tfrac{\max(i,j)-1}{m},$$

and

$$\psi_\rho(s) \equiv \begin{cases} 1 & \text{for } s \in [\rho, 1-\rho] \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

When $f(s) = c$ for all s, we write $\mathbf{I}(\mathbf{A}, c) \equiv \mathbf{I}(\mathbf{A}, f)$ and note that $\mathbf{I}(\mathbf{A}, c) = c\mathbf{I}(\mathbf{A}, 1)$ and that $\{\mathbf{I}(\mathbf{A}, c)\}_{i,j=1}^m = \mathbf{A}_{ij}(1 - \iota_{ij})c$.

Theorem 2 *Given* (**N**), (**V**) *and* (**T**), *then* $E(\hat{\gamma}') = (IV + 2m\omega^2, -(m-1)2\omega^2, 0, \ldots, 0)$ *and*

$$\text{cov}(\hat{\gamma}) = \mathbf{I}(\mathbf{A}, \omega^4)m - 2\omega^4\mathbf{C} + \omega^2\mathbf{I}(\mathbf{B}, \sigma^2) + \mathbf{I}(\mathbf{C}, \sigma^4\tau')\tfrac{1}{m} + \mathbf{H}o(\tfrac{1}{m}),$$

where the $m \times m$ matrices (assuming $\kappa = 3$) are given by

$$\mathbf{A} = \begin{pmatrix} 12 & -16 & 4 & 0 & \cdots \\ -16 & 28 & -16 & 4 & \ddots \\ 4 & -16 & 24 & -16 & \ddots \\ 0 & 4 & -16 & 24 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 8 & -8 & 0 & 0 & \cdots \\ -8 & 16 & -8 & 0 & \ddots \\ 0 & -8 & 16 & -8 & \ddots \\ 0 & 0 & -8 & 16 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

$$\mathbf{C} = \text{diag}(2, 4, 4, 4, \ldots), \qquad \mathbf{H} = \text{diag}(1, 1, 2, 3, 4, \ldots).$$

Remark 1 *The matrix* $\mathbf{H}$ *has a lower-right element of $m-1$, such that $\mathbf{H}o(\frac{1}{m})$ is not $o(\frac{1}{m})$. However, for the first q autocovariances, where q is a fixed number the reminder term for this submatrix of* $\text{cov}(\hat{\gamma})$ *is simply $o(\frac{1}{m})$, because all terms of this submatrix are at most $o(\frac{q}{m}) = o(\frac{1}{m})$. Later where we let $q = q_m \to \infty$ as $m \to \infty$, the last terms is $o(\frac{q_m}{m})$.*

Remark 2 *The variance simplifies considerably under* (**T**$'$) *where $IV^2 = IQ$, in which case we have that*

$$\text{cov}(\hat{\gamma}) = (\bar{\mathbf{A}}m - 2\mathbf{C})\omega^4 + \bar{\mathbf{B}}\omega^2 IV + \bar{\mathbf{C}}\tfrac{1}{m}IV^2,$$

where

$$\bar{\mathbf{A}} \equiv \mathbf{I}(\mathbf{A},1) = \begin{pmatrix} 12 & -16\frac{m-1}{m} & 4\frac{m-2}{m} & 0 & \cdots \\ -16\frac{m-1}{m} & 28\frac{m-1}{m} & -16\frac{m-2}{m} & 4\frac{m-3}{m} & \ddots \\ 4\frac{m-2}{m} & -16\frac{m-2}{m} & 24\frac{m-2}{m} & -16\frac{m-3}{m} & \ddots \\ 0 & 4\frac{m-3}{m} & -16\frac{m-3}{m} & 24\frac{m-3}{m} & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

and similar for $\bar{\mathbf{B}}$ *and* $\bar{\mathbf{C}}$*. Thus* $\bar{\mathbf{A}}_{ij} = (1-\iota_{ij})\mathbf{A}_{ij} = \frac{m-\max(i,j)+1}{m}\mathbf{A}_{ij}$ *for all* $i, j = 1, \ldots, m$.

Remark 3 *Theorem 2 is formulated for the case where* $\kappa = 3$*. The result for the general case where* κ *is arbitrary, requires the upper left* 2×2 *submatrix of* $\mathbf{A}$ *to be written as*

$$\begin{pmatrix} 4\kappa & -4(\kappa+1) \\ -4(\kappa+1) & 4(\kappa+4) \end{pmatrix},$$

whereas all other elements of $\mathbf{A}$ *are unchanged, see the proof of Theorem 2. Restricting our attention to the case where* $\kappa = 3$ *has no important implication for our analysis, because the bias properties require that* $\omega_0, \omega_1 \to 1$ *as* $m \to \infty$*, which eliminates the* κ*-terms in* A *(since* $4\kappa + 4(\kappa+4) - 8(\kappa+1) = 8$ *does not involve* κ*, see Hansen & Lunde (2004*a*)).*

Several results in the existing literature now follow as special cases of Theorem 2. If $\omega^2 = 0$ we have the result by Jacod (1994) and Barndorff-Nielsen & Shephard (2002) that $\text{var}(RV^{(m)}) = 2IQ\frac{1}{m} + o(\frac{1}{m})$, see also Jacod & Protter (1998). When $\omega^2 > 0$ we have the expressions $\text{bias}(RV^{(m)}) = 2m\omega^2$ and $\text{var}(RV^{(m)}) = 12m\omega^4 + O(1)$ by Bandi & Russell (2004) and Zhang et al. (2004). More generally we have the following result by Hansen & Lunde (2004*a*) that $\text{var}(RV^{(m)}) = (12m-4)\omega^4 + 8\omega^2 IV + 2IQ\frac{1}{m} + o(\frac{1}{m})$, and the result by Zhou (1996) that $\text{var}(RV^{(m)}_{\text{AC}_1}) = (8m-12)\omega^4 + 8\omega^2 IV + 6IQ\frac{1}{m} + o(\frac{1}{m})$, for $RV^{(m)}_{\text{AC}_1} \equiv \hat{\gamma}_0 + 2\hat{\gamma}_1$, which now follows from Theorem 2 as special cases.

The interesting aspect of Theorem 2 is that adding estimates of autocovariance terms (that have a population value that is known to be zero) can increase the precision whenever $\omega^2 > 0$. The following Corollary contains results for the cases where the second and third autocovariances are included, using weights that minimize the asymptotic variance. For notational convenience we define $v_\rho \equiv \int_0^\rho \sigma^2(s)ds + \int_{1-\rho}^1 \sigma^2(s)ds$ and we note that $v_{\frac{h}{m}} = \sigma_1^2 + \cdots + \sigma_h^2 + \sigma_{m-h+1}^2 + \cdots + \sigma_m^2$ for integers of h.

Corollary 1 *Define* $RV_{\mathrm{AC}_2}^{(m)} \equiv \hat{\gamma}_0 + 2\hat{\gamma}_1 + \hat{\gamma}_2$, $RV_{\mathrm{AC}_3}^{(m)} \equiv \hat{\gamma}_0 + 2\hat{\gamma}_1 + \frac{7}{5}\hat{\gamma}_2 + \frac{3}{5}\hat{\gamma}_3$. *Under the assumptions of Theorem 2 both estimators have bias of* $2\omega^2$ *while*

$$\begin{aligned}\mathrm{var}(RV_{\mathrm{AC}_2}^{(m)}) &= 2m\omega^4 + 4\omega^2 IV + 7IQ\tfrac{1}{m} + 2\omega^2(v_{\frac{2}{m}} + \omega^2) + o(\tfrac{1}{m}),\\ \mathrm{var}(RV_{\mathrm{AC}_3}^{(m)}) &= \tfrac{4}{5}m\omega^4 + \tfrac{68}{25}\omega^2 IV + \tfrac{208}{25}IQ\tfrac{1}{m} + 8\tfrac{21}{100}\omega^2 v_{\frac{2}{m}} + 8\tfrac{12}{100}\omega^2 v_{\frac{3}{m}} + \tfrac{98}{25}\omega^4 + o(\tfrac{1}{m}).\end{aligned}$$

Corollary 1 shows that by adding (a linear combination of) higher-order autocovariances can reduce the variance without affecting the bias (for m sufficiently large), as the higher-order terms (or linear combination of these) have a zero mean and are negatively correlated with $\hat{\gamma}_0 + 2\hat{\gamma}_1$, such that adding a proper linear combination will lead to a reduction of the total variance.

The linear combinations of the higher-order autocovariances that were included in Corollary 1, $1\hat{\gamma}_2$ and $\frac{7}{5}\hat{\gamma}_2 + \frac{3}{5}\hat{\gamma}_3$, where chosen in order to minimize the asymptotic variance that is of order $\omega^4 m$. This also led to a reduction of the variance term that is of order m^0 (from 8 to 4 and $\frac{68}{25}$ times $\omega^2 IV$ respectively), whereas the m^{-1}-variance term was increased, and the last observation highlights the need to study all terms in our analysis of kernel-based estimators.

For notational convenience we define $IV_\rho \equiv \int_0^1 \psi_\rho(s)\sigma^2(s)ds$ and $IQ_\rho \equiv \int_0^1 \psi_\rho(s)\sigma^4(s)ds$, and we note that $IV - IV_\rho = \frac{1}{2}v_\rho$, and that $IQ - IQ_\rho = O(\rho)$, such that $\frac{1}{m}(IQ - IQ_{\frac{h}{m}}) = O(\frac{h}{m^2}) = o(\frac{1}{m})$.

Corollary 2 *Let* $\mathbf{w} = (w_0, \ldots, w_{m-1})'$. *The bias of* $RV_{\mathbf{w}}$ *is given by*

$$\mathrm{bias}(RV_{\mathbf{w}}) = (w_0 - 1)IV + (w_0 - \tfrac{m-1}{m}w_1)2\omega^2 m = w'(IV\mathbf{d} + 2m\omega^2\mathbf{f}) - IV,$$

where $\mathbf{d} = (1, 0, \ldots, 0)'$ *and* $\mathbf{f} = (1, -\frac{m-1}{m}, 0, \ldots, 0)'$; *whereas the variance is given by*

$$\mathrm{var}(RV_{\mathbf{w}}) = V_1\omega^4 m + V_0\omega^2 + V_{-1}\tfrac{1}{m} + o(\tfrac{1}{m}),$$

where

$$\begin{aligned}V_1(\mathbf{w}) = 12w_0^2 &+ \tfrac{m-1}{m}w_1 4(7w_1 - 8w_0) + \sum_{j=2}^{m-1}\tfrac{m-j}{m}w_j 8(3w_j - 4w_{j-1} + w_{j-2})\\ &- \tfrac{4}{m}w_0^2 - \tfrac{8}{m}\sum_{j=1}^{m-1}w_j^2,\end{aligned}$$

$$V_0(\mathbf{w}) = 8IVw_0^2 + \sum_{j=1}^{m-1}16IV_{\frac{j}{m}}w_j(w_j - w_{j-1}), \quad \textit{and} \quad V_{-1}(\mathbf{w}) = 2IQw_0^2 + \sum_{j=1}^{m-1}4IQ_{\frac{j}{m}}w_j^2.$$

Thus, $V_1 = o(\frac{1}{m})$ is a necessary condition for the variance of $RV_{\mathbf{w}}$ to vanish, and $w_0 \to 1$ as $m \to \infty$ is clearly required for $RV_{\mathbf{w}}$ to be generally consistent for IV. While there are other requirements, such as $V_0 = o(1)$ and $V_{-1} = o(m)$, we shall initially focus on the requirement that

$V_1 = o(\frac{1}{m})$, which appears to be the most stringent requirement. For this reason, we seek the kernel that minimizes $V_1(\mathbf{w})$ subject to the constraint that $w_0 = 1$.

Theorem 3 (Variance Bound for Regular Kernel-Based Estimators) *It holds that*

$$\mathbf{w}^\star \equiv \arg\min_{\mathbf{w}\in\mathbb{R}^m} V_1(\mathbf{w}), \quad \textit{subject to } w_0 = 1,$$

is given by $\mathbf{w}^\star = (1, \mathbf{w}_2^{\star\prime})'$ *where* $\mathbf{w}_2^\star \equiv -\mathbf{M}_{22}^{-1}\mathbf{M}_{21}$ *and* $\mathbf{M}_{22}$ *and* $\mathbf{M}_{21}$ *are submatrices of*

$$\bar{\mathbf{A}}m - 2\mathbf{C} = \begin{pmatrix} M_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix},$$

with dimensions $(m-1)\times(m-1)$ *and* $(m-1)\times 1$, *respectively. Further, it holds that*

$$mV_1(\mathbf{w}^\star) = \mathbf{w}^{\star\prime}(\bar{\mathbf{A}}m - 2\mathbf{C})\mathbf{w}^\star \to 4, \qquad \textit{as } m \to \infty.$$

Theorem 3 shows that it is not possible to drive the variance of a regular kernel-based estimator to zero, as $m \to \infty$. The result shows that $4\omega^4$ is a lower bound for the asymptotic variance. So the existence of a consistent regular kernel-based estimator is ruled out.[3] While consistency is clearly important, it is worth noticing that the non-vanishing variance term, $4\omega^4$, is likely to be very small in practice. For example, Hansen & Lunde (2004*a*) estimate ω^4 to be of an order in the neighborhood of 10^{-8} for the stocks of the Dow Jones Industrial Average. Consistency is convenient because it justifies the δ-method, such that a central limit theorem (CLT) for $\log(\mathbf{w}'\hat{\gamma})$, say, follows from a CLT for $\mathbf{w}'\hat{\gamma}$. Naturally, if $4\omega^4$ is negligible relative to $\text{var}(\mathbf{w}'\hat{\gamma})$, the distortions from using the δ-method to approximate the distribution of $\log(\mathbf{w}'\hat{\gamma})$, say, will be extremely modest. Nevertheless, the mere existence of consistent estimator – the subsample estimator of Zhang et al. (2004) – does challenge the usefulness of regular kernel-based estimators. So in the following two sections we shall study the subsample-based estimator and a modified class of kernel-based estimators, where the latter is motivated by the relation between the subsample estimator and a particular kernel-based estimator. But first we evaluate how far we can push the precision of regular kernel-based estimators.

Theorem 3 provides a lower bound for the asymptotic variance of regular kernel-based estimators, derived from V_1. Since the variance also involves the terms, V_0 and V_{-1} it is unclear whether this bound can be obtained by any kernel. This question is addressed by the following Lemma that gives a simple example of a scheme for $\mathbf{w}$ which achieves the lower bound. This estimator is almost

[3] While consistency does not require the variance to vanish, consistency is indeed ruled out in the present setting, because r_k/k (to be defined later) does not vanish in probability.

identical to that introduced to this context by Hansen & Lunde (2003), and later applied by Hansen & Lunde (2004*b*).

Lemma 4 *Consider the Bartlett-type kernel, where the elements of* $\mathbf{w}_\mathrm{B}$ *are given by*

$$w_0 = \tfrac{m-1}{m}\tfrac{q-1}{q}, \quad w_j = \tfrac{q-j}{q} \quad \textit{for } j = 1, \ldots, q, \quad w_j = 0 \quad \textit{for } j > q,$$

where $w_0 = \frac{m-1}{m} w_1$ *in order to eliminate the bias. Given* $(\mathbf{N})$, $(\mathbf{V})$, *and* $(\mathbf{T}')$ *it holds that*

$$V_1 = 4\tfrac{1}{m} + O(\tfrac{1}{q^2}), \quad V_0 = O(\tfrac{q}{m}), \quad V_{-1} = O(\tfrac{q^2}{m}),$$

such that $\mathrm{var}(RV_{\mathbf{w}_\mathrm{B}}) = 4\omega^4 + O(\frac{m}{q^2}) + O(\frac{q}{m})$, *which tends to* $4\omega^4$ *provided that* $q/m \to 0$ *and* $q^2/m \to \infty$ *as* $q, m \to \infty$.

Since the Bartlett-type kernel in Lemma 4 achieves the lower bound, it is asymptotically efficient in the class of regular kernel estimators.

3.1. *Bias Eliminating Regular Kernels*

Lemma 5 *We define* $\lambda \equiv \omega^2/IV$,

$$\Sigma_\lambda \equiv (\bar{\mathbf{A}}m - 2\mathbf{C})\lambda^2 + \bar{\mathbf{B}}\lambda + \bar{\mathbf{C}}\tfrac{1}{m} \quad \textit{and} \quad \Xi_\lambda \equiv (\mathbf{d} + 2m\lambda\mathbf{f})(\mathbf{d} + 2m\lambda\mathbf{f})',$$

where $\mathbf{d}$ *and* $\mathbf{f}$ *where defined in Corollary 2. Under the assumptions of Theorem 2 and* $(\mathbf{T}')$, *we have that* $\mathrm{MSE}(RV_\mathbf{w})/IV^2 = \mathbf{w}'(\Sigma_\lambda + \Xi_\lambda)\mathbf{w} - 2\mathbf{w}'(\mathbf{d} + 2m\lambda\mathbf{f}) + 1$.

