Rounding Errors and Volatility Estimation

YINGYING LI
Department of Information Systems, Business Statistics and Operations Management, Hong Kong University of Science and Technology

PER A. MYKLAND
Department of Statistics, University of Chicago

ABSTRACT
Financial prices are often discretized—with smallest tick size of one cent, for example. Thus prices involve rounding errors. Rounding errors affect the estimation of volatility, and understanding them is critical, particularly when using high frequency data. We study the asymptotic behavior of realized volatility (RV), which is commonly used as an estimator of integrated volatility. We prove the convergence of the RV and scaled RV under various conditions on the rounding level and the number of observations. A bias-corrected volatility estimator is proposed and an associated central limit theorem is shown. The simulation and empirical results demonstrate that the proposed method can yield substantial statistical improvement. (JEL: C02, C13, C14)

KEYWORDS: rounding errors, bias-correction, diffusion process, market microstructure, realized volatility (RV)

High frequency data analysis has received substantial attention in recent years and volatility estimation is a central topic of interest. The primary difficulty in estimating daily volatilities by using high frequency data is the presence of market microstructure noise; notable developments have been made in this area. Zhang et al. (2003); Zhang (2006); Barndorff-Nielsen et al. (2008); and Xiu (2010) proposed and evaluated various volatility estimators that exhibited favorable financial support provided by the HK SAR RGC (grants SBR09/10.BM17 and GRF-602710), and the National Science Foundation (grants DMS 06-04758, SES 06-31605, and SES 11-24526) is gratefully acknowledged. We thank Professor Michael J. Wichura for very helpful comments related to the Lemma 4 of this paper; and we thank the Editor, Associate Editor, and two anonymous referees for their constructive comments. Address correspondence to Yingying Li, Department of Information Systems, Business Statistics and Operations Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, or e-mail: yyli@ust.hk.

doi:10.1093/jjfinec/nbu005 Advanced Access publication February 27, 2014 © The Author, 2014. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
convergence properties, assuming that the microstructure noise was additive, and independent and identically distributed (i.i.d.). Li and Mykland (2007) and Jacod et al. (2009) studied cases in which the market microstructure noise was a combination of additive noise and rounding error. Rosenbaum (2009) proposed a novel volatility estimation approach, using absolute values of the increments when rounding is the only source of the market microstructure noise.

Rounding is a crucial source of market microstructure noise that should not be ignored. Because stocks are traded using discrete price grids, their observations are effectively rounded. In certain cases, particularly when the stock prices are low, rounding can be the main source of market microstructure noise. Figure 1 shows the second-by-second stock prices of Citigroup Inc on May 1, 2007, indicating that the log prices of the stock did not exhibit the pattern of a diffusion process or a diffusion process with additive noise. Rather, these prices seem like samples from a rounded diffusion.

In this article, we focus on the extreme case where there is pure rounding. We explore the commonly used volatility estimator, the realized volatility (RV), and how it can be bias-corrected to yield consistent volatility estimates. RV goes back to the path breaking work of Andersen and Bollerslev (1997), Andersen et al. (2001, 2003), Barndorff-Nielsen and Shephard (2002), Jacod and Protter (1998), among others.

We consider a security price process $S$, whose logarithm $X = \log S$ follows

$$dX_t = \mu_t dt + \sigma_t dW_t.$$  

(1)

In other words, $S$ is the solution to the following stochastic differential equation:

$$dS_t = (\mu_t + \frac{1}{2} \sigma_t^2)S_t dt + \sigma_t S_t dW_t, \quad t \in [0,1]$$  

(2)

where $W_t$ is a standard Brownian motion. We assume that $\mu_t$ and $\sigma_t$ are continuous random processes satisfying the regularity conditions specified in Section 4.

It is a common practice in finance to use the sum of frequently sampled squared returns, the RV, to estimate the integrated volatility $\int_0^1 \sigma_t^2 dt$. However, empirical studies have shown that because of market microstructure noise, RV can be severely biased when prices are sampled at high frequencies, whereas sampling sparsely yields more reasonable estimates (see, for example, the signature plots introduced by Andersen et al. (2000)). In this study, we investigate the case in which the contamination caused by market microstructure results solely from rounding errors.

Let $\alpha_n$ be a sequence of positive numbers representing the accuracy of measurement when the price process is observed $n$ times during a time period [0,1]. Suppose that at time $i/n$ ($i=0, \ldots, n$), the value $k\alpha_n$ is observed when the true value $S_{i/n}$ is in $[k\alpha_n, (k+1)\alpha_n)$ with $k \in \mathbb{Z}$. For every real $s$, we denote by $s(\alpha_n) = \alpha_n \lfloor s/\alpha_n \rfloor$ its rounded-off value at level $\alpha_n$. Taking the Citigroup data as in Figure 1, for example, the rounding level is $\alpha_0 = 0.01$. On the day shown, the $k$ ranged from 296 to 317.
We investigate the asymptotic behavior of the RV, as follows:

$$V_n = \sum_{i=1}^{n} (Y_i - Y_{i-1})^2,$$

(3)

where $Y_i = \log(c(a_n))$, $i=0,\ldots,n$ represents the observed log prices. Our main results are presented for the case in which the rounding is down as previously described. This is primarily to facilitate presenting the proofs, where results of \cite{Delattre1997} for round-off errors are applied. Same or similar results also apply to rounding up or rounding to the nearest multiple of $\alpha_n$ (see Remark 5 and Section 2.3.2 for additional details).

\cite{Jacod1996} and \cite{Delattre1997} previously studied volatility interference based on a rounded Itô diffusion; although their work inspired the current study, we seek in this article to spell out what ensues when rounding occurs on the original (e.g., the US dollar, euro, etc.) scale rather than the log scale. As we shall see later in this article, our findings indicate that this more realistic rounding yields a bias which requires a somewhat more complicated correction. For example, in the simple case that the volatility is constant, the bias is no longer a function of the volatility (see Remark 1).

We shall provide the limit of $V_n$, demonstrating that the RV can be problematic when rounding errors are present, and elucidating why “sparse sampling” could be a practical way to estimate volatility (however, sparsely sampling does not solve all the problems). We subsequently propose a bias-corrected estimator, and prove an associated central limit theorem. The simulation results demonstrate that the proposed bias-corrected estimator yields substantially enhanced statistical accuracy. Empirical studies show that the bias-correction can facilitate financial risk management. Our main bias correction applies to the case of “small rounding” as in \cite{Delattre1997} and \cite{Rosenbaum2009}. Such asymptotics are realistic in practice, cf. the findings for additive error in \cite{Zhang2011}. Small rounding asymptotics has also been studied in \cite{Kolassa1990}, where it is
shown to be related to additive error. We also discuss the effects on RV when the rounding is not “small”.

The theoretical results are presented in Section 1; the simulation studies are presented in Section 2, and the empirical studies in Section 3; Section 4 concludes. The proofs are shown in the Appendix.