While Lemma 5 is useful in order to evaluate the MSE for a given kernel estimator, it does not constitute a useful way to define an optimal kernel, such as $\mathbf{w}^\star \equiv \arg\min_\mathbf{w} \mathrm{MSE}(RV_\mathbf{w}) = (\Sigma_\lambda + \Xi_\lambda)^{-1}(\mathbf{d} + 2m\lambda\mathbf{f})$, because such a kernel would be extremely sensitive to small variations in λ.[4] Instead we restrict attention to kernels for which $w_0 = \frac{m-1}{m} w_1$ and $w_0 \to 1$ as $m \to \infty$. These restrictions guarantees that the resulting estimator is asymptotically unbiased, as can we verified from $E(\hat{\gamma})$ that was stated in Theorem 2. Note that the Bartlett-type kernel in Lemma 4 satisfies this criterion. The reason that we do not impose the constraint $w_0 = 1$, is that the MSE may be reduced by allowing w_0 to be slightly smaller than one, (i.e. trading a small increase in the bias for a reduction of the variance).

[4] This issue can be understood by considering the kernel given by: $w_0 = IV/(IV + 2\omega^2 m) = 1(1 + 2\lambda m)$ and $w_h = 0$ for $h \geq 1$. For $\lambda m = 4.5$ we have $w_0 = 1/10$, which is unbiased if indeed $\lambda m = 4.5$, but can be severely biased for other values of λ.

We define the $m-1 \times 1$ vector, $\mathbf{v} = (v_1, \ldots, v_{m-1})' = \mathbf{D}\mathbf{w}$, where $\mathbf{D}$ is the $m-1 \times m$ matrix given by

$$\mathbf{D} = \begin{pmatrix} \frac{m}{m-1} & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \\ 0 & & 0 & 1 & \ddots \\ \vdots & & & \ddots & \ddots \end{pmatrix},$$

and solve the constrained optimization problem, $\min_{\mathbf{v}} \mathbf{v}'\mathbf{D}\Sigma_\lambda\mathbf{D}'\mathbf{v}$ s.t. $\tilde{w}_1 = 1$, using the same technique as in Theorem 3. Thus we determine $\mathbf{v}_2^* = -\mathbf{M}_{22}^{-1}\mathbf{M}_{21}$, where $\mathbf{M}_{22}$ and $\mathbf{M}_{21}$ are submatrices of $\mathbf{D}\Sigma_\lambda\mathbf{D}'$, and define the kernel $\mathbf{w}_\lambda^* = (\frac{m-1}{m}, 1, \mathbf{v}_2^{*\prime})'$.

FIGURE 1 ABOUT HERE

Elements of $\mathbf{w}^*$ plotted against $i/\sqrt{m}$ (x-axis: $[0, 2]$)

$m = 78,\ 390,\ 1560$ and $\lambda = 0.01,\ 0.001$

Although our kernel is derived under the independent noise assumption, we note that the kernel has some degree of robustness to mild time dependence in the noise process. Time dependence in the noise process causes higher-order covariances to have an expected value that is different from zero, since the kernel above has $w_i > 0$, for $i = 2, 3, \ldots$ it is somewhat capable of capturing this deviation from the indpendence assumption.

The rate at which the variance of $RV_{\mathbf{w}_\lambda^*}$ converges to $4\omega^4$ can be determined numerically from an ancillary regression and we find this rate to be $m^{-1/2}$. We describe the ancillary regressions towards the end of Section 5.

4. Subsample-Based Estimator

Zhang et al. (2004) have proposed a very stimulating subsample-based estimator of integrated variance. In an unpublished paper Müller (1993) also studied the use of subsampling to estimate the variability of financial prices. His motivation was the same as Zhang et al. (2004), but his analysis was much less formalized, so we will focus entirely on the contribution from Zhang et al. (2004). The subsample estimator can be constructed from the grid, $\mathcal{G} \equiv \{t_0, t_2, \ldots, t_m\}$, [5] and the (non-overlapping) subgrids,

$$\mathcal{G}_{k_j} = \{t_{j-1}, t_{j-1+k}, \ldots, t_{j-1+c_j k}\}, \qquad \text{for } j = 1, \ldots, k \quad \text{where } c_j \equiv \left\lfloor \tfrac{m-j+1}{k} \right\rfloor,$$

[5] In the following we will often suppress the subscript-m to simplify our expressions.

and $\lfloor a \rfloor$ denotes the largest integer that is smaller than or equal to a. So the subgrids are such that $\mathcal{G}_{k_i} \cap \mathcal{G}_{k_j} = \emptyset$ for $i \neq j$ and $\mathcal{G} = \bigcup_{j=1}^{k} \mathcal{G}_{k_j}$ for any $k \leq m$. For each subsample we can calculate the realized variance

$$RV_{\mathcal{G}_{k_j}} \equiv \sum_{t_i \in \mathcal{G}_{k_j}} y_{t_i,t_{i+k}}^2, \qquad \text{where } y_{t_i,t_{i+k}} \equiv p_{t_{i+k}} - p_{t_i},$$

with the convention that $y_{t_i,t_j} = 0$ if $j > m$. The k-subsampling estimator by Zhang et al. (2004) is given by

$$RV_{\text{sub}_k} = \tfrac{1}{k} \sum_{j=1}^{k} RV_{\mathcal{G}_{k_j}} - \tfrac{m-k+1}{mk} RV_{\mathcal{G}}.$$

Theorem 6 *It holds that*

$$RV_{\text{sub}_k} = (1 - \tfrac{m-k+1}{mk})\hat{\gamma}_0 + \sum_{h=1}^{k} \tfrac{k-h}{k}(\hat{\gamma}_{-h} + \hat{\gamma}_h) - \tfrac{1}{k} r_k,$$

where $r_1 \equiv 0$ *and* $r_k \equiv r_{k-1} + (y_1 + \cdots + y_{k-1})^2 + (y_{m-k+2} + \cdots + y_m)^2$ *for* $k \geq 2$.

It is very interesting that the subsample-based estimator is almost identical to the kernel-based estimator that employs the Bartlett-type kernel:

$$\mathbf{w}_{\text{sub}_k} = (1 - \tfrac{m-k+1}{mk}, \tfrac{k-1}{k}, \tfrac{k-2}{k}, \ldots, \tfrac{1}{k}, 0, \ldots, 0)'.$$

The difference is the presence of r_k.

Remark 4 *Theorem 6 provide a way to implement the subsampling estimator, as* RV_{sub_k} *(for any* k*) can be calculated from the empirical autocovariances and the recursive formula for* r_k.

Remark 5 *The close relationship between* RV_{sub_k} *and kernel-based estimators, stems from the fact that* $y_{t_i,t_{i+k}} = y_{i+1} + \cdots + y_{i+k}$*, such that* RV_{sub_k} *is simply a linear combination of cross products of intraday returns,* $y_{i,m} y_{j,m}$, $i, j = 1, \ldots, m$*, as is the case for all kernel-based estimators. That the subsample estimator is closely related to the Bartlett kernel is perhaps not too surprising, because Bartlett (1950) motivated the Bartlett kernel with the subsampling idea, see also Anderson (1971, p. 512) and Priestley (1981, pp. 439–440). Interestingly Politis, Romano & Wolf (1999) noted that the subsample estimator (of the long-run variance) of Carlstein (1986) is identical to both the moving block bootstrap estimator and the Jackknife estimator in this case, see Künsch (1989) and Liu & Singh (1992). Further, the term,* $\tfrac{1}{k} r_k$*, that makes* RV_{sub_k} *distinct from kernel-based estimators is related to the* end effects*, see e.g. Priestley (1981, p. 440).*

Remark 6 *The really surprising result of Theorem 6 is that $\frac{1}{k}r_k$, which is innocuous in the contest of conventional stationary time series, is indeed crucial for the consistency of RV_{sub_k}. Zhang et al. (2004) show that $\lim_{m\to\infty}\text{var}(RV_{\text{sub}_k})=0$ for a suitable choice of $k=k_m$. So $\frac{1}{k}r_k$ is responsible for the increased precision beyond the lower bound, $4\omega^4$, that we established for kernel-based estimators in Theorem 3. It is interesting here to note the results in Müller (2004) that shows that the most 'robust' quadratic estimator is not a kernel estimator.*

Lemma 7 *Given* (**N**) *and* (**V**) *it holds that*

$$\begin{aligned} E(r_k) &= \sum_{h=1}^{k-1} h(\sigma_h^2+\sigma_{m+1-h}^2)+4(k-1)\omega^2, \\ \text{var}(\tfrac{1}{k}r_k) &= 4\tfrac{k-1}{k}\omega^4+O(\tfrac{k}{m}), \\ \text{cov}(\tfrac{1}{k}r_k,\widehat{\gamma}') &\overset{(1)}{=} \omega^4(12\tfrac{k-1/3}{k},-8\tfrac{k-1}{k},0,\ldots,0). \end{aligned}$$

Here we have used $\overset{(1)}{=}$ to denote equality in terms of the ω^4-terms, while other terms that involve $\sigma_{i,m}^2$ and $\sigma_{i,m}^4$ are omitted as these are $O(m^{-1})$ and $O(m^{-2})$, respectively.

Lemma 7 shows that

$$\begin{aligned} \text{var}(RV_{\text{sub}_k}) &= \text{var}(RV_{\mathbf{w}_{\text{sub}_k}}-\tfrac{1}{k}r_k) \\ &\overset{(1)}{=} 4\omega^4+4\tfrac{k-1}{k}\omega^4-2\text{cov}(\tfrac{1}{k}r_k,\widehat{\gamma}')\mathbf{w} \\ &\to 4\omega^4+4\omega^4-2(12\omega^4-8\omega^4)=0 \qquad \text{as } k,m\to\infty, \end{aligned}$$

confirming that RV_{sub_k} is consistent whereas the Bartlett type estimator is inconsistent.

Another result that follows from Lemma 7 is that the bias of RV_{sub_k} is given by

$$\begin{aligned} \text{bias}(RV_{\text{sub}_k}) &= (1-\tfrac{m-k+1}{mk}-1)IV+(1-\tfrac{m-k+1}{mk}-\tfrac{m-1}{m}\tfrac{k-1}{k})2\omega^2 m-\tfrac{1}{k}E(r_k) \\ &= -\tfrac{m-k+1}{mk}IV-\tfrac{1}{k}\sum_{h=1}^{k-1}h(\sigma_h^2+\sigma_{m+1-h}^2), \end{aligned} \tag{2}$$

which can be verified to be of order $O(\frac{m+k^2}{mk})$. Thus $\text{bias}(RV_{\text{sub}_k})=o(1)$ if $k/m=o(1)$ as $k,m\to\infty$.

5. Modified Kernel-Based Estimators

Having understood the connection between a regular kernel estimator and subsampling and gained an appreciation of why subsampling is consistent, we are now in a position to modify the regular kernel-based estimator to inherit that property. Our hope is to deliver a consistent estimator which is reasonably efficient even in small samples.

For $h \geq 1$ we define

$$z_h \equiv y_h^2 + 2y_h(y_{h-1} + \cdots + y_1), \quad \text{and} \quad \tilde{z}_h \equiv y_{m-h+1}^2 + 2y_{m-h+1}(y_{m-h+2} + \cdots + y_m),$$

then it can be shown that

$$r_k = \sum_{j=1}^{k-1}(k-j)z_j + \sum_{j=1}^{k-1}(k-j)\tilde{z}_j,$$

(see the proof of Lemma 7) such that

$$\begin{aligned} RV_{\text{sub}_k} &= (1 - \tfrac{m-k+1}{mk})\hat{\gamma}_0 + \sum_{h=1}^{k-1} \tfrac{k-h}{k} 2\hat{\gamma}_h - \tfrac{1}{k} r_k \\ &= (1 - \tfrac{m-k+1}{mk})\hat{\gamma}_0 + \sum_{h=1}^{k-1} \tfrac{k-h}{k}(2\hat{\gamma}_h - z_h - \tilde{z}_h) = \mathbf{w}'_{\text{sub}_k}\tilde{\gamma}, \end{aligned}$$

where we use the vector of modified autocovariances estimators,

$$\tilde{\gamma} \equiv (\hat{\gamma}_0, 2\tilde{\gamma}_1, \ldots, 2\tilde{\gamma}_{m-1})', \qquad 2\tilde{\gamma}_h \equiv 2\hat{\gamma}_h - z_h - \tilde{z}_h, \quad \text{for } h \geq 1.$$

Thus inspired by the subsample estimator, we consider a *modified class of kernel estimators*, given by $\{\tilde{\mathbf{w}}'\tilde{\gamma} : \tilde{\mathbf{w}} \in \mathbb{R}^m\}$. This class of estimators contains at least one consistent estimator of IV. Theorem 8 gives the properties of the underlying $\tilde{\gamma}$.

Theorem 8 *Given* (**N**), (**V**) *and* (**T**$'$), *it holds that*

$$E(\tilde{\gamma}) = (IV + 2m\omega^2, -2(m+1)\omega^2 - \frac{2}{m}IV, -\frac{2}{m}IV, \ldots, -\frac{2}{m}IV)'$$

and

$$\text{cov}(\tilde{\gamma}) = \text{cov}(\hat{\gamma}) + \tilde{\mathbf{A}}\omega^4 + \tfrac{1}{m}\tilde{\mathbf{B}}\omega^2 IV + \tfrac{1}{m^2}\tilde{\mathbf{C}}IV^2,$$

where the upper left $q \times q$ *sub-matrices of* $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, *and* $\tilde{\mathbf{C}}$ *are given by*

$$\tilde{\mathbf{A}}_q = \begin{pmatrix} 0 & -20 & 8 & 0 & \cdots \\ -20 & 48 & -28 & 8 & \ddots \\ 8 & -28 & 40 & -28 & \ddots \\ 0 & 8 & -28 & 40 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}, \quad \tilde{\mathbf{B}}_q = \begin{pmatrix} 0 & -8(2) & 0 & 0 & \cdots \\ -8(2) & 16(2) & -8(3) & 0 & \ddots \\ 0 & -8(3) & 16(3) & -8(4) & \ddots \\ 0 & 0 & -8(4) & 16(4) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

$$\tilde{\mathbf{C}}_q = \begin{pmatrix} 0 & -4 & -4 & -4 & \cdots \\ -4 & 8(.5) & -8 & -8 & \ddots \\ -4 & -8 & 8(1.5) & -8 & \ddots \\ -4 & -8 & -8 & 8(2.5) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

With Theorem 8 in place it is now simple to determine the number of subsamples that minimizes the mean squared error (MSE).

Corollary 9 *Given the assumptions of 8, it holds that*

$$\mathrm{bias}(RV_{\mathrm{sub}_q}) = \mathbf{w}'_{\mathrm{sub}_q} E(\tilde{\gamma}) = -\tfrac{m+(q-1)^2}{mq} IV, \tag{3}$$

such that mean squared error of RV_{sub_q} *is given by*

$$\mathrm{MSE}(RV_{\mathrm{sub}_q})/IV^2 = \tilde{\mathbf{w}}'_{\mathrm{sub}_q} \tilde{\Sigma}^q_\lambda \tilde{\mathbf{w}}_{\mathrm{sub}_q} + [\tfrac{m+(q-1)^2}{mq}]^2,$$

where $\tilde{\Sigma}^q_\lambda \equiv \Sigma^q_\lambda + \tilde{\mathbf{A}}_q \lambda^2 + \tilde{\mathbf{B}}_q \lambda + \tilde{\mathbf{C}}_q \frac{1}{m}$, Σ^q_λ *is the upper left* $q \times q$ *submatrix of* Σ_λ, *and* $\tilde{\mathbf{w}}_{\mathrm{sub}_q} = (1 - \frac{m-q+1}{mq}, \frac{q-1}{q}, \frac{q-2}{q}, \ldots, \frac{1}{q})$.

We observe that (3) is equivalent to (2) given $(\mathbf{T}')$.

Next we seek the optimal unbiased estimator in this modified class of kernel-based estimators. We define the $q \times q+1$ matrix

$$\tilde{\mathbf{D}}_q \equiv \begin{pmatrix} 1 & \frac{m+1}{m} & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \ddots \\ \vdots & & & \ddots & \ddots \end{pmatrix}.$$

Now we solve the constrained optimization problem, $\min_{\tilde{\mathbf{v}}} \tilde{\mathbf{v}}' \tilde{\mathbf{D}}_q \tilde{\Sigma}^q_\lambda \tilde{\mathbf{D}}'_q \tilde{\mathbf{v}}$ subject to $\tilde{v}_1 = 1$, using the same technique as in Theorem 3. Thus we determine $\tilde{\mathbf{v}}^*_2 = -\mathbf{M}^{-1}_{22}\mathbf{M}_{21}$, where $\mathbf{M}_{22}$ and $\mathbf{M}_{21}$ are submatrices of $\tilde{\mathbf{D}}_q \tilde{\Sigma}^q_\lambda \tilde{\mathbf{D}}'_q$, and define the kernel $\tilde{\mathbf{w}}^*_\lambda = (1, \frac{m}{m+1}, \tilde{\mathbf{v}}^{*\prime}_2)'$.

FIGURE 2 ABOUT HERE

Elements of $\tilde{\mathbf{w}}^*_\lambda$ plotted against $i/\sqrt{m}$ (x-axis: [0, 5])

$m = 78, 390, 1560$ and $\lambda = 0.01, 0.001$

Truncation: $q = 4\sqrt{m}$

5.1. *Ancillary Regressions*

Our analytical (matrix) expressions for $\mathrm{var}(\mathbf{w}'\hat{\gamma})$ and $\mathrm{var}(\tilde{\mathbf{w}}'\tilde{\gamma})$ do not reveal their dependence on m in closed form. However, this dependence can be determined numerically by ancillary regressions.

For the regular kernel estimator we found that $\mathrm{var}(\mathbf{w}_\lambda^{*\prime}\hat{\gamma}) \to 4\omega^4$ as $m \to \infty$, and the rate at which the variance converges to the lower bound can be determined from the ancillary regression

$$\log(\mathbf{w}_\lambda^{*\prime}\Sigma_{\boldsymbol{\lambda}}\mathbf{w}_\lambda^* - 4\lambda^2) = \alpha + \beta \log m + \varepsilon_m, \qquad \text{for } m = m_{\min}, \ldots, m_{\max}.$$

Similarly for the modified kernel estimator and the subsampling estimator where $\log(\tilde{\mathbf{w}}_\lambda^{*\prime}\tilde{\Sigma}_\lambda^q\tilde{\mathbf{w}}_\lambda^*)$ and $\log(\tilde{\mathbf{w}}_{\mathrm{sub}_{q*}}^{\prime}\tilde{\Sigma}_\lambda^q\tilde{\mathbf{w}}_{\mathrm{sub}_{q*}})$ are the relevant dependent variables. For the latter $q^* = q^*(\lambda, m)$ denotes the number of subsamples that minimized the variance.

1. Let $Y_{m_i} = \log(\mathbf{w}_\lambda^{*\prime}\Sigma_{\boldsymbol{\lambda}}\mathbf{w}_\lambda^* - 4\lambda^2)$, $\log(\tilde{\mathbf{w}}_\lambda^{*\prime}\tilde{\Sigma}_\lambda^q\tilde{\mathbf{w}}_\lambda^*)$ (using truncation $4\sqrt{m}$) or $\log(\tilde{\mathbf{w}}_{\mathrm{sub}_{q*}}^{\prime}\tilde{\Sigma}_\lambda^q\tilde{\mathbf{w}}_{\mathrm{sub}_{q*}})$ (using optimal q).

   For $m = 10^3,\ 10^4,\ 10^5,\ 10^6$, run the regressions:

   $$Y_{m_i} = \alpha_m + \beta_m \log m_i + \varepsilon_{m_i}, \qquad \text{for } m_i = \tfrac{1}{4}m, \tfrac{1}{2}m, m, 2m, 4m,$$

   which yields $(\hat{\alpha}_m, \hat{\beta}_m)$.

2. By imposing $\beta = -1/2$ (or $\beta = -1/3$) reestimate α_m by

   $$\hat{\alpha}_m = \frac{1}{5}\sum_{i=1}^{5}(Y_{m_i} - \beta \log m_i),$$

TABLE 1 ABOUT HERE

Ancillary Regression Results:

One Panels for each of $RV_{\mathbf{w}_\lambda^*}$ $RV_{\tilde{\mathbf{w}}_\lambda^*}$ $RV_{\mathrm{sub}_{q*}}$.

Table 1 shows that $m^{1/4}(\tilde{\mathbf{w}}_\lambda^{*\prime}\tilde{\gamma} - IV)$ has an asymptotic variance that equals $\exp(\hat{\alpha}_\infty)IV^2$ under (T'). The results in the table is consistent with Zhang et al. (2004) who show that the subsampling estimator converges at the slower rate $m^{1/6}$, which corresponds to $\beta_\infty = -1/3$

FIGURE 3 ABOUT HERE

For $\lambda = 0.0001$ make a scatter plot of:

$\mathbf{w}_\lambda^{*\prime}\Sigma_{\boldsymbol{\lambda}}\mathbf{w}_\lambda^*$, $\tilde{\mathbf{w}}_\lambda^{*\prime}\tilde{\Sigma}_\lambda^q\tilde{\mathbf{w}}_\lambda^*$ and $\tilde{\mathbf{w}}_{\mathrm{sub}_{q*}}^{\prime}\tilde{\Sigma}_\lambda^q\tilde{\mathbf{w}}_{\mathrm{sub}_{q*}}$

against $m = 2^3, 2^4, \ldots, 2^{21}$, in log-log scale

Later we might add the lines:

$\exp(\hat{\alpha}_{\mathrm{reg},\infty})m^{-1/2} + 4\lambda^2$, $\exp(\hat{\alpha}_{\mathrm{mod},\infty})m^{-1/2}$, and $\exp(\hat{\alpha}_{\mathrm{sub},\infty})m^{-1/3}$

We use Theorem 8 to determine the number of subsamples that minimized the variance, and define

$$q^*(\lambda, m) \equiv \arg\min_q \mathrm{var}(RV_{\mathrm{sub}_q}/IV^2).$$

Zhang et al. (2004) show that $q^*(\lambda, m) \propto (\lambda m)^{2/3}$, and in an unrestricted ancillary regression of $\log q^*(\lambda, m)$ on a range of values for $\log\lambda$ and $\log m$, we find that $q^*(\lambda, m) \simeq a * (\lambda m)^{2/3}$.

FIGURE ABOUT HERE

SCATTER PLOT of $q^*(\lambda, m)$ against m using log-log scale.

Plot the $3 \times 5 = 15$ data points resulting from combining pairs of $\lambda = 10^{-2}$, 10^{-3}, and 10^{-4} and $m = 10^2, 10^3, 10^4, 10^5$, and 10^6.

add the three lines $\log a + \frac{2}{3}\log\lambda + \frac{2}{3}\log m$, using the 3 values of λ.

5.2. *Maximum Likelihood Estimator of Integrated Variance*

We now compare the rate of convergence of the modified kernel estimator to the rate that is achieved by a maximum likelihood estimator of IV. So we consider a simple framework where the noise is assumed to be iid and Gaussian distributed, i.e. $u_i \sim N(0, \omega^2)$. Given ($\mathbf{T}'$) it now follows that

$$(y_i, \ldots, y_m)' \sim N_m\left(0, \Sigma_{IV,\omega^2}\right),$$

where the $m \times m$ covariance matrix, , is given by

$$\Sigma_{IV,\omega^2} \equiv \begin{pmatrix} \frac{IV}{m} & 0 & 0 & \cdots \\ 0 & \frac{IV}{m} & 0 & \\ 0 & 0 & \frac{IV}{m} & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix} + \begin{pmatrix} 2\omega^2 & -\omega^2 & 0 & \cdots \\ -\omega^2 & 2\omega^2 & -\omega^2 & \ddots \\ 0 & -\omega^2 & 2\omega^2 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Let $\hat{\sigma}^2_{\mathrm{ML}}$ and $\hat{\omega}^2_{\mathrm{ML}}$ denote the maximum likelihood estimators of IV and ω^2, respectively. The asymptotic properties of $\hat{\sigma}^2_{\mathrm{ML}}$ and $\hat{\omega}^2_{\mathrm{ML}}$ are given from classical results about the MA(1) process.[6]

[6]Setting $IV = 0$ takes the root of the underlying MA(1) process to -1. So for the interesting case with $IV > 0$, the local-to-zero of IV/m leads to a local-to -1 root, as analyzed by Anderson & Takemura (1986), Tanaka & Satchell (1989), and Shephard (1993). However, IV/m is sufficiently "non-local" to zero that it does not affect the limiting (Gaussain) distribution of the maximum likelihood estimators.

By adopting the expression given in Aït-Sahalia et al. (2003, proposition 1) to our notation, we have that asymptotic covariance matrix for $(\hat{\sigma}^2_{\text{ML}}, \hat{\omega}^2_{\text{ML}})$ is given by

$$\frac{IV^2}{m^2}\begin{pmatrix} 2m + 4m\sqrt{4\lambda m + 1} & -\left(2m\lambda + 1 + \sqrt{4m\lambda + 1}\right) \\ \bullet & \frac{1}{2m}(2m\lambda + 1)\left(2m\lambda + 1 + \sqrt{4m\lambda + 1}\right) \end{pmatrix}.$$

So for $\lambda > 0$ we have

$$\text{avar}\begin{pmatrix} m^{1/4}\hat{\sigma}^2_{\text{ML}} \\ m^{1/2}\hat{\omega}^2_{\text{ML}} \end{pmatrix} = IV^2\begin{pmatrix} 8\sqrt{\lambda} & 0 \\ 0 & 2\lambda^2 \end{pmatrix},$$

where avar$(\cdot)$ denotes the asymptotic covariance matrix. This shows that the maximum likelihood estimator of *IV* converges at the same rate, $m^{1/4}$, as the modified kernel estimator, which indeed has been show to be the optimal rate of convergence in this context, see Gloter & Jacod (2001*a*, 2001b). Further $\hat{\omega}^2_{\text{ML}}$ converges at the faster rate, $m^{1/2}$, and since there limit distribution is Gaussian, see e.g. Aït-Sahalia et al. (2003), we note that the two estimators are asymptotically independent.

The special case where there is no market microstructure noise, $(\lambda = 0)$ results in faster rates of convergence. Specifically we find that,

$$\text{avar}\begin{pmatrix} m^{1/2}\hat{\sigma}^2_{\text{ML}} \\ m^{3/2}\hat{\omega}^2_{\text{ML}} \end{pmatrix} = IV^2\begin{pmatrix} 6 & -2 \\ -2 & 1 \end{pmatrix},$$

and it is interesting to note that $\text{avar}(m^{1/2}\hat{\sigma}^2_{\text{ML}}) = 6IV^2$. So the asymptotic variance of $\hat{\sigma}^2_{\text{ML}}$ is in this case three times that of the realized variance, which is the constrained $(\lambda = 0)$ maximum likelihood estimator. Thus the loss in estimating the nuisance parameter ω^2, when it is truly zero, is identical to that of $RV^{(m)}_{\text{AC}_1} \equiv \hat{\gamma}_0 + 2\hat{\gamma}_1$, which also has $\text{var}(RV^{(m)}_{\text{AC}_1}) = 6IV^2\frac{1}{m} + o(\frac{1}{m})$ when $\omega^2 = 0$, see Zhou (1996).

6. Practical Implementation

In practice λ is not known, however it is straightforward to estimate ω^2. Combining results of Theorem 2 concerning $RV \equiv \hat{\gamma}_0$ and our results for $RV_{\tilde{\mathbf{w}}} \equiv \tilde{\mathbf{w}}'\tilde{\gamma}$ shows that

$$\hat{\omega}^2 \equiv \frac{RV - RV_{\tilde{\mathbf{w}}}}{2m} \xrightarrow{p} \omega^2,$$

since $E(RV) = IV + 2\omega^2 m$, $\text{var}(RV) = O(m)$ and $RV_{\tilde{\mathbf{w}}} \xrightarrow{p} IV$. Given the consistency of $RV_{\tilde{\mathbf{w}}}$ it follows that

$$\hat{\lambda} \equiv \frac{RV - RV_{\tilde{\mathbf{w}}}}{2mRV_{\tilde{\mathbf{w}}}} = \frac{\hat{\gamma}_0 - \tilde{\mathbf{w}}'\tilde{\gamma}}{2m \cdot \tilde{\mathbf{w}}'\tilde{\gamma}} \xrightarrow{p} \lambda.$$

This leads to a two-step estimator of integrated variance.

1. Given some initial value for λ (λ^o say), we construct $\tilde{\mathbf{w}}^*_{\lambda^o}$, and estimate

$$\hat{\lambda} = \hat{\lambda}(\lambda^o) \equiv \max\left(\frac{\hat{\gamma}_0 - \tilde{\mathbf{w}}^{*\prime}_{\lambda^o}\tilde{\gamma}}{2m \cdot \tilde{\mathbf{w}}^{*\prime}_{\lambda^o}\tilde{\gamma}}, 0\right).$$

2. Given $\hat{\lambda}$ we determine $\tilde{\mathbf{w}}^*_{\hat{\lambda}}$, and define our two-step estimator of IV to be:

$$RV_{\tilde{\mathbf{w}}^*_{\hat{\lambda}}} \equiv \tilde{\mathbf{w}}^{*\prime}_{\hat{\lambda}}\tilde{\gamma}.$$

Naturally this procedure could be iterated, increasing the precision of our estimate of λ.

What is the MSE loss of this procedure, compared to knowing the true value of λ? [Simulation Study to be added].

7. Conclusion

We have provided a systematic analysis of regular kernel-based estimators under the assumption that market microstructure noise is independent of the efficient prices and independent of itself (at different points in time). While this assumption is reasonable when prices are not sampled too frequently, such as every 15 ticks or so, there is overwhelming evidence that market microstructure noise has a more sophisticated dependence structure when sampling occurs at ultra-high frequencies, such as every tick. We are therefore, in separate papers, extending our analysis of kernel-based estimators to the situation with more general assumptions about the noise process.

We have showed that regular kernel-based estimators can be quite accurate estimators of quadratic variation, however they are always inconsistent. Taking inspiration from the consistent subsampling estimator, a new modified kernel estimator is suggested which is consistent and has good finite sample properties.

A. Proof of Theorem 2 and Related Results

Proof of Lemma 1. First we note that $t_{\lceil sm\rceil-1,m} = t_{\lceil (s-\frac{1}{m})m\rceil,m}$ and by (**T**) we have that

$$\sup_{s\in[0,1]}\left|(t_{\lceil sm\rceil,m} - t_{\lceil sm\rceil-1,m}) - (\tau(s) - \tau(s-\tfrac{1}{m}))\right| = o(m^{-1}),$$

such that

$$\sup_{s\in[0,1]}\left|\tfrac{t_{\lceil sm\rceil,m}-t_{\lceil sm\rceil-1,m}}{1/m} - \tau'(s)\right| \leq \sup_{s\in[0,1]}\left|\tfrac{t_{\lceil sm\rceil,m}-t_{\lceil sm\rceil-1,m}}{1/m} - \tfrac{\tau(s)-\tau(s-\frac{1}{m})}{1/m}\right| + \sup_{s\in[0,1]}\left|\tau'(s) - \tfrac{\tau(s)-\tau(s-\frac{1}{m})}{1/m}\right| = o(1) + o(1), \qquad \text{(A.1)}$$

where the last term is $o(1)$ since τ' is bounded. (A.1) clearly implies the result stated in the Lemma, (the two are equivalent given the continuity of $\tau'(s)$). ■

Lemma A.1 *We define $x_{i,h} \equiv y_i y_{i+h}$. Given* (**N**) *and* (**V**) *we have that*

Part I	$E(x_{i,h})$	$\text{var}(x_{i,h})$	$\text{cov}(x_{i,h}, x_{i\pm 1,h})$
$h = 0$	$\sigma_i^2 + 2\omega^2$	$(2\kappa + 2)\omega^4 + 8\omega^2\sigma_i^2 + 2\sigma_i^4$	$(\kappa - 1)\omega^4$
$h = 1$	$-\omega^2$	$(\kappa + 2)\omega^4 + 2\omega^2(\sigma_i^2 + \sigma_{i+1}^2) + \sigma_i^2\sigma_{i+1}^2$	ω^4
$h \geq 2$	0	$4\omega^4 + 2\omega^2(\sigma_i^2 + \sigma_{i+h}^2) + \sigma_i^2\sigma_{i+h}^2$	ω^4

while $\text{cov}(x_{i,h}, x_{i\pm k,h}) = 0, \quad k \geq 2$ *for all* $h = 0, 1, \ldots$.

Part II	$\text{cov}(x_{i,h}, x_{i,h+1})$	$\text{cov}(x_{i,h}, x_{i-1,h+1})$	$\text{cov}(x_{i,h}, x_{i-1,h+2})$
$h = 0$	$-(\kappa + 1)\omega^4 - 2\omega^2\sigma_i^2$	$-(\kappa + 1)\omega^4 - 2\omega^2\sigma_i^2$	$2\omega^4$
$h \geq 1$	$-2\omega^4 - \omega^2\sigma_i^2$	$-2\omega^4 - \omega^2\sigma_{i+h}^2$	ω^4,

while all other covariance terms are zero.

Proof. (Part I) The expected values are given from

$$E(x_{i,h}) = E(y_i y_{i+h}) = E(y_i^* + u_i - u_{i-1})(y_{i+h}^* + u_{i+h} - u_{i+h-1}),$$

which shows that $E(x_{i,0}) = E(y_i^{*2}) + E(u_i)^2 + E(u_{i-1}^2) = \sigma_i^2 + 2\omega^2$, since y_i^*, u_i, and u_{i-1} are pairwise uncorrelated. Similarly we find that $E(x_{i,1}) = E[(u_i)(-u_i)] + 0 = -\omega^2$ and that $E(x_{i,h}) = 0$ for $h \geq 2$.