1 ASYMPTOTIC RESULTS

We assume that the latent security price process $S_t$ follows (2), where $\sigma_t$ is a nonrandom function of $S_t$, of class $C^5$ on $[0, \infty)$ (In the Black-Scholes model, $\sigma_t = \sigma$ is a constant). Assume further that $\mu_t$ is a continuous random process (in particular, it is locally bounded).

Let $\beta_n = \alpha_n \sqrt{n}$.

**Theorem 1:** When $\alpha_n \to 0$ as $n \to \infty$ such that $\beta_n \to \beta \in [0, \infty)$, we have

$$V^n \to \frac{1}{\sqrt{2\pi} \sigma_t^2} \left( \int_0^1 \sigma_t^2 dt + \frac{\beta^2}{6} \int_0^1 \frac{1}{S_t^2} dt - \frac{\beta^2}{\pi^2} \int_0^1 \frac{1}{S_t^2} dt \right) \exp \left( - \frac{2\pi^2 k^2 \sigma_t^2 S_t^2}{\beta^2} \right) dt.$$

One sees from this result that the bias is always positive when $\beta \neq 0$, rapidly increasing as $\beta$ grows. Also, the bias is smaller when the stock price $S_t, t \in [0, 1]$ is larger. Figure 2 gives a visual representation of this. This result captures the empirical features that

a) subsampling helps (the same $\alpha$ value and a smaller $n$ value yields a smaller $\beta$ and correspondingly smaller bias); and

b) the rounding effect is less severe for more expensive stocks (i.e., higher $S_t$ values correspond to smaller biases).

Theorem 1 shows that when $\beta_n \to 0$, one has the consistency of $V^n$. If $\beta_n$ decays polynomially in $n$, we have the following central limit theorem.

**Theorem 2:** When $\beta_n = O(n^{-\gamma})$ for some $\gamma > 0$, we have

$$\sqrt{n} \left( V^n - \int_0^1 \sigma_t^2 dt - \frac{\beta_n^2}{6} \int_0^1 \frac{1}{S_t^2} dt \right) \to \mathcal{L}_{\text{stable}} \int_0^1 \sqrt{2\pi} \sigma_t^2 dB_t,$$

where $B$ is a standard Brownian motion independent of $W$.

In this case, a finite sample bias of $\frac{\beta_n^2}{\pi} \int_0^1 \frac{1}{S_t^2} dt$ remains. The bias can be estimated and a bias-corrected estimator can be determined as follows.
Figure 2 RV $V^n$ versus $\beta$ based on Theorem 1 and three simulated sample paths ($\mu_t = 0, \sigma_t = 0.01$) with starting prices $S_0 = 1$, $S_0 = 10$ and $S_0 = 20$. The dashed line represents the true integrated volatility, which is 0.0001; the solid curves represent the limits of the RV. The curve shapes indicate that the bias is increasing in $\beta$, and comparing the ranges of the y axes of the plots demonstrates that the bias is smaller when $S_t, t \in [0, 1]$ is larger.

Theorem 3: Assume that $\beta_n = O(n^{-\gamma})$ for some $\gamma > 0$, and let

$$V^n_0 := V^n - \frac{1}{6} \sum_{i=1}^{n} \frac{1}{(\sigma_i S_{i/n})^2}.$$

Then as $n \rightarrow \infty$,

$$\sqrt{n} \left( V^n_0 - \int_0^1 \sigma_t^2 dt \right) \rightarrow \mathcal{L} - \text{stably} \int_0^1 \sqrt{2} \sigma_t^2 dB_t,$$

where $B$ is a standard Brownian motion independent of $W$.

The simulation results in the subsequent section demonstrate that this bias-correction yields substantially improved estimates. The empirical studies further show that the bias correction can be quite helpful in risk analysis.
Remark 1: In the case where $\sigma_t \equiv \sigma$, it is documented in section 4 of Delattre and Jacod (1997) that when rounding is implemented on the log scale, the bias is a function of $\sigma^2 = \int_0^1 \sigma^2_t dt$. We see from the above results that in the presence of this more realistic type of rounding, the bias is no longer a function of the targeted integrated volatility even when $\sigma_t \equiv \sigma$. Therefore this requires somewhat more complicated bias correction.

Remark 2: The condition of small rounding is necessary for the asymptotic results above. In practice, we apply these asymptotic results via expansion—we observe only one $\alpha_n$ and one $n$ for a particular price process in a specified time period. When small rounding is relevant, we can make a correction as in Theorem 3, yielding a superior estimator. In practice, we usually cannot change $\alpha_n$. If the $\beta_n$ is too big due to high sampling frequency, we can analyze a subsample with a moderate frequency to establish a situation with small $\beta_n$. (Ref. simulation studies for additional detail.)

Remark 3: The condition that the random process $\sigma_t$ is a nonrandom function of $S_t$ is assumed so that the framework of Delattre and Jacod (1997) can be applied. In Sections 2 and 3, we see in simulation and empirical studies that even when this condition is not necessarily satisfied, the proposed bias correction can still be very helpful. We conjecture that similar results hold also in stochastic volatility settings.

When the small rounding condition is not satisfied, the RV would blow up as the sampling frequency becomes larger. Theorem 4 illustrates the asymptotic result of a simple case where $\sigma_t \equiv \sigma$. In this case, simple bias correction is insufficient. A correction after subsampling may help.

Theorem 4: Let the accuracy of measurement $\alpha_n \equiv \alpha$ be independent of the number of observations $n$. Consider the case when $\sigma_t \equiv \sigma$ for $t \in [0, 1]$. Redefine $S_{1/n}^{(\alpha)} = \alpha$ if $S_{1/n}^{(\alpha)} = 0$. As $n \to \infty$,

$$
\frac{1}{\sqrt{n}} V^n \to \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} L_k \log((k+1)\alpha) \left( \log \frac{k+1}{k} \right)^2,
$$

where $L_k$ is the local time at the level $a$ of the continuous semimartingale $X_t = \log S_t$ (see Revuz and Yor (1999), page 222).

Remark 4: Redefining $S_{1/n}^{(\alpha)} = \alpha$ if $S_{1/n}^{(\alpha)} = 0$ rules out the possibility of yielding a logarithm of zero for log prices. In practice, this simply means that the security price does not go below the smallest rounding grid (1 cent if $\alpha = 0.01$) during the specified time period.

1 We emphasize that our derivation builds on the general results of Delattre and Jacod (1997).
2 A formal extension to this more general case can use a simple parametric approximation to the process, perhaps via the contiguity arguments in Mykland and Zhang (2011).
Remark 5: When rounding is not down, but rather to the nearest multiple of \( \alpha_n \), the results of Theorems 1–3 remain the same, but a small adjustment must be made to the limit of Theorem 4: the local times will be at levels \( \log((k + \frac{1}{2})\alpha) \) instead of \( \log((k+1)\alpha) \).

2 SIMULATION STUDIES

2.1 Moderate Sampling Frequencies

Consider first the simplest case that \( \sigma_t \equiv \sigma \) for \( t \in [0, 1] \). Denote by \( V^n_{-CI} \) and \( V^n_0_{-CI} \) the nominal 95% confidence interval (CI) based on \( V^n \) and \( V^n_0 \), respectively.