Next, we turn to the variance and covariance terms, where we make use of the identities, $\text{var}(e_i) = E(e_i^2) = 2\omega^2$ and

$$E(e_i^4) = E[u_i^4 + u_{i-1}^4 + 6u_i^2 u_{i-1}^2 - 4u_i u_{i-1}^3 - 4u_i^3 u_{i-1}] = (2\kappa + 6)\omega^4.$$

For $h = 0$ we have

$$E(x_{i,0}^2) = E(y_i^4) = E(y_i^* + e_i)^4 = E(y_i^{*4}) + E(e_i^4) + 6E(y_i^{*2} e_i^2) = 3\sigma_i^4 + (2\kappa + 6)\omega^4 + 6\sigma_i^2 2\omega^2,$$

and

$$\begin{aligned}
E(x_{i,0} x_{i+1,0}) &= E(y_i^2 y_{i+1}^2) = E(y_i^* + e_i)^2 (y_{i+1}^* + e_{i+1})^2 \\
&= E[(y_i^{*2} + e_i^2 + 2y_i^* e_i)(y_{i+1}^{*2} + e_{i+1}^2 + 2y_{i+1}^* e_{i+1})] \\
&= E[(y_i^{*2} + e_i^2)(y_{i+1}^{*2} + e_{i+1}^2)] = E[(y_i^{*2} + u_i^2 + u_{i-1}^2)(y_{i+1}^{*2} + u_{i+1}^2 + u_i^2)]
\end{aligned}$$

$$= \sigma_i^2\sigma_{i+1}^2 + (\sigma_i^2 + \sigma_{i+1}^2)2\omega^2 + (\kappa+3)\omega^4, \tag{A.2}$$

$$E(x_{i,0}x_{i+h,0}) = E(y_i^2 y_{i+h}^2) = \sigma_i^2\sigma_{i+h}^2 + (\sigma_i^2 + \sigma_{i+h}^2)2\omega^2 + 4\omega^4, \quad \text{for } h \geq 2, \tag{A.3}$$

such that

$$\begin{aligned} \text{var}(x_{i,0}) &= E(x_{i,0}^2) - [E(x_{i,0})]^2 = [3\sigma_i^4 + (2\kappa+6)\omega^4 + 12\sigma_i^2\omega^2] - [\sigma_i^2 + 2\omega^2]^2 \\ &= 2\sigma_i^4 + (2\kappa+2)\omega^4 + 8\sigma_i^2\omega^2, \\ \text{cov}(x_{i,0}, x_{i+1,0}) &= (\kappa-1)\omega^4, \quad \text{and} \quad \text{cov}(x_{i,0}, x_{i+h,0}) = 0 \quad \text{for } h \geq 2. \end{aligned}$$

For $h = 1$ we find $E(x_{i,1}^2) = E(y_i^2 y_{i+1}^2) = E(x_{i,0}x_{i+1,0})$ which is derived in (A.2),

$$\begin{aligned} E(x_{i,1}x_{i+1,1}) &= E(y_i^* + u_i - u_{i-1})(y_{i+1}^* + u_{i+1} - u_i)^2(y_{i+2}^* + u_{i+2} - u_{i+1}) \\ &= E[(u_i)(-2u_{i+1}u_i)(-u_{i+1})] = 2E[u_i u_{i+1} u_i u_{i+1}] = 2\omega^4, \end{aligned}$$

and $E(x_{i,1}x_{i+2,1}) = \omega^4$. Since $E(x_{i,1})E(x_{j,1}) = (-\omega^2)(-\omega^2) = \omega^4$ for all $i, j = 1, \ldots, m$, we find that

$$\begin{aligned} \text{var}(x_{i,1}) &= \sigma_i^2\sigma_{i+1}^2 + (\sigma_i^2 + \sigma_{i+1}^2)2\omega^2 + (\kappa+3)\omega^4 - (\omega^2)^2 \\ &= \sigma_i^2\sigma_{i+1}^2 + (\sigma_i^2 + \sigma_{i+1}^2)2\omega^2 + (\kappa+2)\omega^4, \end{aligned}$$

and $\text{cov}(x_{i,1}, x_{i+1,1}) = \omega^4$ and $\text{cov}(x_{i,1}, x_{i+h,1}) = 0$, for $h \geq 2$.

For $h \geq 2$ we have $E(x_{i,h}^2) = E(y_i^2 y_{i+h}^2) = E(x_{i,0}x_{i+h,0})$ which is derived in (A.3), such that $\text{var}(x_{i,h}) = \sigma_i^2\sigma_{i+h}^2 + 2(\sigma_i^2 + \sigma_{i+h}^2)\omega^2 + 4\omega^4$. Next, we have that

$$E(x_{i,h}x_{i+1,h}) = E(e_i e_{i+h} e_{i+1} e_{i+1+h}) = \omega^4,$$

while $E(x_{i,h}x_{i+k,h}) = 0$ for $k \geq 2$. So $\text{cov}(x_{i,h}, x_{i\pm1,h}) = \omega^4$ and $\text{cov}(x_{i,h}, x_{i\pm k,h}) = 0$ for $k \geq 2$.

(Part II) We consider

$$\begin{aligned} E(x_{i,0}x_{i,1}) &= E(y_i^2 y_i y_{i+1}) = E[(y_i^* + e_i)^3(y_{i+1}^* + e_{i+1})] \\ &= E[(y_i^{*2} + 2y_i^* e_i + e_i^2)(y_i^* + e_i)(-u_i)] \\ &= E[(y_i^{*2} + 2y_i^* e_i + e_i^2)(y_i^* + u_i - u_{i-1})(-u_i)] \\ &= -\sigma_i^2\omega^2 - 2\sigma_i^2\omega^2 + E[e_i^2(u_i - u_{i-1})(-u_i)] \\ &= -\sigma_i^2\omega^2 - 2\sigma_i^2\omega^2 + E[(u_i^2 + u_{i-1}^2 - 2u_{i-1}u_i)(u_i - u_{i-1})(-u_i)] \\ &= -\sigma_i^2\omega^2 - 2\sigma_i^2\omega^2 - \kappa\omega^4 - \omega^4 - 2\omega^4 = -3\sigma_i^2\omega^2 - (\kappa+3)\omega^4, \end{aligned}$$

such that $\text{cov}(x_{i,0}, x_{i,1}) = -3\sigma_i^2\omega^2 - 6\omega^4 - (\sigma_i^2 + 2\omega^2)(-\omega^2) = -2\sigma_i^2\omega^2 - (\kappa+1)\omega^4$, and similarly

$$E(x_{i,0}x_{i-1,1}) = E(y_i^2 y_{i-1} y_i) = E[(y_i^{*2} + 2y_i^* e_i + e_i^2)(y_i^* + u_i - u_{i-1})(u_{i-1})]$$

$$
\begin{aligned}
&= -\sigma_i^2\omega^2 - 2\sigma_i^2\omega^2 + E[e_i^2(u_i - u_{i-1})(u_{i-1})] \\
&= -\sigma_i^2\omega^2 - 2\sigma_i^2\omega^2 + E[(u_i^2 + u_{i-1}^2 - 2u_{i-1}u_i)(u_i - u_{i-1})(u_{i-1})] \\
&= -\sigma_i^2\omega^2 - 2\sigma_i^2\omega^2 - \omega^4 - \kappa\omega^4 - 2\omega^4 = -3\sigma_i^2\omega^2 - (\kappa+3)\omega^4,
\end{aligned}
$$

which shows that $\mathrm{cov}(x_{i,0}, x_{i-1,1}) = -2\sigma_i^2\omega^2 - (\kappa+1)\omega^4$. For $k \geq 1$ we have

$$
\begin{aligned}
E(x_{i,0}x_{i+k,1}) &= E[(y_i^{*2} + 2y_i^* e_i + e_i^2)(y_{i+k}^* + e_{i+k})(y_{i+k+1}^* + e_{i+k+1})] \\
&= E[(y_i^{*2} + 2y_i^* e_i + e_i^2)(-u_{i+k}^2)] \\
&= -E[(y_i^{*2} + u_i^2 + u_{i-1}^2)(u_{i+k}^2)] = -\sigma_i^2\omega^2 - 2\omega^4,
\end{aligned}
$$

and similarly for $k \leq -2$. Thus $\mathrm{cov}(x_{i,0}, x_{i+k,1}) = 0$ for $k \geq 1$ and $k \leq -2$.

The only non-zero covariance between $x_{i,0}$ and $x_{i+k,2}$, is

$$
\mathrm{cov}(x_{i,0}, x_{i-1,2}) = \mathrm{cov}(e_i^2, e_{i-1}e_{i+1}) = E(e_i^2 u_{i-1}(-u_i)) = E(2u_{i-1}^2 u_i^2) = 2\omega^4,
$$

and for $j \geq 3$ we find that $\mathrm{cov}(x_{i,0}, x_{i+k,j}) = 0$ for all k.

For $h \geq 1$ we have

$$
\begin{aligned}
\mathrm{cov}(x_{i,h}, x_{i,h+1}) &= E(y_i y_{i+h} y_i y_{i+h+1}) = E[y_i^2 u_{i+h}(-u_{i+h})] = -(\sigma_i^2 + 2\omega^2)\omega^2, \\
\mathrm{cov}(x_{i,h}, x_{i-1,h+1}) &= E(y_i y_{i+h} y_{i-1} y_{i+h}) = E[-u_{i-1}u_{i-1}y_{i+h}^2] = -(\sigma_{i+h}^2 + 2\omega^2)\omega^2,
\end{aligned}
$$

and similarly $\mathrm{cov}(x_{i,h}, x_{i-1,h+2}) = E(e_i e_{i+h} e_{i-1} e_{i+h+1}) = E(-u_{i-1})(u_{i+h})(u_{i-1})(-u_{i+h}) = \omega^4$. ■

Lemma A.2 *(a)* $\sum_{i=1}^{m-h}(\sigma_i^2 + \sigma_{i+h}^2) = 2\int_0^1 \psi_{\frac{h}{m}}(s)\sigma^2(s)ds$*, and (b) given* (**V**) *and* $q_m = O(m^{1/2})$ *it holds that*

$$
m\sum_{i=1}^{m-q_m} \sigma_i^2\sigma_{i+q_m}^2 - \int_0^1 \psi_{\frac{q_m}{m}}(s)\sigma^4(s)ds = o(1).
$$

Proof. (a) Since $\sigma_i^2 = \int_{\frac{i-1}{m}}^{\frac{i}{m}} \sigma^2(s)ds$, the first result follows from the identity $\sum_{i=1}^{m-h}\sigma_i^2 = \int_0^{\frac{m-h}{m}} \sigma^2(s)ds$. (b) We note that $\sum_{i=1}^{m-q_m}\sigma_i^2\sigma_{i+q_m}^2 = \sum_{i=1}^{m-q_m}[\sigma_i^4 + \sigma_i^2(\sigma_{i+q_m}^2 - \sigma_i^2)]$ and similarly that $\sum_{i=1}^{m-q_m}\sigma_i^2\sigma_{i+q_m}^2 = \sum_{i=1}^{m-q_m}[\sigma_{i+q_m}^4 - \sigma_{i+q_m}^2(\sigma_{i+q_m}^2 - \sigma_i^2)]$ such that

$$
\sum_{i=1}^{m-q_m}\sigma_{i,m}^2\sigma_{i+q_m,m}^2 = \frac{1}{2}\sum_{i=1}^{m-q_m}(\sigma_{i,m}^4 + \sigma_{i+q_m,m}^4) - \frac{1}{2}\sum_{i=1}^{m-q_m}(\sigma_{i+q_m,m}^2 - \sigma_{i,m}^2)^2. \tag{A.4}
$$

First we consider the first term on the right hand side. Let $\delta_{i,m} \equiv t_{i,m} - t_{i-1,m}$ and note that $\delta_{i,m} = O(m^{-1})$ given (**T**). So for arbitrary pairs $(s_{i,m}, \tilde{s}_{i,m})$, $i = 1, \ldots, m$ of points, where $s_{i,m}, \tilde{s}_{i,m}, \in [t_{i-1,m}, t_{i,m}]$ we have that

$$
m\sum_{i=1}^{m-q_m}|\sigma^4(s_{i,m}) - \sigma^4(\tilde{s}_{i,m})|\delta_{i,m}^2 = m^{-1/2}\sum_{i=1}^{m-q_m}|\sigma^4(s_{i,m}) - \sigma^4(\tilde{s}_{i,m})|\delta_{i,m}^2 m^{3/2}
$$

$$\leq \quad m^{-1/2}\sum_{i=1}^{m}|\sigma^4(s_{i,m})-\sigma^4(\tilde{s}_{i,m})| = o(1),$$

where the equality holds for m sufficiently large given (**V**).

Next, we let $s_{i,m}$ and $\tilde{s}_{i,m}$ be the points in $[t_{i-1,m}, t_{i,m}]$ that are such that $\sigma^2(s_{i,m})\delta_{i,m} = \int_{t_{i-1,m}}^{t_{i,m}} \sigma^2(s)ds$ and $\sigma^4(\tilde{s}_{i,m})\delta_{i,m} = \int_{t_{i-1,m}}^{t_{i,m}} \tau'(s)\sigma^4(s)ds$, and we note that these points exist given the continuity of σ^2 and τ'. In now follows that

$$\begin{aligned}
m\sum_{i=1}^{m-q_m}\sigma^4_{i,m} &= m\sum_{i=1}^{m-q_m}\left(\int_{t_{i-1,m}}^{t_{i,m}}\sigma^2(s)ds\right)^2 = m\sum_{i=1}^{m-q_m}\sigma^4(s_{i,m})\delta^2_{i,m}, \\
&= \sum_{i=1}^{m-q_m}\tfrac{t_{i,m}-t_{i-1,m}}{1/m}\sigma^4(\tilde{s}_{i,m})\delta_{i,m} + o(1) \\
&= \sum_{i=1}^{m-q_m}\tau'(\tilde{s}_{i,m})\sigma^4(\tilde{s}_{i,m})\delta_{i,m} + \sum_{i=1}^{m-q_m}[\tfrac{\tau(s)-\tau(s-\frac{1}{m})}{1/m} - \tau'(\tilde{s}_{i,m})]\sigma^4(\tilde{s}_{i,m})\delta_{i,m} + o(1) \\
&= \int_0^{1-\frac{q_m}{m}}\tau'(s)\sigma^4(s)ds + o(1),
\end{aligned}$$

where we used that

$$\begin{aligned}
\sum_{i=1}^{m-q_m}[\tfrac{\tau(s)-\tau(s-\frac{1}{m})}{1/m} - \tau'(\tilde{s}_{i,m})]\sigma^4(\tilde{s}_{i,m})\delta_{i,m} &\leq \sup_s\left|\tfrac{\tau(s)-\tau(s-\frac{1}{m})}{1/m} - \tau'(\tilde{s}_{i,m})\right|\sum_{i=1}^{m}\sigma^4(\tilde{s}_{i,m})\delta_{i,m} \\
&= o(1)\int_0^1\sigma^4(s)ds = o(1).
\end{aligned}$$

By similar arguments we find that $m\sum_{i=1}^{m-q_m}\sigma^4_{i+q_m,m} = \int_{\frac{q_m}{m}}^1\sigma^4(s)ds + o(1)$, such that the first term on the right hand side of (A.4) can be expressed as $\frac{m}{2}\sum_{i=1}^{m-q_m}(\sigma_i^4+\sigma^4_{i+q_m}) = \int_0^1\psi_{\frac{q_m}{m}}(s)\sigma^4(s)ds + o(1)$.

Now consider the second term on the right hand side of (A.4).

$$\begin{aligned}
m\sum_{i=1}^{m-q_m}(\sigma^2_{i,m}-\sigma^2_{i+q_m,m})^2 &= m\sum_{i=1}^{m-q_m}[\delta_{i,m}\sigma^2(s_{i,m})-\delta_{i+q_m,m}\sigma^2(s_{i+q_m,m})]^2 \\
&= m^{-1/2}\sum_{i=1}^{m-q_m}[m^{3/4}\delta_{i,m}\sigma^2(s_{i,m})-m^{3/4}\delta_{i+q_m,m}\sigma^2(s_{i+q_m,m})]^2, \\
&\leq c_m^2 m^{-1/2}\sum_{i=1}^{m-q_m}[\sigma^2(s_{i,m})-\sigma^2(s_{i+q_m,m})]^2, \qquad \text{(A.5)}
\end{aligned}$$

where

$$c_m \equiv \sup_i m^{3/4}\delta_{i,m} = m^{-1/4}\sup_i\tfrac{\delta_{i,m}}{1/m} \leq m^{-1/4}[\sup_s\tau'(s)+\sup_i|\tfrac{\delta_{i,m}}{1/m}-\tau'(\tfrac{i}{m})|] = O(m^{-1/4}).$$

Now for m sufficiently large it holds that

$$[\sigma^2(s_{i,m})-\sigma^2(s_{i+q_m,m})]^2 \quad \leq \quad |\sigma^2(s_{i,m})-\sigma^2(s_{i+q_m,m})|$$

$$\leq \quad |\sigma^2_{(s_i)} - \sigma^2_{(t_i)}| + |\sigma^2_{(t_i)} - \sigma^2_{(t_{i+1})}| + \cdots + |\sigma^2_{(t_{i+q_m-1})} - \sigma^2_{(s_{i+q_m})}|,$$

where we write $\sigma^2_{(s_i)}$ as short for $\sigma^2(s_{i,m})$. So (A.5) can be written as q_m sums that each are of order $c_m^2 o(1)$ given (**V**), which shows that (A.5) is $o(m^{-1/2}q_m)$. So it now follows that $m\sum_{i=1}^{m-q_m}(\sigma^2_{i,m} - \sigma^2_{i+q_m,m})^2 = o(1)$ provided that $q_m = O(m^{1/2})$. This completes the proof. ■