The naïve CI based on \( V^n \) relies on the classical theory for RV, which indicates the following when there is no observation error:

\[
\sqrt{n}[V^n - \sigma^2] \rightarrow \mathcal{L}(0, 2\sigma^4).
\]

The resulting nominal 95% CI is as follows:

\[
V^n_{-CI} = \left[ V^n - 1.96 \sqrt{2(V^n)^2/n}, V^n + 1.96 \sqrt{2(V^n)^2/n} \right].
\]

Our findings in the previous section indicate that the RV is no longer reliable when rounding errors are present. We proposed the following simple bias-corrected estimator that should function when \( \alpha_n \sqrt{n} \) is reasonably small:

\[
V^n_0 = V^n - \frac{\alpha_n^2}{6} \sum_{i=1}^{n} \frac{1}{(S^n_{i/n})^2}.
\]

By Theorem 3

\[
\sqrt{n}[V^n_0 - \sigma^2] \rightarrow \mathcal{L}(0, 2\sigma^4).
\]

The adjusted nominal 95% CI is as follows:

\[
V^n_0_{-CI} = \left[ V^n_0 - 1.96 \sqrt{2(V^n_0)^2/n}, V^n_0 + 1.96 \sqrt{2(V^n_0)^2/n} \right].
\]

To examine the performance of the volatility estimators \( V^n \) and \( V^n_0 \), we perform the following simulation study.

We simulate sample paths from (2) with \( \mu = 0, \sigma = 0.01 \). We examine two price levels, with starting prices of \( S_0 = \$10 \) and \( S_0 = \$50 \), respectively. At each price level, 10,000 simulations were conducted for a one-day period. We use a fixed rounding level of \( \alpha_n \equiv \alpha = 0.01 \), to mimic the financial market, in which stock prices are often rounded to the cent.
Table 1 Performance of the nominal 95% CIs based on $V_n$ and $V_{n0}$ for stocks priced at approximately $10 (S_0 = $10)

<table>
<thead>
<tr>
<th>samp. freq.</th>
<th>samp. intv.</th>
<th>‘$\beta_n$’</th>
<th>$V_n$ CI</th>
<th>$V_{n0}$ CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>78</td>
<td>5 min</td>
<td>0.088</td>
<td>f: 94.29%</td>
<td>89.57%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: 7.02 x 10^{-5}</td>
<td>6.20 x 10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: 1.18 x 10^{-5}</td>
<td>-1.21 x 10^{-6}</td>
</tr>
<tr>
<td>130</td>
<td>3 min</td>
<td>0.114</td>
<td>f: 78.59%</td>
<td>87.78%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: 5.87 x 10^{-5}</td>
<td>4.81 x 10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: 2.06 x 10^{-5}</td>
<td>-1.02 x 10^{-6}</td>
</tr>
<tr>
<td>195</td>
<td>2 min</td>
<td>0.140</td>
<td>f: 31.29%</td>
<td>85.08%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: 5.23 x 10^{-5}</td>
<td>3.94 x 10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: 3.18 x 10^{-5}</td>
<td>-6.57 x 10^{-7}</td>
</tr>
<tr>
<td>390</td>
<td>1 min</td>
<td>0.197</td>
<td>f: 0</td>
<td>74.75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: 4.61 x 10^{-5}</td>
<td>2.79 x 10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: 6.4 x 10^{-5}</td>
<td>-6.71 x 10^{-7}</td>
</tr>
<tr>
<td>780</td>
<td>30 sec</td>
<td>0.279</td>
<td>f: 0</td>
<td>46.91%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: 4.44 x 10^{-5}</td>
<td>1.85 x 10^{-5}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: 1.23 x 10^{-4}</td>
<td>-6.54 x 10^{-6}</td>
</tr>
</tbody>
</table>

"f": actual coverage frequency of the CIs; "l": average CI length; "b": finite sample bias.

Tables 1 and 2 show the simulation results. The first columns show the sample frequencies (samp. freq.), the second columns show the corresponding sample intervals (samp. intv.), and the third columns show the pre-limiting $\beta_n = \alpha \sqrt{n}$, demonstrating how the proposed small rounding asymptotic theory functions at a finite sample size and fixed rounding level. The final two columns display three items each. The notation “f” denotes the “actual coverage frequency,” which is used to record the frequency at which the true parameter is covered by the CIs based on the corresponding volatility estimators $V_n$ and $V_{n0};$ “l” denotes the “average length of the CI,” which indicates the CI width; and “b” denotes the “finite sample bias,” which indicates how much and in which direction the estimators are biased.

Comparing $V_n$ with $V_{n0}$ indicates that when the sample frequency is relatively low (e.g., a 5-min sampling interval for stocks priced at approximately $10, or 1–5 min for stocks priced at approximately $50), both $V_n$ and $V_{n0}$ perform well because their actual coverage frequency is near the nominal rate of 95%. This is consistent with the empirical evidence that subsampling is beneficial. But since the convergence rate is the square root of $n$, the CIs are wide when $n$ is small. When the sample frequency increases slightly (e.g., 3 min - 1 min for stocks priced at approximately $10 or 30 sec - 20 sec for stocks priced at approximately $50), the problems with the RV become apparent, the coverage frequency decreases from approximately 95% to some much lower rates (or even zero in the former case), whereas the $V_{n0}$ CI continues to exhibit a large coverage frequency. The biases demonstrate that the RV goes to some values much larger than the true value,
Table 2 Performance of the nominal 95% CIs based on $V_n$ and $V_n^0$ for stocks priced at approximately $50 (S_0 = \$50)$

<table>
<thead>
<tr>
<th>samp. freq.</th>
<th>samp. intv.</th>
<th>$\beta$</th>
<th>$\alpha \sqrt{n}$</th>
<th>$V_n$ CI</th>
<th>$V_n^0$ CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>78</td>
<td>5 min</td>
<td>0.088</td>
<td>f: 92.89%</td>
<td>92.48%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $6.23 \times 10^{-5}$</td>
<td>$6.19 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $-7.51 \times 10^{-7}$</td>
<td>$-1.27 \times 10^{-6}$</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>3 min</td>
<td>0.114</td>
<td>f: 94.01%</td>
<td>93.49%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $4.86 \times 10^{-5}$</td>
<td>$4.82 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $9.39 \times 10^{-8}$</td>
<td>$-7.73 \times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>195</td>
<td>2 min</td>
<td>0.140</td>
<td>f: 94.83%</td>
<td>93.86%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $3.99 \times 10^{-5}$</td>
<td>$3.94 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $6.26 \times 10^{-7}$</td>
<td>$-6.74 \times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>390</td>
<td>1 min</td>
<td>0.197</td>
<td>f: 94.9%</td>
<td>93.81%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $2.87 \times 10^{-5}$</td>
<td>$2.80 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $2.31 \times 10^{-6}$</td>
<td>$-2.95 \times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>780</td>
<td>30 sec</td>
<td>0.279</td>
<td>f: 86.01%</td>
<td>93.45%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $2.08 \times 10^{-5}$</td>
<td>$1.98 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $5.08 \times 10^{-6}$</td>
<td>$-1.23 \times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>1170</td>
<td>20 sec</td>
<td>0.342</td>
<td>f: 60.83%</td>
<td>93.12%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $1.74 \times 10^{-5}$</td>
<td>$1.62 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $7.66 \times 10^{-6}$</td>
<td>$-1.37 \times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>2340</td>
<td>10 sec</td>
<td>0.484</td>
<td>f: 0.27%</td>
<td>31.45%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>l: $1.32 \times 10^{-5}$</td>
<td>$1.23 \times 10^{-5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b: $1.55 \times 10^{-5}$</td>
<td>$7.72 \times 10^{-6}$</td>
<td></td>
</tr>
</tbody>
</table>

"f": actual coverage frequency of the CIs; "l": average CI length; "b": finite sample bias.

whereas the $V_n^0$ remains close to the true parameter value. Hence, $V_n^0$ substantially outperforms the uncorrected RV $V_n$.