Proof of Theorem 2. The results of Lemma A.1 are used extensively. First we note that

$$\begin{aligned}
\text{var}(\hat{\gamma}_0) &= \text{var}(\sum_{i=1}^{m} x_{i,0}) = \sum_{i=1}^{m}\text{var}(x_{i,0}) + \sum_{i=2}^{m}\text{cov}(x_{i,0}, x_{i-1,0}) + \sum_{i=1}^{m-1}\text{cov}(x_{i,0}, x_{i+1,0}) \\
&= \sum_{i=1}^{m}[(2\kappa + 2)\omega^4 + 8\sigma_i^2\omega^2 + 2\sigma_i^4] + 2(\kappa - 1)(m - 1)\omega^4 \\
&= (4\kappa m - 2(\kappa - 1))\omega^4 + 8IV\omega^2 + 2IQ\tfrac{1}{m} + o(\tfrac{1}{m}).
\end{aligned}$$

This result is identical to that derived in Hansen & Lunde (2004*a*). Similarly,

$$\begin{aligned}
\text{var}(\hat{\gamma}_1) &= \sum_{i=1}^{m-1}\text{var}(x_{i,1}) + \sum_{i=2}^{m-1}\text{cov}(x_{i,1}, x_{i-1,1}) + \sum_{i=1}^{m-2}\text{cov}(x_{i,1}, x_{i+1,1}) \\
&= \sum_{i=1}^{m-1}[(\kappa + 2)\omega^4 + 2(\sigma_i^2 + \sigma_{i+1}^2)\omega^2 + \sigma_i^2\sigma_{i+1}^2] + 2(m - 2)\omega^4 \\
&= ((\kappa + 4)m - (\kappa + 6))\omega^4 + 4\omega^2 IV_{(\frac{1}{m})} + \tfrac{1}{m}IQ_{(\frac{1}{m})} + o(\tfrac{1}{m}).
\end{aligned}$$

For $h \geq 2$ we find

$$\begin{aligned}
\text{var}(\hat{\gamma}_h) &= \sum_{i=1}^{m-h}\text{var}(x_{i,h}) + \sum_{i=2}^{m-h}\text{cov}(x_{i,h}, x_{i-1,h}) + \sum_{i=1}^{m-h-1}\text{cov}(x_{i,h}, x_{i+1,h})] \\
&= \sum_{i=1}^{m-h}[4\omega^4 + 2(\sigma_i^2 + \sigma_{i+h}^2)\omega^2 + \sigma_i^2\sigma_{i+h}^2] + 2(m - h - 1)\omega^4 \\
&= (6m - 6h - 2)\omega^4 + 4\omega^2 IV_{(\frac{h}{m})} + \tfrac{1}{m}IQ_{(\frac{h}{m})} + o(\tfrac{h}{m}).
\end{aligned}$$

Next, we consider the covariance terms.

$$\begin{aligned}
\text{cov}(\hat{\gamma}_0, \hat{\gamma}_1) &= \text{cov}(\sum_{i=1}^{m} x_{i,0}, \sum_{i=1}^{m-1} x_{i,1}) = \sum_{i=1}^{m-1}\text{cov}(x_{i,0}, x_{i,1}) + \sum_{i=1}^{m-1}\text{cov}(x_{i+1,0}, x_{i,1}) \\
&= \sum_{i=1}^{m-1}[-2\sigma_i^2\omega^2 - (\kappa + 1)\omega^4] + \sum_{i=1}^{m-1}[-2\sigma_{i+1}^2\omega^2 - (\kappa + 1)\omega^4] \\
&= -(2\kappa + 2)(m - 1)\omega^4 - 4\omega^2 IV_{(\frac{1}{m})},
\end{aligned}$$

and similarly

$$\text{cov}(\hat{\gamma}_0, \hat{\gamma}_2) = \text{cov}(\sum_{i=1}^{m} x_{i,0}, \sum_{i=1}^{m-2} x_{i,2}) = \sum_{i=1}^{m-2}\text{cov}(x_{i+1,0}, x_{i,2}) = (2m - 4)\omega^4,$$

while $\text{cov}(\hat{\gamma}_0, \hat{\gamma}_k) = 0$ for $k \geq 3$.

For $h \geq 1$ we find:

$$\begin{aligned}
\text{cov}(\hat{\gamma}_h, \hat{\gamma}_{h+1}) &= \text{cov}(\sum_{i=1}^{m-h} x_{i,h}, \sum_{i=1}^{m-h-1} x_{i,h+1}) = \sum_{i=1}^{m-h-1} \text{cov}(x_{i,h}, x_{i,h+1}) + \sum_{i=1}^{m-h-1} \text{cov}(x_{i+1,h}, x_{i,h+1}) \\
&= -\sum_{i=1}^{m-h-1} (\sigma_i^2 + 2\omega^2)\omega^2 - \sum_{i=1}^{m-h-1} (\sigma_{i+h+1}^2 + 2\omega^2)\omega^2 \\
&= -4(m-h-1)\omega^4 - 2\omega^2 IV_{\frac{h+1}{m}}, \\
\text{cov}(\hat{\gamma}_h, \hat{\gamma}_{h+2}) &= \text{cov}(\sum_{i=1}^{m-h} x_{i,h}, \sum_{i=1}^{m-h-2} x_{i,h+2}) = (m-h-2)\omega^4,
\end{aligned}$$

which $\text{cov}(\hat{\gamma}_h, \hat{\gamma}_{h+k}) = 0$ for $k \geq 3$. ■

Proof of Corollary 1. From Theorem 2 we have that

$$\text{cov}(2\hat{\gamma}_2, \hat{\gamma}_0 + 2\hat{\gamma}_1) = (4-16)(m-2)\omega^4 - 8\omega^2 IV_{\frac{2}{m}} = -12m\omega^4 + 24\omega^4 - 8\omega^2 IV_{\frac{2}{m}},$$

such that

$$\begin{aligned}
\text{var}(\hat{\gamma}_0 + 2\hat{\gamma}_1 + \hat{\gamma}_2) &= \text{var}(\hat{\gamma}_0 + 2\hat{\gamma}_1) + \tfrac{1}{4}\text{var}(2\hat{\gamma}_2) + \tfrac{1}{2}2\text{cov}(\hat{\gamma}_0 + 2\hat{\gamma}_1, 2\hat{\gamma}_2) \\
&= 8(m-1)\omega^4 + 8\omega^2 IV + 6IQ\tfrac{1}{m} + o(\tfrac{1}{m}) \\
&\quad + \tfrac{24}{4}(m-2)m\omega^4 - \tfrac{8}{4}\omega^4 + \tfrac{16}{4}\omega^2 IV_{\frac{2}{m}} + \tfrac{4}{4}IQ_{\frac{2}{m}}\tfrac{1}{m} + o(\tfrac{1}{m}) \\
&\quad -12m\omega^4 + 24\omega^4 - 8\omega^2 IV_{\frac{2}{m}} \\
&= [8 + \tfrac{24}{4} - 12]m\omega^4 + (-8 - 12 - 2 + 24)\omega^4 \\
&\quad +8\omega^2 IV - 4\omega^2 IV_{\frac{2}{m}} + 6IQ\tfrac{1}{m} + IQ_{\frac{2}{m}}\tfrac{1}{m} + o(\tfrac{1}{m}) \\
&= 2m\omega^4 + 4\omega^2 IV + 2\omega^2(v_{\frac{2}{m}} + \omega^2) + 7IQ\tfrac{1}{m} + o(\tfrac{1}{m}).
\end{aligned}$$

The second result follows by defining $\mathbf{w} = (1, 1, \frac{7}{10}, \frac{3}{10})'$ and

$$\mathbf{w}'\mathbf{A}_{[4]}\mathbf{w} = \tfrac{4}{5}, \qquad \mathbf{w}'\mathbf{B}_{[4]}\mathbf{w} = \tfrac{68}{25}, \qquad \mathbf{w}'\mathbf{C}_{[4]}\mathbf{w} = \tfrac{208}{25},$$

where $\mathbf{A}_{[4]}$, $\mathbf{B}_{[4]}$, and $\mathbf{C}_{[4]}$, are the upper left 4×4 submatrices of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, respectively, and the calculations

$$\begin{array}{ll}
[\omega^4] & -1(-32+28) - 2\tfrac{7}{10}(8 - 32 + 24\tfrac{7}{10}) - 3\tfrac{3}{10}(8 - 32\tfrac{7}{10} + 24\tfrac{3}{10}) - 2\mathbf{w}'\mathbf{C}_{[4]}\mathbf{w} = \tfrac{98}{25}, \\
[v_\rho] & (-16 + 16\tfrac{7}{10})\tfrac{7}{10} = -16\tfrac{21}{100} \quad \text{and} \quad (-16\tfrac{7}{10} + 16\tfrac{3}{10})\tfrac{3}{10} = -16\tfrac{12}{100},
\end{array}$$

that quantifies the remaining terms. ■

Proof of Corollary 2. From Theorem 2 it follows that $E(\mathbf{w}'\hat{\gamma}) = w_0(IV + 2\omega^2 m) - w_1 \frac{m-1}{m} 2\omega^2 m$, such that $\text{bias}(\mathbf{w}'\hat{\gamma}) = (w_0 - 1)IV + (w_0 - \frac{m-1}{m} w_1) 2\omega^2 m$. The result for the variance follows by the structure of the matrices **A**, **B**, and **C**. ■

Proof of Theorem 3. It follows directly that

$$mV_1(\mathbf{w}) = \mathbf{w}'(\bar{\mathbf{A}}m - 2\mathbf{C})\mathbf{w} = M_{11} + \mathbf{w}_2'\mathbf{M}_{22}\mathbf{w}_2 + 2\mathbf{M}_{12}\mathbf{w}_2, \tag{A.6}$$

using the constraint $w_0 = 1$, and the decomposition of the $m \times m$ matrix

$$\bar{\mathbf{A}}m - 2\mathbf{C} = \begin{pmatrix} M_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix}.$$

By the first order condition of the right hand side of (A.6) yields $\mathbf{w}_2^\star \equiv -\mathbf{M}_{22}^{-1}\mathbf{M}_{21}$, and by substitution it follows that

$$mV_1(\mathbf{w}^\star) = \mathbf{w}^{\star\prime}(\bar{\mathbf{A}}m - 2\mathbf{C})\mathbf{w}^\star = M_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{M}_{21}.$$

While a closed-form expression for $mV_1(\mathbf{w}^\star)$ is unavailable it is easy to establish that $mV_1(\mathbf{w}^\star) \to 4$ as $m \to \infty$, numerically. The following table gives $mV_1(\mathbf{w}^\star)$ for some values of m.

m	10	50	100	200	500	1000	2000	5000
$mV_1(\mathbf{w}^\star)$	4.8837	4.1732	4.0850	4.0418	4.0165	4.008	4.0041	4.0016

■

Proof of Lemma 4. The first result follows from the identity

$$\begin{aligned} V_1 &= 12\tfrac{m-1}{m}\tfrac{q-1}{q} + \tfrac{m-1}{m}(\tfrac{q-1}{q})^2 4(7 - 8\tfrac{m-1}{m}) + \tfrac{m-2}{m}\tfrac{q-2}{q} 8(3\tfrac{q-2}{q} - 4\tfrac{q-1}{q} + \tfrac{m-1}{m}\tfrac{q-1}{q}) \\ &\quad + \sum_{j=3}^{q} \tfrac{m-j}{m}\tfrac{q-j}{q} 8(3\tfrac{q-j}{q} - 4\tfrac{q-j+1}{q} + \tfrac{q-j+2}{q}) - \tfrac{4}{m}(\tfrac{m-1}{m}\tfrac{q-1}{q})^2 - \tfrac{8}{m}\sum_{j=1}^{q}\tfrac{(q-j)^2}{q^2} \\ &= 4\tfrac{-1+2q+5m-q^2-6m^2-6qm+2m^3+q^2m+m^2q^2+m^2q}{m^3q^2} \\ &= \frac{4}{m} + \frac{1}{m}O(\tfrac{1}{m^2q^2} + \tfrac{1}{m^2q} + \tfrac{1}{mq^2} + \tfrac{1}{m^2} + \tfrac{1}{q^2} + \tfrac{1}{mq} + \tfrac{m}{q^2} + \tfrac{1}{m} + \tfrac{1}{q}) \end{aligned}$$

Similarly we have that

$$\begin{aligned} V_0 &= 8(\tfrac{m-1}{m}\tfrac{q-1}{q})^2 + \tfrac{m-1}{m}(\tfrac{q-1}{q})^2 16(1 - \tfrac{m-1}{m}) + \sum_{j=2}^{q}\tfrac{m-j}{m}\tfrac{q-j}{q}16(\tfrac{q-j}{q} - \tfrac{q-j+1}{q}) \\ &= \tfrac{8}{3}\tfrac{-3+6m+6q+3m^2q-3m^2-3q^2+q^3m-7qm}{m^2q^2} = O(\tfrac{q}{m}), \end{aligned}$$

and

$$V_{-1} = 2(\tfrac{m-1}{m}\tfrac{q-1}{q})^2 + \sum_{j=1}^{q}\tfrac{m-j}{m}4(\tfrac{q-j}{q})^2$$

$$= \frac{1}{3}\frac{6-12m-12q-10m^2q+6m^2+6q^2+24qm-11mq^2+4m^2q^3-mq^4}{m^2q^2} = O(\frac{q^2}{m})$$

■

B. Proofs of Section 4

Proof of Theorem 6. The first couple of subgrids are given by

$$\begin{aligned}
\mathcal{G}_{2_1} &= \{t_0, t_2, \ldots, t_{m-1}\}, \quad \mathcal{G}_{2_2} = \{t_1, t_3, \ldots, t_m\}, \\
\mathcal{G}_{3_1} &= \{t_0, t_3, \ldots, t_{m-2}\}, \quad \mathcal{G}_{3_2} = \{t_1, t_4, \ldots, t_{m-1}\}, \quad \mathcal{G}_{3_3} = \{t_2, t_5, \ldots, t_m\}, \\
\mathcal{G}_{4_1} &= \{t_0, t_4, \ldots, t_{m-3}\}, \quad \mathcal{G}_{4_2} = \{t_1, t_5, \ldots, t_{m-2}\}, \quad \ldots .
\end{aligned}$$

Since $y_{t_i,t_{i+j}} = y_{i+1} + \cdots + y_{i+j}$, we find that

$$\begin{aligned}
RV_{\mathcal{G}_{2_1}} + RV_{\mathcal{G}_{2_2}} &= (y_1 + y_2)^2 + \cdots + (y_{m-2} + y_{m-1})^2 + \\
&\quad (y_2 + y_3)^2 + \cdots + (y_{m-1} + y_m)^2 \\
&= 2\sum_{i=1}^{m} y_i^2 + 2\sum_{i=1}^{m-1} y_i y_{i+1} + r_2 = 2(\hat{\gamma}_0 + \hat{\gamma}_1) - r_2
\end{aligned}$$

where $r_2 = y_1^2 + y_m^2$. Similarly for $q = 3$ we have

$$\begin{aligned}
\sum_{j=1}^{3} RV_{\mathcal{G}_{3_j}} &= (y_1 + y_2 + y_3)^2 + \cdots + (y_{m-4} + y_{m-3} + y_{m-2})^2 + \\
&\quad (y_2 + y_3 + y_4)^2 + \cdots + (y_{m-3} + y_{m-2} + y_{m-1})^2 + \\
&\quad (y_3 + y_4 + y_5)^2 + \cdots + (y_{m-2} + y_{m-1} + y_m)^2 \\
&= 3\sum_{i=1}^{m} y_i^2 + 4\sum_{i=1}^{m-1} y_i y_{i+1} + 2\sum_{i=1}^{m-2} y_i y_{i+2} + r_3 \\
&= 3\hat{\gamma}_0 + 4\hat{\gamma}_1 + 2\hat{\gamma}_2 - r_3,
\end{aligned}$$

where the remainder is given by $r_3 = y_1^2 + y_m^2 + (y_1 + y_2)^2 + (y_{m-1} + y_m)^2 = r_2 + (y_1 + y_2)^2 + (y_{m-1} + y_m)^2$.

Similarly for $k = 4$ we find

$$\sum_{j=1}^{4} RV_{\mathcal{G}_{4_j}} = 4\sum_{i=1}^{m} y_i^2 + 6\sum_{i=1}^{m-1} y_i y_{i+1} + 4\sum_{i=1}^{m-3} y_i y_{i+2} + 2\sum_{i=1}^{m-4} y_i y_{i+3} - r_4,$$

where $r_4 = r_3 + (y_1 + y_2 + y_3)^2 + (y_{m-2} + y_{m-1} + y_m)^2$ and in the general case we

$$\sum_{j=1}^{k} RV_{\mathcal{G}_{k_j}} = k\sum_{i=1}^{m} y_i^2 + 2(k-1)\sum_{i=1}^{m-1} y_i y_{i+1} + \cdots + 2\sum_{i=1}^{m-k} y_i y_{i+k} - r_k$$

$$= \quad k\hat{\gamma}_0 + \sum_{h=1}^{k} 2(k-h)\hat{\gamma}_h - r_k,$$

where $r_k = r_{k-1} + (y_1 + \cdots + y_{k-1})^2 + (y_{m-k+2} + \cdots + y_m)^2$. So it follows that

$$k^{-1}\sum_{j=1}^{k} RV_{\mathcal{G}_{k_j}} = \tfrac{1}{k}[k\hat{\gamma}_0 + \sum_{h=1}^{k} 2(k-h)\hat{\gamma}_h] - \tfrac{r_k}{k} = \hat{\gamma}_0 + 2\sum_{h=1}^{k} \tfrac{k-h}{k}\hat{\gamma}_h - \tfrac{r_k}{k},$$

which completes the proof. ■

Lemma A.3 *Define* $z_j \equiv x_{j,0} + 2\sum_{i=1}^{j-1} x_{j-i,i}$ *for* $j = 1, \ldots, m-1$. *Then it holds that* $\mathrm{var}(z_1) = 8\omega^4 + 8\omega^2\sigma_1^2 + 2\sigma_1^4$, *whereas*

$$\mathrm{var}(z_j) = 12\omega^4 + 8\omega^2(\sigma_1^2 + \cdots + \sigma_j^2) + \sigma_j^2(4\sigma_1^2 + \cdots + 4\sigma_{j-1}^2 + 2\sigma_j^2), \qquad \text{for } j \geq 2.$$

The covariances are given by: $\mathrm{cov}(z_j, z_{j+1}) = -6\omega^4 - 4\omega^2(\sigma_1^2 + \cdots + \sigma_j^2)$ *while* $\mathrm{cov}(z_j, z_{j+h}) = 0$ *for* $|h| \geq 2,\ j = 1, 2, \ldots$.

(*Under* ($\mathbf{T}'$) *where* $\sigma_i^2 = \sigma^2/m$ *we have* $\mathrm{var}(z_j) = 12\omega^4 + 8j\omega^2 IV/m + 4(j - \frac{1}{2})IV^2/m^2$ *for* $j \geq 2$ *and* $\mathrm{cov}(z_j, z_{j+1}) = -6\omega^4 - 4j\omega^2 IV/m$ *for all* $j \geq 1$).

Proof of Lemma A.3. From Lemma A.1 we have that

$$\begin{aligned}
\mathrm{var}(z_1) &= \mathrm{var}(x_{1,0}) = 8\omega^4 + 8\omega^2\sigma_1^2 + 2\sigma_1^4, \\
\mathrm{var}(z_2) &= \mathrm{var}(x_{2,0}) + 4\mathrm{var}(x_{1,1}) + 4\mathrm{cov}(x_{2,0}, x_{1,1}) \\
&= [8\omega^4 + 8\omega^2\sigma_2^2 + 2\sigma_2^4] + 4[5\omega^4 + 2(\sigma_1^2 + \sigma_2^2)\omega^2 + \sigma_1^2\sigma_2^2] \\
&\quad +4[-4\omega^4 - 2\omega^2\sigma_2^2] \\
&= 12\omega^4 + 8\omega^2(\sigma_1^2 + \sigma_2^2) + \sigma_2^2(4\sigma_1^2 + 2\sigma_2^2), \\
\mathrm{var}(z_3) &= \mathrm{var}(x_{3,0}) + 4\mathrm{var}(x_{2,1}) + 4\mathrm{var}(x_{1,2}) \\
&\quad +4\mathrm{cov}(x_{3,0}, x_{2,1}) + 4\mathrm{cov}(x_{3,0}, x_{1,2}) + 8\mathrm{cov}(x_{2,1}, x_{1,2}) \\
&= [8\omega^4 + 8\omega^2\sigma_3^2 + 2\sigma_3^4] + 4[5\omega^4 + 2(\sigma_2^2 + \sigma_3^2)\omega^2 + \sigma_2^2\sigma_3^2] \\
&\quad +4[4\omega^4 + 2(\sigma_1^2 + \sigma_3^2)\omega^2 + \sigma_1^2\sigma_3^2] \\
&\quad +4[-4\omega^4 - 2\omega^2\sigma_3^2] + 4[0] + 8[-2\omega^4 - \omega^2\sigma_3^2] \\
&= 12\omega^4 + 8\omega^2(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) + \sigma_3^2(4\sigma_1^2 + 4\sigma_2^2 + 2\sigma_3^2), \\
\mathrm{var}(z_4) &= \mathrm{var}(x_{4,0}) + 4\mathrm{var}(x_{3,1}) + 4\mathrm{var}(x_{2,2}) + 4\mathrm{var}(x_{1,3}) \\
&\quad +4\mathrm{cov}(x_{4,0}, x_{3,1}) + 4\mathrm{cov}(x_{4,0}, x_{2,2}) + 4\mathrm{cov}(x_{4,0}, x_{1,3})
\end{aligned}$$

$$
\begin{aligned}
&+8\mathrm{cov}(x_{3,1},x_{2,2})+8\mathrm{cov}(x_{3,1},x_{1,3})+8\mathrm{cov}(x_{2,2},x_{1,3})\\
=\ &[8\omega^4+8\omega^2\sigma_4^2+2\sigma_4^4]+4[5\omega^4+2(\sigma_3^2+\sigma_4^2)\omega^2+\sigma_3^2\sigma_4^2]\\
&+4[4\omega^4+2(\sigma_2^2+\sigma_4^2)\omega^2+\sigma_2^2\sigma_4^2]+4[4\omega^4+2(\sigma_1^2+\sigma_4^2)\omega^2+\sigma_1^2\sigma_4^2]\\
&+4[-4\omega^4-2\omega^2\sigma_4^2]+4[0]+4[0]\\
&+8[-2\omega^4-\omega^2\sigma_4^2]+8[0]+8[-2\omega^4-\omega^2\sigma_4^2]\\
=\ &12\omega^4+8\omega^2(\sigma_1^2+\sigma_2^2+\sigma_3^2+\sigma_4^2)+\sigma_4^2(4\sigma_1^2+4\sigma_2^2+4\sigma_3^2+2\sigma_4^2),
\end{aligned}
$$

and the general result follows by the correlation structure of $x_{i,j}$. Next, we note that

$$
\begin{aligned}
\mathrm{cov}(z_1,z_2) &= \mathrm{cov}(x_{1,0},x_{2,0}+2x_{1,1})=[2\omega^4]+2[-2(\sigma_1^2+2\omega^2)\omega^2]=-6\omega^4-4\omega^2\sigma_1^2,\\
\mathrm{cov}(z_2,z_3) &= \mathrm{cov}(x_{2,0}+2x_{1,1},x_{3,0}+2x_{2,1}+2x_{1,2})\\
&= [2\omega^4]+2[-2(\sigma_2^2+2\omega^2)\omega^2]+2[2\omega^4]+2[0]+4[\omega^4]+4[-(\sigma_1^2+2\omega^2)\omega^2]\\
&= -6\omega^4-4\omega^2(\sigma_1^2+\sigma_2^2),
\end{aligned}
$$

and the general result follows by induction. The higher order covariance are verified to be zero from the correlation structure of $x_{i,j}$. ■

Lemma A.4 *Given the assumptions of Theorem 2, it holds that*

$$
\begin{aligned}
\mathrm{cov}(\hat\gamma_0,z_1) &= 10\omega^4+8\omega^2\sigma_1^2+2\sigma_1^4\\
\mathrm{cov}(\hat\gamma_0,z_2) &= -4\omega^4+4\omega^2(\sigma_2^2-\sigma_1^2)+2\sigma_2^4\\
\mathrm{cov}(\hat\gamma_0,z_j) &= 4\omega^2(\sigma_j^2-\sigma_{j-1}^2)+2\sigma_j^4 \qquad \text{for } j\geq 3
\end{aligned}
$$

and in general we have for $i \geq 1$ that

$$
\begin{aligned}
\mathrm{cov}(\hat\gamma_i,z_{i-1}) &= 0,\\
\mathrm{cov}(\hat\gamma_i,z_i) &= -2\omega^4-2\omega^2\sigma_1^2, \qquad \text{for } i\geq 2 \qquad \mathrm{cov}(\hat\gamma_1,z_1)=-4\omega^4-2\omega^2\sigma_1^2\\
\mathrm{cov}(\hat\gamma_i,z_{i+1}) &= 4\omega^4+4\omega^2\sigma_1^2+2\omega^2(\sigma_{i+1}^2-\sigma_2^2)\ +2\sigma_1^2\sigma_{i+1}^2,\\
\mathrm{cov}(\hat\gamma_i,z_{i+2}) &= -2\omega^4+2\omega^2[(\sigma_2^2-\sigma_1^2)-(\sigma_3^2-\sigma_2^2)]\ +2\sigma_2^2\sigma_{i+2}^2,\\
\mathrm{cov}(\hat\gamma_i,z_{i+j}) &= 2\omega^2[(\sigma_j^2-\sigma_{j-1}^2)-(\sigma_{j+1}^2-\sigma_j^2)]+2\sigma_j^2\sigma_{i+j}^2, \qquad \text{for } j\geq 3.
\end{aligned}
$$

Proof. The structure follows from the correlation structure of $x_{i,j}$.

$\hat{\gamma}_0$	$x_{1,0}$	$x_{2,0}$	$x_{1,1}$	$x_{3,0}$	$x_{2,1}$	$x_{1,2}$	$x_{4,0}$	$x_{3,1}$	$x_{2,2}$	$x_{1,3}$	$x_{5,0}$	$x_{4,1}$	$x_{3,2}$	$x_{2,3}$	$x_{1,4}$
$x_{1,0}$	8	2	−4												
$x_{2,0}$	2	8	−4	2	−4	2									
$x_{3,0}$		2		8	−4		2	−4	2						
$x_{4,0}$				2			8	−4			2	−4	2		
$x_{5,0}$							2				8	−4			
$x_{6,0}$											2				
$x_{7,0}$															

The table above identify the non-zero correlations, where the multiple of the ω^4-term is given, whereas the other two terms (involving $\omega^2\sigma_i^2$ and $\sigma_i^2\sigma_j^2$) are given from Lemma A.1. Since $\hat{\gamma}_0 = \sum x_{i,0}$ we find by the definition of z_i, that

$$\begin{aligned}
\text{cov}(\hat{\gamma}_0, z_1) &= [8\omega^4 + 8\omega^2\sigma_1^2 + 2\sigma_1^4] + [2\omega^4] = 10\omega^4 + 8\omega^2\sigma_1^2 + 2\sigma_1^4, \\
\text{cov}(\hat{\gamma}_0, z_2) &= [2\omega^4] + [8\omega^4 + 8\omega^2\sigma_2^2 + 2\sigma_2^4] + [2\omega^4] \\
&\quad + 2[-4\omega^4 - 2\omega^2\sigma_1^2] + 2[-4\omega^4 - 2\omega^2\sigma_2^2] \\
&= -4\omega^4 + 4\omega^2(\sigma_2^2 - \sigma_1^2) + 2\sigma_2^4, \\
\text{cov}(\hat{\gamma}_0, z_3) &= [2\omega^4] + [8\omega^4 + 8\omega^2\sigma_3^2 + 2\sigma_3^4] + [2\omega^4] \\
&\quad + 2[-4\omega^4 - 2\omega^2\sigma_2^2] + 2[-4\omega^4 - 2\omega^2\sigma_3^2] + 2[2\omega^2] \\
&= 4\omega^2(\sigma_3^2 - \sigma_2^2) + 2\sigma_3^4.
\end{aligned}$$

$\hat{\gamma}_1$	$x_{1,0}$	$x_{2,0}$	$x_{1,1}$	$x_{3,0}$	$x_{2,1}$	$x_{1,2}$	$x_{4,0}$	$x_{3,1}$	$x_{2,2}$	$x_{1,3}$	$x_{5,0}$	$x_{4,1}$	$x_{3,2}$	$x_{2,3}$	$x_{1,4}$
$x_{1,1}$	−4	−4	5		1	−2									
$x_{2,1}$		−4	1	−4	5	−2		1	−2	1					
$x_{3,1}$				−4	1		−4	5	−2			1	−2	1	
$x_{4,1}$							−4	1			−4	5	−2		
$x_{5,1}$											−4	1			
$x_{6,1}$															

Since $\hat{\gamma}_1 = \sum x_{i,1}$ it follows that

$$\begin{aligned}
\text{cov}(\hat{\gamma}_1, z_1) &= -4\omega^4 - 2\omega^2\sigma_1^2, \\
\text{cov}(\hat{\gamma}_1, z_2) &= [-4\omega^4 - 2\omega^2\sigma_2^2] + [-4\omega^4 - 2\omega^2\sigma_2^2]
\end{aligned}$$

$$
\begin{aligned}
& +2[5\omega^4 + 2\omega^2(\sigma_1^2 + \sigma_2^2) + \sigma_1^2\sigma_2^2] + 2[\omega^2] \\
&= 4\omega^4 + 4\omega^2\sigma_1^2 + 2\sigma_1^2\sigma_2^2, \\
\mathrm{cov}(\hat{\gamma}_1, z_3) &= [-4\omega^4 - 2\omega^2\sigma_3^2] + [-4\omega^4 - 2\omega^2\sigma_3^2] \\
& +2[\omega^4] + 2[5\omega^4 + 2(\sigma_2^2 + \sigma_3^2) + \sigma_2^2\sigma_3^2] + 2[\omega^2] \\
& +2[-2\omega^4 - \omega^2\sigma_1^2] + 2[-2\omega^4 - \omega^2\sigma_3^2] \\
&= -2\omega^4 + 2\omega^2[(\sigma_2^2 - \sigma_1^2) - (\sigma_3^2 - \sigma_2^2)] + 2\sigma_2^2\sigma_3^2, \\
\mathrm{cov}(\hat{\gamma}_1, z_{1+j}) &= \quad 2\omega^2[(\sigma_j^2 - \sigma_{j-1}^2) - (\sigma_{j+1}^2 - \sigma_j^2)] + 2\sigma_j^2\sigma_{i+j}^2.
\end{aligned}
$$

$\hat{\gamma}_2$	$x_{1,0}$	$x_{2,0}$	$x_{1,1}$	$x_{3,0}$	$x_{2,1}$	$x_{1,2}$	$x_{4,0}$	$x_{3,1}$	$x_{2,2}$	$x_{1,3}$	$x_{5,0}$	$x_{4,1}$	$x_{3,2}$	$x_{2,3}$	$x_{1,4}$
$x_{1,2}$		2	−2		−2	4			1	−2					
$x_{2,2}$				2	−2	1		−2	4	−2			1	−2	1
$x_{3,2}$							2	−2	1			−2	4	−2	
$x_{4,2}$											2	−2	1		
$x_{5,2}$															

Since $\hat{\gamma}_2 = \sum x_{i,2}$ it now follows that

$$
\begin{aligned}
\mathrm{cov}(\hat{\gamma}_2, z_1) &= 0, \\
\mathrm{cov}(\hat{\gamma}_2, z_2) &= [2\omega^4] + 2[-2\omega^4 - \omega^2\sigma_1^2] = -2\omega^4 - 2\omega^2\sigma_1^2, \\
\mathrm{cov}(\hat{\gamma}_2, z_3) &= [2\omega^4] + 2[-2\omega^4 - \omega^2\sigma_3^2] + 2[-2\omega^4 - \omega^2\sigma_2^2] \\
& +2[4\omega^4 + 2\omega^2(\sigma_1^2 + \sigma_3^2) + \sigma_1^2\sigma_3^2] + 2[\omega^2] \\
&= 4\omega^4 + 4\omega^2\sigma_1^2 + 2\omega^2(\sigma_3^2 - \sigma_2^2) + 2\sigma_1^2\sigma_3^2, \\
\mathrm{cov}(\hat{\gamma}_2, z_4) &= [2\omega^4] + 2[-2\omega^4 - \omega^2\sigma_4^2] + 2[-2\omega^4 - \omega^2\sigma_3^2] + 2[\omega^4] \\
& +2[4\omega^4 + 2\omega^2(\sigma_2^2 + \sigma_4^2) + \sigma_2^2\sigma_4^2] + 2[\omega^2] \\
& +2[-2\omega^4 - \omega^2\sigma_1^2] + 2[-2\omega^4 - \omega^2\sigma_4^2] \\
&= -2\omega^4 + 2\omega^2[(\sigma_2^2 - \sigma_1^2) - (\sigma_3^2 - \sigma_2^2)] + 2\sigma_2^2\sigma_4^2, \\
\mathrm{cov}(\hat{\gamma}_2, z_{2+j}) &= \quad +2\omega^2[(\sigma_j^2 - \sigma_{j-1}^2) - (\sigma_{j+1}^2 - \sigma_j^2)] + 2\sigma_j^2\sigma_{i+j}^2
\end{aligned}
$$

Next we omit the $x_{h,0}$-columns as these correlations are all zero.