At extremely high frequencies (less than 30 sec for $10 stocks or less than 10 sec for $50 stocks), the bias-corrected volatility estimator $V_n^0$ demonstrates decreased performance levels, although its bias remains substantially smaller compared with the RV). This is expected, because the bias-corrected estimator is built on asymptotic theory, which requires the condition $n \alpha \sqrt{n} \rightarrow 0$, which is hypothetical, since in practice one is faced with a fixed data set and a fixed tick size. If the sample frequency were to continue to increase at a fixed rounding level, the proposed bias correction would eventually fail. The failure at extremely high frequency is expected among other RV-based volatility estimators as well and is a direct consequence of Theorem 4 (Theorem 2 in [Li and Mykland (2007)] provides the result for the two scales realized volatility of [Zhang et al. (2005)]). The above simulation shows that for a given price level and rounding level, the proposed bias correction method is effective when the sample frequency is not excessively high.
2.2 Large Sampling Frequencies at a Fixed Rounding Level

At a fixed rounding level, when $n$ is excessively large, the conditions of Theorems 1–3 are no longer met and the feature described in Theorem 4 appears (Figures 3 and 4).

Figures 3 and 4 demonstrate that as the sampling frequency increases (sampling interval decreases), both the RV and the proposed bias corrected estimator $V^n_0$ will
Table 3  Performance of $V_n$ and $V_n^0$ when volatility is not a function of price

<table>
<thead>
<tr>
<th></th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Mean</th>
<th>Root Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_n - V$</td>
<td>$1.28\times10^{-4}$</td>
<td>$2.05\times10^{-4}$</td>
<td>$3.15\times10^{-4}$</td>
<td>$2.47\times10^{-4}$</td>
<td>$3.02\times10^{-4}$</td>
</tr>
<tr>
<td>$V_0^0 - V$</td>
<td>$-4.36\times10^{-5}$</td>
<td>$-1.02\times10^{-6}$</td>
<td>$1.74\times10^{-5}$</td>
<td>$-1.85\times10^{-5}$</td>
<td>$6.98\times10^{-5}$</td>
</tr>
</tbody>
</table>

The estimation errors are summarized by their 1st quartile, median, 3rd quartile, mean, and root mean squares.

rapidly increase as described in Theorem 4. Figure 4 most clearly exhibits the rate of divergence. However, the $V_0^0$ demonstrates a clear advantage over the RV at a large range of moderate-sized sampling intervals (20s - 2 min in the case illustrated in Figure 4).

2.3 When Conditions Deviate From the Requirements

2.3.1 When volatility is not a function of price. The theoretical results are established under conditions specified in Section 1; it is worth investigating how the bias correction performs if the required conditions are not met. Therefore, we conduct simulations based on a stochastic model in which the volatility process evolves by itself and is not a function of the price process. The Heston model \cite{Heston} was adopted to determine the log price:

\[
dX_t = (\mu - \nu/2)dt + \sigma dB_t \\
d\nu_t = \kappa (\eta - \nu_t)dt + \gamma \nu_t^{1/2} dW_t
\]

where $\nu_t = \sigma_t^2$. $B$ and $W$ are standard Brownian motions with $E(dB_t dW_t) = \rho dt$, and the parameters $\mu, \eta, \kappa, \gamma, \rho$, and the starting log-price $X_0$ are set at 0.05/252, 0.1/252, 5/252, 0.5/252, -0.5, and log(9), respectively. \cite{Aït-Sahalia and Kimmel} and \cite{Aït-Sahalia et al.} were referenced when selecting these parameter values.

We used a moderate leverage effect parameter $\rho = -0.5$ to represent an individual stock. We simulate 10,000 days and obtained pairs of the latent observations $X_{t_i}, \sigma_{t_i}$ for $t_0 = 0, t_1 = \frac{1}{365}, \ldots, t_n = 1$ for each day (one observation per minute, $n = 390$). We compute the integrated volatility $V = n^{-1} \sum_{i=1}^{n} \sigma_{t_i}^2$, and use this value as the reference measure. The observed log prices are $\log(\exp(X_{t_i}^{(a)}))$ where $a = 0.01$ (rounded to cent). We compute the RV $V^n$ and the proposed bias-corrected estimator $V_0^n$; and summarize their performance in Table 3.

The results indicate that although this model fails to meet the conditions for the theoretical results, the estimator $V_0^n$ demonstrates a clear advantage.

2.3.2 When rounding to the nearest multiple of $\alpha$. As mentioned in Remark 5, rounding down and rounding to the nearest multiple of $\alpha$ should yield
Table 4  Performance of $V^n$ and $V^n_0$ when rounding to the nearest multiple of $\alpha$

<table>
<thead>
<tr>
<th></th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Mean</th>
<th>Root Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V^n - V$</td>
<td>$1.28 \times 10^{-4}$</td>
<td>$2.01 \times 10^{-4}$</td>
<td>$3.13 \times 10^{-4}$</td>
<td>$2.45 \times 10^{-4}$</td>
<td>$3.00 \times 10^{-4}$</td>
</tr>
<tr>
<td>$V^n_0 - V$</td>
<td>$-4.54 \times 10^{-5}$</td>
<td>$-1.14 \times 10^{-6}$</td>
<td>$1.79 \times 10^{-5}$</td>
<td>$-1.89 \times 10^{-5}$</td>
<td>$7.09 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

The estimation errors are summarized by their 1st quartile, median, 3rd quartile, mean, and root mean squares.

2.3.3 When jumps exist. We further investigate volatility estimation in the presence of jumps. We simulated the following model:

$$dX_t = (\mu - \nu_t/2)dt + \nu_t^{1/2}dB_t + J_t dN_t$$  \hspace{1cm} (4)
$$d\nu_t = \kappa (\eta - \nu_t)dt + \gamma \nu_t^{1/2}dW_t,$$  \hspace{1cm} (5)

where $B_t$ and $W_t$ are Brownian motions with correlation $\rho$, $N_t$ is a Poisson process with intensity $\lambda$, and $J_t$ denotes the jump size, which is assumed to follow an independent $N(0, \sigma_J^2)$. The prices are again rounded to cents: $Y_i/n = \log(\exp(X_i/n)/\alpha)$ $(i=1, \ldots, n)$.

One way to remove the impact of jumps in volatility estimation is using the truncated RV which is defined as follows (Aït-Sahalia and Jacod, 2012)\footnote{See also Mancini (2001), Lee and Mykland (2008), and Jing et al. (2012). Bi- and multipower methods (Barndorff-Nielsen and Shephard, 2004, 2006) may also work, but we have not investigated this.}

$$V^{n, tr}_n = \sum_{i=1}^n (Y_{i/n} - Y_{(i-1)/n})^2 1\{|Y_{i/n} - Y_{(i-1)/n}| \leq a\pi - r\}$$  \hspace{1cm} (6)

for some $\pi \in (0, 1/2)$ and $a > 0$.

We define

$$V^{n, tr}_n = V^{n, tr}_0 - \frac{\alpha^2}{6} \sum_{i=1}^n \left(\exp(Y_{i/n})\right)^2$$

as the bias-corrected version of the truncated RV. The parameters $\eta = 0.1$, $\gamma = 0.5/252$, $\kappa = 5/252$, $\rho = -0.5$, $\mu = 0.05/252$, $\lambda = 5$, $\sigma_J = 0.015$, and $\alpha = 0.01$.
Table 5 Performance of $V_{n,lr}$ and $V_{0,lr}$ when jumps exist

<table>
<thead>
<tr>
<th></th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Mean</th>
<th>Root mean squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{n,lr} - V$</td>
<td>$1.47 \times 10^{-4}$</td>
<td>$2.24 \times 10^{-4}$</td>
<td>$3.34 \times 10^{-4}$</td>
<td>$2.59 \times 10^{-4}$</td>
<td>$3.53 \times 10^{-4}$</td>
</tr>
<tr>
<td>$V_{0,lr} - V$</td>
<td>$-8.46 \times 10^{-5}$</td>
<td>$-2.96 \times 10^{-5}$</td>
<td>$6.03 \times 10^{-5}$</td>
<td>$-7.12 \times 10^{-5}$</td>
<td>$2.29 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

The estimation errors are summarized by their 1st quartile, median, 3rd quartile, mean and root mean squares.

$x_0 = \log(9)$ were used, and the truncation level was set at $an^{-\alpha} = 4\sqrt{\alpha n^{-1/2}}$ (ref. Aït-Sahalia and Jacod 2012, Aït-Sahalia et al. 2013). The results are summarized in Table 5 which indicates that when jumps exist, the proposed bias correction method plays a significant role in reducing the bias of the truncated RV.

Remark 6: We notice that there is some deterioration in small sample behavior relative to the no jump-no truncation case. Naturally, the truncation leads to both higher bias and higher variance in small samples. The issue may relate to whether jumps get over-detected in small samples (Bajgrowicz et al. 2013), but also to the fact that one loses intervals that have continuous evolution (whether or not they have jumps). This latter problem has been discussed, with possible solutions, in Lee (2005) and Lee and Hannig (2010), and is beyond the scope of this article.

3 EMPIRICAL STUDY

To further illustrate the effectiveness of the proposed bias correction method, we conduct an empirical analysis of Citigroup Inc. (NYSE:C), CBS Corporation (NYSE:CBS), Dell Inc. (NYSE:DELL), Host Hotels and Resorts Inc. (NYSE:HST), and KeyCorp (NYSE:KEY) stock data from 2009. We collected the stock prices every minute (390 observations per day), computing the $V_n$ and $V_{0,n}$ values for each day. Based on the estimated volatilities, and the assumption that the return on each day is normally distributed, exhibiting approximately zero mean and variance as estimated (as is commonly assumed in risk management), we computed the 5% value at risk (VaR) for each day, counting the total number of days that the VaR was violated. Table 6 lists the VaR violations that occurred among the 252 days considered.

Because we considered the 5% VaRs, the expected rate of violation was 5%. The examined stocks based on the $V_{0,n}$ demonstrated rates closer to the expected rate than did those based on $V_n$. The $V_n$ tends to dramatically overestimate the daily volatilities, causing over-cautious VaRs.

See, for example, Christoffersen (2011). We used the estimated volatility of the same day for computing VaR of that day as our main purpose here is to examine the accuracy of the volatility estimators instead of to predict VaR.
Table 6 5% VaR violation rate based on the minute-by-minute stock prices of C, CBS, DELL, HST, and KEY stocks in 2009

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>CBS</th>
<th>DELL</th>
<th>HST</th>
<th>KEY</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V^n$</td>
<td>1.59%</td>
<td>2.78%</td>
<td>3.57%</td>
<td>3.57%</td>
<td>2.38%</td>
</tr>
<tr>
<td>$V_0^n$</td>
<td>2.78%</td>
<td>3.17%</td>
<td>4.76%</td>
<td>3.97%</td>
<td>3.97%</td>
</tr>
</tbody>
</table>

4 CONCLUSIONS AND DISCUSSION

We have explored the estimation of the integrated volatility when rounding is the primary source of market microstructure noise. We established asymptotic results for the RV based on “small rounding” conditions. We proposed a bias-corrected estimator for which consistency and central limit theorems were established. Results were also presented for the case when “small rounding” conditions were not satisfied. The effectiveness and practicality of the proposed bias correction method was demonstrated in both simulation and empirical studies.

Note that while we work with observations on a time interval $[0, 1]$, results for the more general case of time interval $[0, T]$ are obtained by rescaling. The case of unequal observation times can be studied based on the methods of Jacod and Protter (1998) and Mykland and Zhang (2006).

APPENDIX

A.1 Preparation

We assume without loss of generality (see Section A.4 for further justification) that $\mu_t = 0$, in which case

$$d \log S_t = \sigma_t dW_t; \quad (A.1)$$

and that there exist nonrandom constants $L_\sigma, U_\sigma \in (0, \infty)$, such that

$L_\sigma \leq \sigma_t \leq U_\sigma$ for $t \in [0, 1]$.

More Notation:

$$A_m := \left\{ \omega \in \Omega : S_t(\omega)_{t \in [0, 1]} \in \left[ \frac{1}{m}, m \right] \right\};$$

$$B_n := \left\{ \omega \in \Omega : \max_{1 \leq i \leq n} \sqrt{n} \frac{S_i/n - S_{(i-1)/n}}{S_{(i-1)/n}} \leq 2 \log n \right\};$$

$$Y_{i,n} := \sqrt{n} \left( S_{(i-1)/n} - S_{i/n} \right);$$

$$U(n, \phi) := \frac{1}{n} \sum_{i=1}^{n} \phi \left( S_{(i-1)/n}^{(a_n)}, Y_{i,n} \right) \text{ for } \phi : \mathbb{R}^2 \to \mathbb{R}; \quad (A.2)$$
Lemma 1: \( P(B_n) \to 1 \) as \( n \to \infty \).