$\hat{\gamma}_3$	$x_{1,1}$	$x_{2,1}$	$x_{1,2}$	$x_{3,1}$	$x_{2,2}$	$x_{1,3}$	$x_{4,1}$	$x_{3,2}$	$x_{2,3}$	$x_{1,4}$	$x_{5,1}$	$x_{4,2}$	$x_{3,3}$	$x_{2,4}$	$x_{1,5}$
$x_{1,3}$		1	−2		−2	4			1	−2					
$x_{2,3}$				1	−2	1		−2	4	−2			1	−2	1
$x_{3,3}$							1	−2	1			−2	4	−2	
$x_{4,3}$											1	−2	1		
$x_{5,3}$															

Since $\hat{\gamma}_3 = \sum x_{i,3}$ it now follows that

$$
\begin{aligned}
\text{cov}(\hat{\gamma}_3, z_1) &= \text{cov}(\hat{\gamma}_3, z_2) = 0, \\
\text{cov}(\hat{\gamma}_3, z_3) &= 2[\omega^4] + 2[-2\omega^4 - \omega^2\sigma_1^2] = -2\omega^4 - 2\omega^2\sigma_1^2, \\
\text{cov}(\hat{\gamma}_3, z_4) &= 2[\omega^4] + 2[-2\omega^4 - \omega^2\sigma_4^2] + 2[-2\omega^4 - \omega^2\sigma_2^2] \\
&\quad +2[4\omega^4 + 2\omega^2(\sigma_1^2 + \sigma_4^2) + \sigma_1^2\sigma_4^2] + 2[\omega^2] \\
&= 4\omega^4 + 4\omega^2\sigma_1^2 + 2\omega^2(\sigma_4^2 - \sigma_2^2) + 2\sigma_1^2\sigma_4^2, \\
\text{cov}(\hat{\gamma}_3, z_5) &= 2[\omega^4] + 2[-2\omega^4 - \omega^2\sigma_5^2] + 2[-2\omega^4 - \omega^2\sigma_3^2] + 2[\omega^4] \\
&\quad 2[4\omega^4 + 2\omega^2(\sigma_2^2 + \sigma_5^2) + \sigma_2^2\sigma_5^2] + 2[\omega^2] \\
&\quad +2[-2\omega^4 - \omega^2\sigma_1^2] + 2[-2\omega^4 - \omega^2\sigma_5^2] \\
&= -2\omega^4 + 2\omega^2[(\sigma_2^2 - \sigma_1^2) - (\sigma_3^2 - \sigma_2^2)] + 2\sigma_2^2\sigma_5^2, \\
\text{cov}(\hat{\gamma}_3, z_{3+j}) &= \quad +2\omega^2[(\sigma_j^2 - \sigma_{j-1}^2) - (\sigma_{j+1}^2 - \sigma_j^2)] + 2\sigma_j^2\sigma_{i+j}^2
\end{aligned}
$$

Results for $\hat{\gamma}_i$, $i \geq 4$ follows similarly. ■

Proof of Lemma 7. First note that $E(y_1 + \cdots + y_{k-1}) = 2\omega^2 + \sum_{i=1}^{k-1} \sigma_i^2$, such that

$$
E(r_k) = E(r_{k-1}) + \sum_{i=1}^{k-1}(\sigma_i^2 + \sigma_{m+1-i}^2) + 4\omega^2, \qquad E(r_1) = 0,
$$

which proves the first result. In the constant-volatility case the expression simplifies to

$$
E(r_k) = 2(1 + 2 + \cdots + k - 1)\frac{\sigma^2}{m} + (k-1)4\omega^2 = k(k-1)\frac{\sigma^2}{m} + (k-1)4\omega^2.
$$

To establish the results for the variance and covariance, it is convenient to define $z_j \equiv x_{j,0} + 2\sum_{i=1}^{j-1} x_{j-i,i}$ for $j = 1, 2, \ldots$ So $z_1 = x_{1,0}$, $z_2 = x_{2,0} + 2x_{1,1}$, $z_3 = x_{3,0} + 2x_{2,1} + 2x_{1,2}$, etc. Similarly, we define $\tilde{z}_1 \equiv x_{m,0}$, $\tilde{z}_2 \equiv x_{m-1,0} + 2x_{m-2,1}$, etc.

From calculations, such as

$$
y_1^2 + (y_1 + y_2)^2 + (y_1 + y_2 + y_3)^2 \quad = \quad 3y_1^2 + 2(y_2^2 + 2y_1y_2) + (y_3^2 + 2y_2y_3 + 2y_1y_3)
$$

$$= \quad 3x_{1,0} + 2(x_{2,0} + 2x_{1,1}) + (x_{3,0} + 2x_{2,1} + 2x_{1,2})$$
$$= \quad 3z_1 + 2z_2 + z_3,$$

it follows that

$$\tfrac{1}{k} r_k = \tfrac{1}{k} \sum_{j=1}^{k-1} (k-j) z_j + \tfrac{1}{k} \sum_{j=1}^{k-1} (k-j) \tilde{z}_j .$$

From Lemma A.3 it follows that

$$\begin{aligned}
\mathrm{var}\left(\sum_{j=1}^{k-1} \tfrac{k-j}{k} z_j \right) &= \sum_{j=1}^{k-1} \left(\tfrac{k-j}{k} \right)^2 \mathrm{var}\left(z_j\right) + 2 \sum_{j=1}^{k-1} \tfrac{k-j}{k} \tfrac{k-j-1}{k} \mathrm{cov}\left(z_j, z_{j+1}\right) \\
&= \sum_{j=1}^{k-1} \tfrac{k-j}{k} \left[\tfrac{k-j}{k} \mathrm{var}\left(z_j\right) + \tfrac{k-j-1}{k} 2\mathrm{cov}\left(z_j, z_{j+1}\right) \right] \\
&= \sum_{j=1}^{k-1} \tfrac{k-j}{k} \left[\tfrac{k-j}{k} - \tfrac{k-j-1}{k} \right] 12\omega^4 - \tfrac{k-1}{k} 4\omega^4 \\
&\quad + \sum_{j=1}^{k-1} \tfrac{k-j}{k} \left[\tfrac{k-j}{k} - \tfrac{k-j-1}{k} \right] 8\omega^2 (\sigma_1^2 + \cdots + \sigma_j^2) \\
&\quad + \sum_{j=1}^{k-1} \left(\tfrac{k-j}{k} \right)^2 \sigma_j^2 (4\sigma_1^2 + \cdots + 4\sigma_{j-1}^2 + 2\sigma_j^2).
\end{aligned}$$

Since

$$\begin{aligned}
\sum_{j=1}^{k} \tfrac{k-j}{k} \tfrac{1}{k} &= \tfrac{1}{2} \tfrac{k-1}{k}, \\
\sum_{j=1}^{k} \tfrac{k-j}{k} \tfrac{1}{k} \tfrac{j}{m} &= \tfrac{1}{6} \tfrac{k-1}{k} \tfrac{k+1}{m} = O(\tfrac{k}{m}), \qquad \text{and} \\
\sum_{j=1}^{k} \tfrac{k-j}{k} \tfrac{k-j}{k} \tfrac{j}{m^2} &= \tfrac{1}{12} \tfrac{k^2-1}{m^2} = O(\tfrac{k^2}{m^2}),
\end{aligned}$$

and $(\sigma_1^2 + \cdots + \sigma_j^2) = O(\frac{j}{m})$ and that $\sigma_j^2 (4\sigma_1^2 + \cdots + 4\sigma_{j-1}^2 + 2\sigma_j^2) = O(\frac{j}{m^2})$ under our assumptions, we find that

$$\begin{aligned}
\sum_{j=1}^{k-1} \tfrac{k-j}{k} \left[\tfrac{k-j}{k} - \tfrac{k-j-1}{k} \right] 12\omega^4 - \tfrac{k-1}{k} 4\omega^4 &= (6-4) \tfrac{k-1}{k} \omega^4 = 2 \tfrac{k-1}{k} \omega^4, \\
\sum_{j=1}^{k-1} \tfrac{k-j}{k} \left[\tfrac{k-j}{k} - \tfrac{k-j-1}{k} \right] 8\omega^2 (\sigma_1^2 + \cdots + \sigma_j^2) &= O(\tfrac{k}{m}), \qquad \text{and} \\
\sum_{j=1}^{k-1} \left(\tfrac{k-j}{k} \right)^2 \sigma_j^2 (4\sigma_1^2 + \cdots + 4\sigma_{j-1}^2 + 2\sigma_j^2) &= O(\tfrac{k^2}{m^2}),
\end{aligned}$$

which shows that $\mathrm{var}(\sum_{j=1}^{k-1}\frac{k-j}{k}z_j) = 2\frac{k-1}{k}\omega^4 + O(\frac{k}{m})$. Finally, by adding the contributions from the term $\sum_{j=1}^{k-1}\frac{k-j}{k}\tilde{z}_j$ that are derived in the same manner, and using that z_i and $\tilde{z}_j$ are uncorrelated for $i, j < m/2$, the result for $\mathrm{var}(\frac{1}{k}r_k)$ follows.

Next, we consider the covariance between r_k and $\hat{\gamma}_i$, for $i = 0, 1, \ldots$ From Lemma A.4 it follows that

$$\begin{aligned} & \mathrm{cov}(\hat{\gamma}_0, \tfrac{1}{k}r_k) \overset{(1)}{=} \tfrac{k-1}{k}\mathrm{cov}(\hat{\gamma}_0, z_1 + \tilde{z}_1) + \tfrac{k-2}{k}\mathrm{cov}(\hat{\gamma}_0, z_2 + \tilde{z}_2) + 0 \\ = \; & 2\omega^4(\tfrac{k-1}{k}10 - \tfrac{k-2}{k}\tfrac{k}{m}) = 2\omega^4(6\tfrac{k-1}{k} + \tfrac{4}{k}) = 2\omega^4(6\tfrac{k-1+4/6}{k}) = 12\omega^4(\tfrac{k-1/3}{k}). \end{aligned}$$

For the remaining elements of $\widehat{\gamma}$ that involve $\hat{\gamma}_h + \hat{\gamma}_{-h} = 2\sum_{i=1}^{m-h} x_{i,h}$, we find similarly that

$$\mathrm{cov}(\hat{\gamma}_1 + \hat{\gamma}_{-1}, \tfrac{1}{k}r_k) \overset{(1)}{=} 2\cdot 2\left[-\tfrac{k-1}{k}4 + \tfrac{k-2}{k}4 - \tfrac{k-3}{k}2\right] = -8\tfrac{k-1}{k},$$

whereas

$$\mathrm{cov}(\hat{\gamma}_h + \hat{\gamma}_{-h}, \tfrac{1}{k}r_k) \overset{(1)}{=} 4\left[-\tfrac{k-h}{k}2 + \tfrac{k-h-1}{k}4 - \tfrac{k-h-2}{k}2\right] = 0, \qquad \text{for } h \geq 2.$$

This completes the proof. ■

C. **Proof of Results of Section 5**

Proof of Theorem 8.

By the independence of z_i and $\tilde{z}_i$ we have for $i, j \geq 1$ that,

$$\begin{aligned} \mathrm{cov}(\tilde{\gamma}_0, 2\tilde{\gamma}_i) &= \mathrm{cov}(\hat{\gamma}_0, 2\hat{\gamma}_i) - \mathrm{cov}(\hat{\gamma}_0, z_i + \tilde{z}_i) \\ &\overset{a}{=} \mathrm{cov}(\hat{\gamma}_0, 2\hat{\gamma}_i) - 2\mathrm{cov}(\hat{\gamma}_0, z_i), \\ \mathrm{var}(2\tilde{\gamma}_i) &= \mathrm{var}(2\hat{\gamma}_i) + \mathrm{var}(z_i + \tilde{z}_i) - 2\mathrm{cov}(2\hat{\gamma}_i, z_i + \tilde{z}_i) \\ &\overset{a}{=} \mathrm{var}(2\hat{\gamma}_i) + 2\mathrm{var}(z_i) - 8\mathrm{cov}(\hat{\gamma}_i, z_i), \\ \mathrm{cov}(2\tilde{\gamma}_i, 2\tilde{\gamma}_j) &= \mathrm{cov}(2\hat{\gamma}_i, 2\hat{\gamma}_j) + \mathrm{cov}(z_i + \tilde{z}_i, z_j + \tilde{z}_j) - \mathrm{cov}(2\hat{\gamma}_i, z_j + \tilde{z}_j) - \mathrm{cov}(z_i + \tilde{z}_i, 2\hat{\gamma}_j) \\ &\overset{a}{=} \mathrm{cov}(2\hat{\gamma}_i, 2\hat{\gamma}_j) + 2\mathrm{cov}(z_i, z_j) - 4\mathrm{cov}(\hat{\gamma}_i, z_j) - 4\mathrm{cov}(z_i, \hat{\gamma}_j), \end{aligned}$$

where $\overset{a}{=}$ refers to equality under the assumption that $\sigma_i^2 = \sigma_j^2$ for all i, j, in which case the contributions from z_i and $\tilde{z}_i$ are identical.

Thus, the elements of $\tilde{\mathbf{A}}\omega^4 + \tilde{\mathbf{B}}\omega^2\sigma^2 + \tilde{\mathbf{C}}\sigma^4\frac{1}{m}$ are given as follows.

$$\begin{aligned} [0,1] &= -2\mathrm{cov}(\hat{\gamma}_0, z_1) = -20\omega^4 - 16\omega^2\sigma^2/m - 4\sigma^4/m^2, \\ [0,2] &= -2\mathrm{cov}(\hat{\gamma}_0, z_2) = +8\omega^4 + 0 - 4\sigma^4/m^2, \end{aligned}$$

$$\begin{aligned}
[0,i] &= -2\mathrm{cov}(\hat{\gamma}_0, z_i) = -4\sigma^4/m^2, \qquad \text{for } i \geq 2, \\
[1,1] &= 2\mathrm{var}(z_1) - 8\mathrm{cov}(\hat{\gamma}_1, z_1) \\
&= 2[8\omega^4 + 8\omega^2\sigma^2/m + 2\sigma^4/m^2] - 8[-4\omega^4 - 2\omega^2\sigma^2/m] \\
&= 48\omega^4 + 32\omega^2\sigma^2/m + 4\sigma^4/m^2,
\end{aligned}$$

and more generally for $i \geq 1$ we have

$$\begin{aligned}
[i,i+1] &= 2\mathrm{cov}(z_i, z_{i+1}) - 4\mathrm{cov}(\hat{\gamma}_i, z_{i+1}) - 4\mathrm{cov}(z_i, \hat{\gamma}_{i+1}) \\
&= 2[-6\omega^4 - 4i\omega^2\sigma^2/m] - 4[4\omega^4 + 4\omega^2\sigma^2/m + 2\sigma^4/m^2] - 4[0] \\
&= -28\omega^4 - 8(i+2)\omega^2\sigma^2/m - 8\sigma^4/m^2, \\
[i,i+2] &= 2\mathrm{cov}(z_i, z_{i+2}) - 4\mathrm{cov}(\hat{\gamma}_i, z_{i+2}) - 4\mathrm{cov}(z_i, \hat{\gamma}_{i+2}) \\
&= 2[0] - 4[-2\omega^4 + 2\sigma^4/m^2] - 4[0] = 8\omega^4 - 8\sigma^4/m^2, \\
[i,i+j] &= 2\mathrm{cov}(z_i, z_{i+j}) - 4\mathrm{cov}(\hat{\gamma}_i, z_{i+j}) - 4\mathrm{cov}(z_i, \hat{\gamma}_{i+j}) \\
&= 2[0] - 4[2\sigma^4/m^2] - 4[0] = -8\sigma^4/m^2, \qquad \text{for } j \geq 3.
\end{aligned}$$

Further, for $i \geq 2$ we find that

$$\begin{aligned}
[i,i] &= 2\mathrm{var}(z_i) - 8\mathrm{cov}(\hat{\gamma}_i, z_i) \\
&= 2[12\omega^4 + 8\omega^2 i\omega^2\sigma^2/m + 4(i - \tfrac{1}{2})\sigma^4/m^2] - 8[-2\omega^4 - 2\omega^2\sigma^2/m] \\
&= 40\omega^4 + 16(i+1)\omega^2\sigma^2/m + 8(i - \tfrac{1}{2})\sigma^4/m^2.
\end{aligned}$$

■

References

Aït-Sahalia, Y., Mykland, P. A. & Zhang, L. (2003), How often to sample a continuous-time process in the presence of market microstructure noise, Working Paper w9611, NBER.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**(1), 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modeling and forecasting realized volatility', *Econometrica* **71**(2), 579–625.

Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, John Wiley and Sons, New York.

Anderson, T. W. & Takemura, A. (1986), 'Why do noninvertible estimated moving average occur?', *Journal of Time Series Analysis* **7**, 235–254.

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

Bandi, F. M. & Russell, J. R. (2004), Microstructure noise, realized volatility, and optimal sampling, Working paper, Graduate School of Business, The University of Chicago.

Barndorff-Nielsen, O. E. & Shephard, N. (2002), 'Econometric analysis of realised volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society* **B 64**, 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2003), 'Realized power variation and stochastic volatility', *Bernoulli* **9**, 243–265.

Barndorff-Nielsen, O. E. & Shephard, N. (2005), 'Power variation and time change', *Theory of Probability and Its Applications* . Forthcoming.

Bartlett, M. S. (1950), 'Periodogram analysis and continuous spectra', *Biometrica* **37**, 1–16.

Bollen, B. & Inder, B. (2002), 'Estimating daily volatility in financial markets utilizing intraday data', *Journal of Empirical Finance* **9**, 551–562.

Carlstein, E. (1986), 'The use of subseries values for estimating teh variance of a general statistic from a stationary time series', *Annals of Statistics* **14**, 1171–1179.

Gloter, A. & Jacod, J. (2001*a*), 'Diffusions with measurement errors. i. local asymptotic normality', *Probability and Statistics* **5**, 225–242.

Gloter, A. & Jacod, J. (2001*b*), 'Diffusions with measurement errors. i. measurement errors', *Probability and Statistics* **5**, 243–260.

Hansen, P. R. & Lunde, A. (2003), 'An optimal and unbiased measure of realized variance based on intermittent high-frequency data'. Mimeo prepared for the CIREQ-CIRANO Conference: Realized Volatility. Montreal, November 2003.