Proof. By (A.1),

\[
S_{i/n}/S_{(i-1)/n} = \exp\left( \int_{(i-1)/n}^{i/n} \sigma_s dW_s \right).
\]

Note that for any \( i = 1, 2, \ldots, n \),

\[
E\left( \exp\left( \sqrt{n} \int_{(i-1)/n}^{i/n} \sigma_s dW_s \right) \right) 
\leq E\left( \exp\left( \sqrt{n} \int_{(i-1)/n}^{i/n} \sigma_s dW_s - \frac{1}{2} n \int_{(i-1)/n}^{i/n} \sigma_s^2 ds + \frac{1}{2} U_t^2 \right) \right) 
= \exp\left( \frac{1}{2} U_t^2 \right).
\]

Hence for any \( a > 0 \)

\[
P\left( \int_{(i-1)/n}^{i/n} \sigma_s dW_s > a \right) 
= P\left( \exp\left( \sqrt{n} \int_{(i-1)/n}^{i/n} \sigma_s dW_s \right) > \exp\left( \sqrt{n} a \right) \right) 
\leq \exp\left( \frac{1}{2} U_t^2 \right) \exp\left( \sqrt{n} a \right)
\]

Therefore,

\[
P\left( \max_{1 \leq i \leq n} \left( \int_{(i-1)/n}^{i/n} \sigma_s dW_s \right) > \log(2\log n) \right)
= P\left( \max_{1 \leq i \leq n} \left( \frac{S_{i/n}}{S_{(i-1)/n}} - 1 \right) > 2\log n \right)
\leq P\left( \max_{1 \leq i \leq n} \left( \frac{S_{i/n}}{S_{(i-1)/n}} \right) > \frac{2\log n}{\sqrt{n}} + 1 \right)
= P\left( \max_{1 \leq i \leq n} \left( \int_{(i-1)/n}^{i/n} \sigma_s dW_s \right) > \log\left( \frac{2\log n}{\sqrt{n}} + 1 \right) \right).
\]
\[
\frac{\exp\left(\frac{1}{2} \frac{\varepsilon^2}{n}\right)}{\exp\left(\sqrt{n} \left(1 + \frac{2\log n}{\sqrt{n}} + 1\right)\right)} \to 0 \text{ as } n \to \infty.
\]

A parallel argument gives
\[
P\left(\max_{1 \leq i \leq n} \left(\sqrt{n} \left(1 - \frac{S_i}{n} \frac{S_{i-1}}{n}\right)\right) > 2\log n\right) \to 0 \text{ as } n \to \infty,
\]

hence the conclusion.

**Lemma 2:** If \(\sqrt{n} \alpha_n \to \beta \in [0, \infty)\), then for any \(m\), there exist \(N\) large and \(c_m \in (0, 1\] such that for all \(n \geq N\), \(i=0, 1, 2, \cdots, n\),
\[
S^{(n)}_{i/n} \geq c_m \text{ on } A_m.
\]

**Proof.** \(\forall i=0, 1, 2, \cdots, n\), 
\[
\frac{S^{(n)}_{i/n}}{n} \geq \frac{1}{m} \text{ on } A_m, \text{ and } \alpha_n \to 0 \text{ as } n \to \infty,
\]

hence the conclusion.

**Lemma 3:** Suppose that \(\beta_n = \sqrt{n} \alpha_n \to \beta \in [0, \infty)\), then for any fixed \(m > 0\),
\[
\sup_{\omega \in A_m \cap B_n} \frac{Y_{i,n}}{\sqrt{n} S^{(n)}_{(i-1)/n}} = O\left(\frac{\log n}{\sqrt{n}}\right).
\]

**Proof.** On \(A_m \cap B_n\),
\[
|Y_{i,n}| = \sqrt{n} |S_{i/n}^{(n)} - S_{(i-1)/n}^{(n)}| \leq \sqrt{n} |S_{i/n} - S_{(i-1)/n}| + 2\alpha_n \leq 2m \log n + 2\beta_n.
\]

By Lemma 2 one can find a \(c_m \in (0, \frac{1}{m}]\) such that for large \(n\), on \(A_m \cap B_n\),
\[
\frac{|Y_{i,n}|}{\sqrt{n} S^{(n)}_{(i-1)/n}} \leq \frac{2m \log n + 2\beta_n}{\sqrt{n} c_m}.
\]

Since \(\beta_n \to \beta < \infty\), the above inequality implies that for any fixed \(m\),
\[
\sup_{\omega \in A_m \cap B_n} \frac{Y_{i,n}}{\sqrt{n} S^{(n)}_{(i-1)/n}} = O\left(\frac{\log n}{\sqrt{n}}\right).
\]
Lemma 4: Let $\beta > 0$, then for all $\sigma, x > 0$,

$$\int_0^1 \int h(y) \left( \frac{\beta [u + y \sigma x / \beta]}{x} \right)^2 dy du = \sigma^2 + \frac{1}{x^2} \left( \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -2\pi^2 k^2 \sigma^2 \frac{x^2}{\beta^2} \right) \right).$$

Proof.

$$\int_0^1 \int h(y) \left( \frac{\beta [u + y \sigma x / \beta]}{x} \right)^2 dy du$$

$$= E \left( \frac{\beta [U + Y \sigma x / \beta]}{x} \right)^2, U \sim \text{unif}[0, 1], Y \sim N(0,1)$$

$$= \frac{\beta^2}{x^2} E \left( U + Y \sigma x / \beta \right)^2$$

$$= \frac{\beta^2}{x^2} E \left( U + Z \right)^2, Z \sim N \left( 0, \frac{\sigma^2 x^2}{\beta^2} \right)$$

$$= \frac{\beta^2}{x^2} E \left( E \left( U + Z \mid Z \right) \right)$$

$$= \frac{\beta^2}{x^2} \left( EZ^2 + E(1 - [Z]) \right)$$

$$= \sigma^2 + \frac{1}{x^2} \left( \frac{\beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -2\pi^2 k^2 \sigma^2 \frac{x^2}{\beta^2} \right) \right),$$

where $[z] = z - [z]$ is the fractional part of $z$.

The last equality above is proved by using the Fourier expansion for function $f(z) = [z] - [z]^2$. ■

A.2 Proof of Theorem

Recall that $V^n$ is defined in (3). For large $n$,

$$V^n I_{A_m \cap \mathcal{B}_n}$$

$$= \sum_{i=1}^{n} \left( \log S_{i/n}^{(\alpha)} - \log S_{(i-1)/n}^{(\alpha)} \right)^2 I_{A_m \cap \mathcal{B}_n}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left[ \sqrt{n} \log \left( \frac{S_{i/n}^{(\alpha)} - S_{(i-1)/n}}{S_{i/n}^{(\alpha)} + S_{(i-1)/n}} + 1 \right) \right]^2 I_{A_m \cap \mathcal{B}_n} \quad (A.3)$$
\[
\begin{align*}
&= \frac{1}{n} \sum_{i=1}^{n} \left[ \sqrt{n} \log \left( \frac{Y_{i,n}}{\sqrt{\pi S_i(n)})} + 1 \right) \right] I_{A_n^c \cap B_n} \\
&= \frac{1}{n} \sum_{i=1}^{n} \left[ \sqrt{n} \left( \frac{Y_{i,n}}{\sqrt{\pi S_i(n)})} - \frac{1}{2} \left( \frac{Y_{i,n}}{\sqrt{\pi S_i(n)})} \right)^2 + \frac{1}{3} \right) \right] I_{A_n^c \cap B_n},
\end{align*}
\]
for \( \theta \in \left( 0, \frac{Y_{i,n}}{\sqrt{\pi S_i(n)})} \right) \).

By Lemma 2, one can find \( c_m \in (0, \frac{1}{n}] \) such that for large \( n \), \( S_i^{(n)} \geq c_m \) for all \( i = 0, 1, 2, \cdots, n \). (A.4)

Define
\[
\phi_{cm}(x, y) = \begin{cases} 
\frac{y^2}{x}, & \text{when } x \geq c_m; \\
\frac{3}{4} x - \frac{8}{c_m x^2} + \frac{6}{c_m^2}, & \text{when } x < c_m.
\end{cases}
\]
Note in particular that \( \phi_{cm} \) is a function satisfying Hypothesis \( L_r \) in \cite{Delattre and Jacod 1997} with \( r = 2 \).

For \( n \) large enough, by Lemmas 2 and 3, (A.3) can be rewritten as
\[
V^n I_{A_n \cap B_n} \leq \frac{1}{n} \sum_{i=1}^{n} \phi_{cm}(S_i^{(n)}, Y_{i,n}) I_{A_n \cap B_n} + O \left( \frac{(\log n)^3}{n^{1/2}} \right) I_{A_n \cap B_n},
\]
where \( U(\cdot, \cdot) \) is defined in (A.2).

Furthermore,
\[
V^n I_{A_n} = V^n I_{A_n \cap B_n} + V^n I_{A_n \cap B_n}^c \\
\leq U(n, \phi_{cm}) I_{A_n} + (V^n - U(n, \phi_{cm})) I_{A_n \cap B_n}^c + O \left( \frac{(\log n)^3}{n^{1/2}} \right) I_{A_n \cap B_n},
\]
(\text{by Lemma 4}).

By Theorem 3.1 of \cite{Delattre and Jacod 1997},
\[
U(n, \phi_{cm}) \rightarrow p \left\{ \begin{cases} 
\int_0^1 \int_0^1 h(y)\phi_{cm}(S_t, \beta \{ u + y \sigma S_t / \beta \}) dy du dt, & \text{if } \beta > 0; \\
\int_0^1 h(y)\phi_{cm}(S_t, y \sigma S_t) dy dt, & \text{if } \beta = 0.
\end{cases} \right.
\]
Note that since $c_m \leq 1/m$, we have

$$
\phi_{cm}(S_{(i-1)/n}, Y) = \left( \frac{Y}{S_{(i-1)/n}} \right)^2 I_{Ac} + \phi_{cm}(S_{(i-1)/n}, Y)I_{Ac}.
$$

Lemma 4 gives, when $\beta > 0$,

$$
U(n, \phi_{cm})I_{Am} \to_P \int_0^1 \left( \frac{\sigma_t^2 S_t^2 + \beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -\frac{2\pi^2 k^2 \sigma_t^2 S_t^2}{\beta^2} \right) \right) dt I_{Am}.
$$

It is easy to check that the above convergence is also true when $\beta = 0$. Therefore, for $\beta \in [0, \infty)$,

$$
V'^n I_{Am} = U(n, \phi_{cm})I_{Am} + o_p(1)
$$

$$
\to_P \int_0^1 \left( \frac{\sigma_t^2 S_t^2 + \beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -\frac{2\pi^2 k^2 \sigma_t^2 S_t^2}{\beta^2} \right) \right) dt I_{Am}.
$$

That is to say, for any $\delta > 0$, $\epsilon > 0$, there exists $N$, such that for all $n > N$,

$$
P \left( \left| V'^n I_{Am} - \int_0^1 \left( \frac{\sigma_t^2 S_t^2 + \beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -\frac{2\pi^2 k^2 \sigma_t^2 S_t^2}{\beta^2} \right) \right) dt I_{Am} \right| > \delta \right) < \epsilon.
$$

On the other hand, since $A_m \not\subset \Omega$ as $m \to \infty$, there exists $M$ large, such that

$$
P(A_M^c) < \epsilon.
$$

Therefore, for $n > N$,

$$
P \left( \left| V^n - \int_0^1 \left( \frac{\sigma_t^2 S_t^2 + \beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -\frac{2\pi^2 k^2 \sigma_t^2 S_t^2}{\beta^2} \right) \right) dt \right| > \delta \right)
$$

$$
\leq P(A_M^c) +
$$

$$
P \left( \left| V^n I_{Am} - \int_0^1 \left( \frac{\sigma_t^2 S_t^2 + \beta^2}{6} - \frac{\beta^2}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \exp \left( -\frac{2\pi^2 k^2 \sigma_t^2 S_t^2}{\beta^2} \right) \right) dt I_{Am} \right| > \delta \right)
$$

$$
< 2\epsilon.
$$

This proves Theorem 1.
A.3 Proof of Theorem 2 and Theorem 3

By (A.3), for large $n$,

$$
\sqrt{n}V^n I_{\text{Am} \cap B_n} = \sum_{i=1}^{n} \sqrt{n} \left( \frac{Y_{i,n}}{\sqrt{E(a_n)}} - \frac{1}{2} \left( \frac{Y_{i,n}}{\sqrt{E(a_n)}} \right)^2 + \frac{1}{3} \theta^3 \right) I_{\text{Am} \cap B_n},
$$

(A.6)

for $\theta \in \left( 0, \frac{Y_{i,n}}{\sqrt{E(a_n)}} \frac{1}{(i-1)/n} \right)$.

Using the $c_m \in (0, \frac{1}{m}]$ as in (A.4), we define

$$
\psi_{c_m}(x, y) = \begin{cases} 
\left( \frac{y}{x} \right)^3, & \text{when } x \geq c_m; \\
\left( \frac{4}{c_m^3} - \frac{3x}{c_m^4} \right) y^3, & \text{when } x < c_m.
\end{cases}
$$

(A.7)

(A.6) can be further written as

$$
\sqrt{n}V^n I_{\text{Am} \cap B_n} \leq \sqrt{n} U(n, \phi_{c_m}) I_{\text{Am} \cap B_n} - U(n, \psi_{c_m}) I_{\text{Am} \cap B_n} + O\left( \frac{(\log n)^4}{n^{1/2}} \right) I_{\text{Am} \cap B_n};
$$

and

$$
\sqrt{n}V^n I_{\text{Am}} = \sqrt{n}V^n I_{\text{Am} \cap B_n} + \sqrt{n}V^n I_{\text{Am} \cap B_c} \
\leq (\sqrt{n} U(n, \phi_{c_m}) - U(n, \psi_{c_m})) I_{\text{Am}} \
+ (\sqrt{n} V^n - \sqrt{n} U(n, \phi_{c_m}) + U(n, \psi_{c_m})) I_{\text{Am} \cap B_c} + O\left( \frac{(\log n)^4}{n^{1/2}} \right) I_{\text{Am} \cap B_n} \
= \sqrt{n} U(n, \phi_{c_m}) I_{\text{Am}} - U(n, \psi_{c_m}) I_{\text{Am}} + o_p(1),
$$

where $\phi_{c_m}$ is defined in (A.5), $\psi_{c_m}$ in (A.7) and $U(\cdot, \cdot)$ in (A.2), and we have used Lemma 3 in the above.