Hansen, P. R. & Lunde, A. (2004*a*), 'Realized variance and market microstructure noise'. `http://www.stanford.edu/people/peter.hansen`.

Hansen, P. R. & Lunde, A. (2004*b*), 'A realized variance for the whole day based on intermittent high-frequency data'. `http://www.stanford.edu/people/peter.hansen`.

Hansen, P. R. & Lunde, A. (2004*c*), 'An unbiased measure of realized variance'. `http://www.stanford.edu/people/peter.hansen`.

Jacod, J. (1994), 'Limit of random measures associated with the increments of a Brownian semimartingale'. Unpublished paper: Laboratorie de Probabilities, Universite P and M Curie, Paris.

Jacod, J. & Protter, P. (1998), 'Asymptotic error distributions for the Euler method for stochastic differential equations', *Annals of Probability* **26**, 267–307.

Künsch, H. R. (1989), 'The jackknife and the bootstrap for general stationary observations', *Annals of Statistics* **17**, 1217–1241.

Liu, R. Y. & Singh, K. (1992), Moving blocks jackknife and bootstrap capture weak dependence., *in* R. LePage & L. Billard, eds, 'Exploring the Limits of Bootstrap', John Wiley, New York, pp. 225–248.

Müller, U. A. (1993), Statistics of variables observed over overlapping intervals, discussion paper, O&A Research Group.

Müller, U. K. (2004), A theory of robust long-run variance estimation, Mimeo, Economics Department, Princeton University.

Mykland, P. A. & Zhang, L. (2005), 'ANOVA for diffusions', *Annals of Statistics* **33**. Forthcoming.

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

Oomen, R. A. C. (2004*a*), 'Properties of bias corrected realized variance in calendar time and business time'. manuscript, Warwick Business School, The University of Warwick.

Oomen, R. A. C. (2004*b*), 'Properties of realized variance for a pure jump process: Calendar time sampling versus business time sampling'. manuscript, Warwick Business School, The University of Warwick.

Politis, D. N., Romano, J. P. & Wolf, M. (1999), *Subsampling*, Springer, New York.

Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Vol. 1, Academic Press, London.

Shephard, N. (1993), 'Distribution of the ML estimator of an MA(1) and local level model', *Econometric Theory* **9**, 377–401.

Tanaka, K. & Satchell, S. E. (1989), 'Asymptotic properties of the maximum likelihood and nonlinear least squares estimator for noninvertible moving average models', *Econometric Theory* **5**, 333–353.

Zhang, L., Mykland, P. A. & Aït-Sahalia, Y. (2002), A tale of two time scales: Determining integrated volatility with noisy high frequency data?, Tech. rep. university of chicago.

Zhang, L., Mykland, P. A. & Aït-Sahalia, Y. (2004), A tale of two time scales: Determining integrated volatility with noisy high frequency data, Paper presented at NBER summer meeting, July 15, 2004.

Zhou, B. (1996), 'High-frequency data and volatility in foreign-exchange rates', *Journal of Business & Economic Statistics* **14**(1), 45–52.

Table 1: Ancillary Regression Results.

Panel A: Regular Kernel

	$\lambda = 0.1$			$\lambda = 0.01$			$\lambda = 0.001$		
m	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$
10^3	1.276	-0.515	1.175	0.244	-0.551	-0.108	0.217	-0.679	-1.019
10^4	1.237	-0.509	1.166	0.112	-0.532	-0.143	-0.242	-0.612	-1.146
10^5	1.215	-0.506	1.161	0.028	-0.521	-0.162	-0.541	-0.573	-1.212
10^6	1.200	-0.504	1.157	-0.026	-0.514	-0.173	-0.733	-0.550	-1.250

Panel B: Modified Kernel

	$\lambda = 0.1$			$\lambda = 0.01$			$\lambda = 0.001$		
m	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$
10^3	0.998	-0.495	1.031	0.141	-0.541	-0.142	0.185	-0.675	-1.027
10^4	1.005	-0.496	1.037	0.033	-0.525	-0.170	-0.265	-0.610	-1.151
10^5	1.016	-0.497	1.039	-0.034	-0.516	-0.185	-0.558	-0.571	-1.216
10^6	1.022	-0.498	1.040	-0.077	-0.511	-0.194	-0.747	-0.549	-1.253

Panel C: Subsample Estimator

	$\lambda = 0.1$			$\lambda = 0.01$			$\lambda = 0.001$		
m	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$	$\hat{\alpha}_m$	$\hat{\beta}_m$	$\hat{\alpha}_m^{\text{rest}}$
10^3	0.366	-0.371	0.105	-0.101	-0.481	-1.124	0.146	-0.658	**-2.099**
10^4	0.297	-0.361	0.073	-0.398	-0.438	-1.243	-0.455	-0.571	**-2.371**
10^5	0.242	-0.354	0.052	-0.621	-0.409	-1.318	-0.939	-0.508	**-2.544**
10^6									

The Table presents results from the local ancillary regressions that reveal the estimators rates of convergence. The local regressions are each based on five data points, $m_i = m/4, m/2, m, 2m,$ and $4m$, where m is listed in the first column. $\hat{\alpha}_m$ and $\hat{\beta}_m$ are the unrestricted estimates and $\hat{\alpha}_m^{\text{rest}}$ is the estimate of α_m when β_m is fixed at $-1/2$ (Panels A and B) or $-1/3$ (Panel C).

D. **Figures**

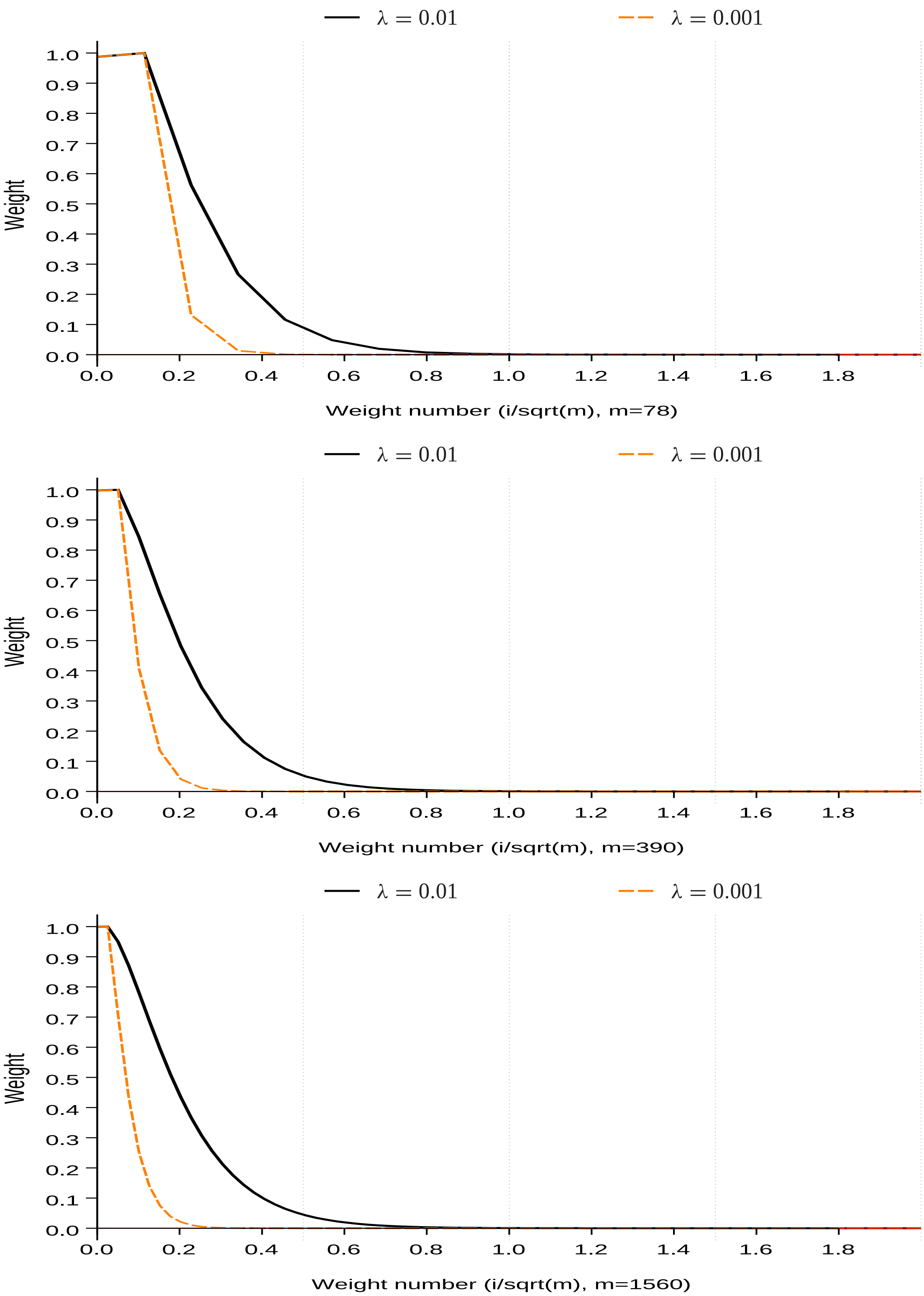

Figure 1: Plot of $\mathbf{w}^*$, the optimal weight. The number of observations, m, equals 78 in the top plot, 390 in the middle ,and in the bottom plot m equals 1560. $\lambda = \omega^2/\sigma^2$ is set to be 0.01 and 0.001 in each subplot.

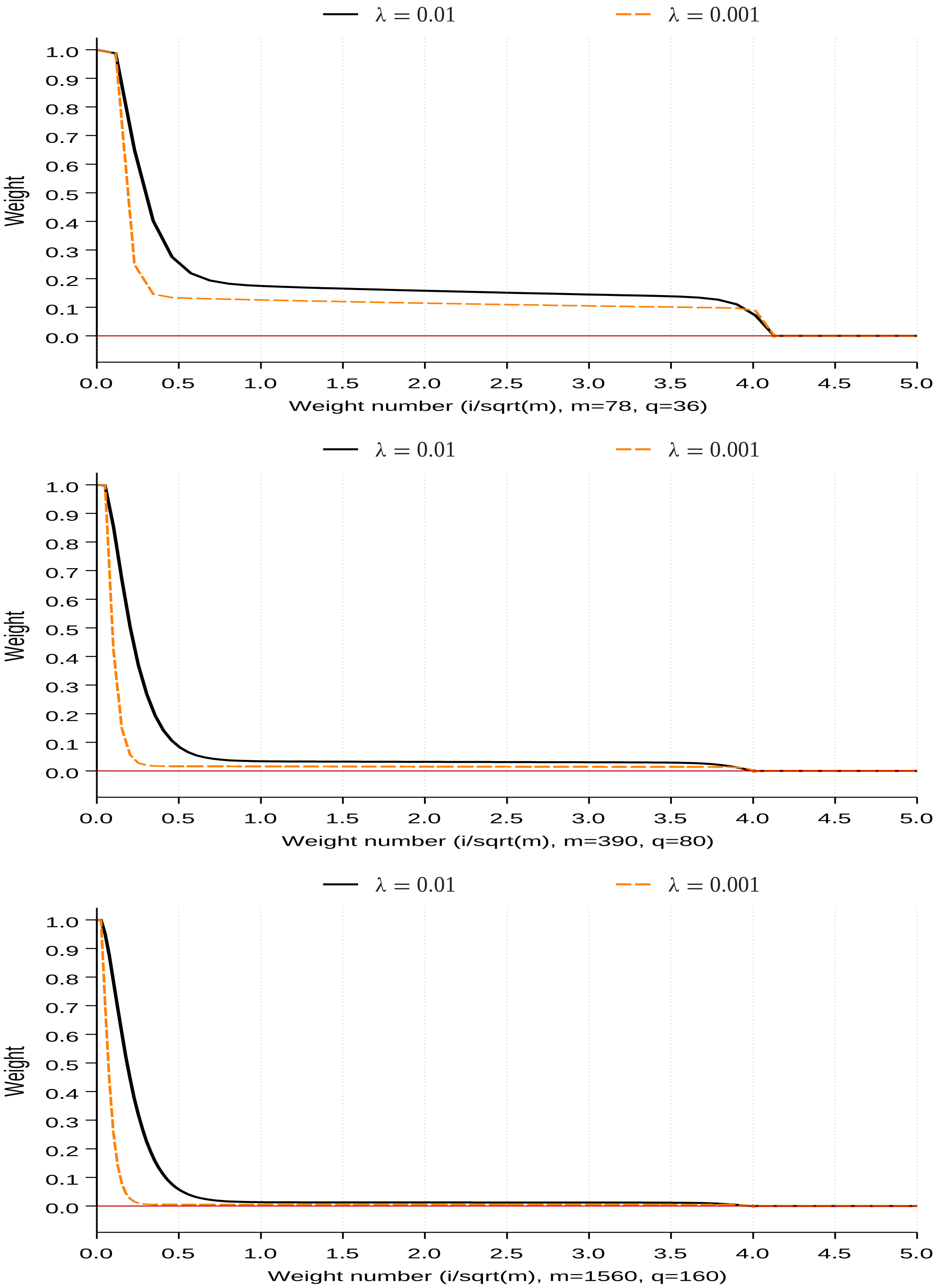

Figure 2: Plot of $\mathbf{w}_{\lambda}^{*}$, the optimal modified weight. The number of observations, m, equals 78 in the top plot, 390 in the middle ,and in the bottom plot m equals 1560. $\lambda = \omega^2/\sigma^2$ is set to be 0.01 and 0.001 in each subplot.

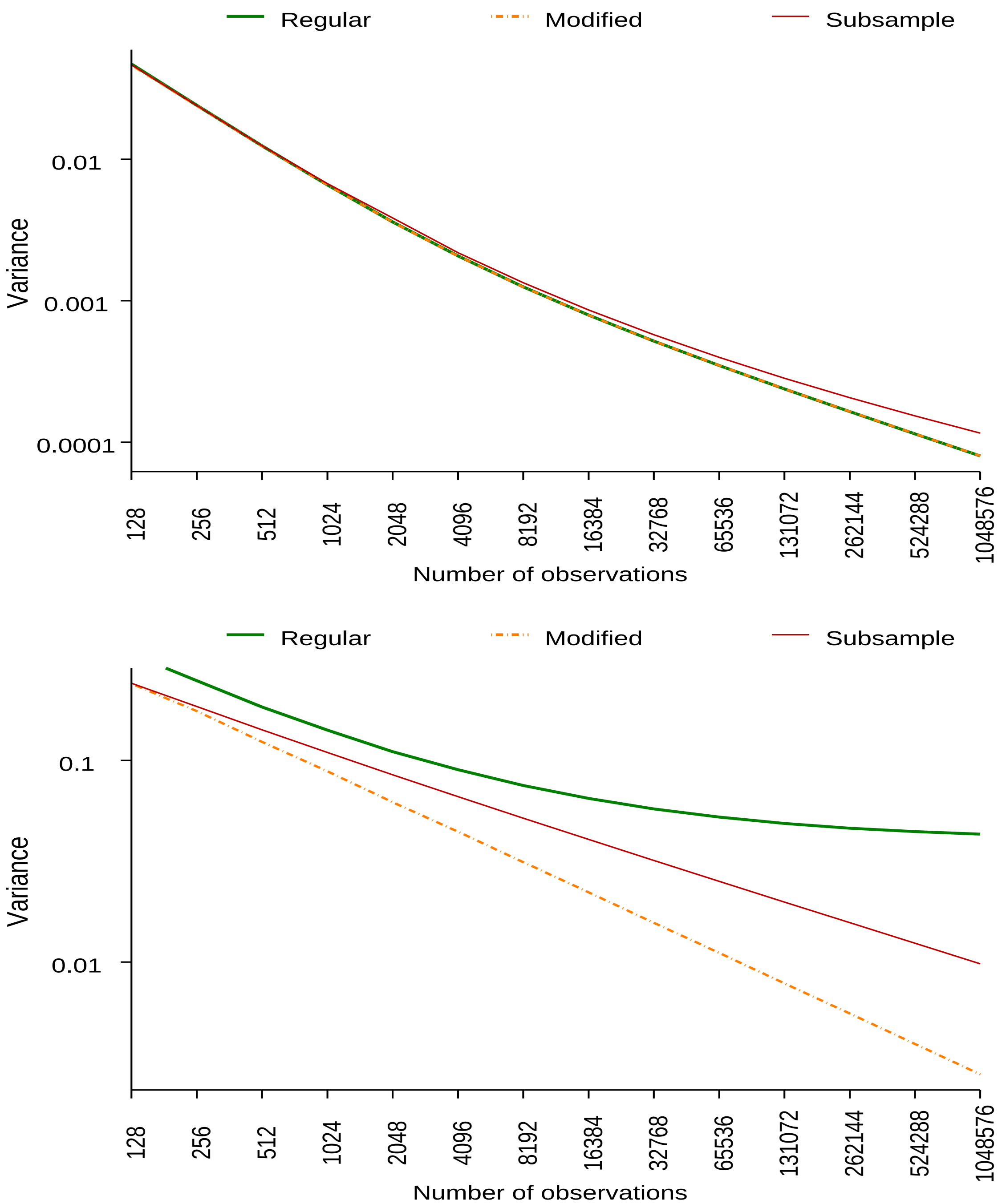

Figure 3: This figure present.... Top: $\lambda = 0.0001$, Bottom: $\lambda = 0.1$