Note that $\psi_{c_m}(S_t, \sigma_t S_y)$ is an odd function of $y$, and $\beta = 0$; by Theorem 3.1 of Delattre and Jacod (1997),

$$
U(n, \psi_{c_m}) \rightarrow p \int_0^1 \int h(y) \psi_{c_m}(S_t, \sigma_t S_y) dy dt = 0.
$$
Therefore, 
\[ U(n, \psi_{cm})|\Lambda_m \to p.0. \]

As a consequence, 
\[ \sqrt{n}V^m|\Lambda_m = \sqrt{n}U(n, \psi_{cm})|\Lambda_m + o_p(1). \] (A.8)

Also by Corollary 3.3 of Delattre and Jacod (1997), since \( \phi_{cm}(x, y) \) is even in \( y \), 
\[ \sqrt{n}[U(n, \psi_{cm}) - \int_0^1 \Gamma \phi_{cm}(S_t, \beta_n)dt] \] 
\[ \to \text{stably in law } \int_0^1 \Delta(\phi_{cm}, \phi_{cm})(S_t, 0)^{1/2}dB_t, \] (A.9)

where \( B \perp W, \) and 
\[ \Gamma \phi_{cm}(S_t, \beta_n) \]
\[ = \int_0^1 \int h(y)\phi_{cm}(S_t, \beta_n | u + y\sigma_t S_t / \beta_n)dydu \]
\[ = \int_0^1 \int h(y) \left( \frac{\beta_n | u + y\sigma_t S_t / \beta_n |}{S_t} \right)^2 I_{A_m} + \phi_{cm}(S_t, \beta_n | u + y\sigma_t S_t / \beta_n)I_{A_m} \right) dydu \]
\[ = \left( \sigma^2 + \frac{\beta_n^2}{6S_t^4} \frac{1}{\pi^2} \frac{1}{S_t^2} \sum_{k=1}^{\infty} \frac{1}{k} \exp \left( -\frac{2\pi^2k^2}{\beta_n^2} \right) \right) I_{A_m} + \]
\[ \int_0^1 \left( \phi_{cm}(S_t, \beta_n | u + y\sigma_t S_t / \beta_n)I_{A_m} \right) dy \]
\[ \right) \] (by Lemma 3) ;
\[ \Delta(\phi_{cm}, \phi_{cm})(S_t, 0) \]
\[ = \int h_{\eta S_t}(y)\phi_{cm}^2(S_t, y)dy - \left( \int h_{\eta S_t}(y)\phi_{cm}(S_t, y)dy \right)^2 \]
\[ = \int h_{\eta S_t}(y) \left( \frac{y}{S_t} \right)^4 I_{A_m} + \phi_{cm}^2(S_t, y)I_{A_m} \right] dy \]
\[ - \left( \int h_{\eta S_t}(y) \left( \frac{y}{S_t} \right)^2 I_{A_m} + \phi_{cm}(S_t, y)I_{A_m} \right) dy \right)^2 \]
\[ = \left[ \int h_{\eta S_t}(y) \left( \frac{y}{S_t} \right)^4 dy - \left( \int h_{\eta S_t}(y) \left( \frac{y}{S_t} \right)^2 dy \right)^2 \right] I_{A_m} \]
\[ + \left[ \int h_{\eta S_t}(y)\phi_{cm}(S_t, y)^2dy - \left( \int h_{\eta S_t}(y)\phi_{cm}(S_t, y)dy \right)^2 \right] I_{A_m}; \]
hence
\[
\Delta(\phi_{cm}, \phi_{cm})(S_t, 0)^{1/2} = \left[ \int h_{n,S_t}(y) \left( \frac{y}{S_t} \right)^4 dy - \left( \int h_{n,S_t}(y) \left( \frac{y}{S_t} \right)^2 dy \right)^2 \right]^{1/2} I_{Am} + \left[ \int h_{n,S_t}(y) \phi_{cm}(S_t, y)^2 dy - \left( \int h_{n,S_t}(y) \phi_{cm}(S_t, y) dy \right)^2 \right]^{1/2} I_{Acm} \tag{A.11}
\]

Plug (A.10) and (A.11) into (A.9), and note that by the assumption that \( \beta_n = O(n^{-\gamma}) \),
\[
\sqrt{n} \left[ \frac{\beta_n}{\pi^2} \int \sum_{k=1}^{\infty} \int h_{n,S_t}(y) \phi_{cm}(S_t, y)^2 dy - \left( \int h_{n,S_t}(y) \phi_{cm}(S_t, y) dy \right)^2 \right]^{1/2} dB_s I_{Acm},
\]
where \( Z \sim \int_0^1 (2\sigma_t^4)^{1/2} dB_s, B \perp W \).

For any continuous function \( g \) that vanishes outside a compact set, the above stable convergence implies that \( \forall E \in \mathcal{F} \),
\[
E \left[ g \left( \sqrt{n} \left[ \int h_{n,S_t}(y) \phi_{cm}(S_t, y)^2 dy - \left( \int h_{n,S_t}(y) \phi_{cm}(S_t, y) dy \right)^2 \right]^{1/2} dB_s I_{Acm} \right) \right] \rightarrow E \left[ g \left( \int_0^1 (2\sigma_t^4)^{1/2} dB_s I_{Acm} \right) \right]. \tag{A.12}
\]
And by defining $\eta_{cm}(\cdot, \cdot)$ to be

$$
\eta_{cm}(x, y) = \begin{cases} 
\left(\frac{1}{x}\right)^2, & \text{when } x \geq c_m; \\
\frac{3}{c_m^2}x^2 - \frac{8}{c_m^3}x + \frac{6}{c_m^4}, & \text{when } x < c_m.
\end{cases}
$$

one has,

$$
V_0^n I_{\text{Am}} = V_0^n I_{\text{Am}} - \frac{\beta_n^2}{6} U(n, \eta_{cm}) I_{\text{Am}}. \tag{A.13}
$$

Again, by Theorem 3.1 of Delattre and Jacod (1997),

$$
U(n, \eta_{cm}) I_{\text{Am}} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{S}{\sqrt{n}} \right)^2 I_{\text{Am}} \to_p \int_0^1 \frac{1}{S_i^2} dt I_{\text{Am}}.
$$

and

$$
\sqrt{n} \left( \frac{\beta_n^2}{6} U(n, \eta_{cm}) - \frac{\beta_n^2}{6} \int_0^1 \frac{1}{S_i^2} dt I_{\text{Am}} \right) I_{\text{Am}} = O_p(\beta_n^2) = o_P(1). \tag{A.14}
$$

By (A.13), (A.14), and (A.14),

$$
\sqrt{n} V_0^n I_{\text{Am}} = \sqrt{n}(U(n, \phi_{cm}) - \frac{\beta_n^2}{6} \int_0^1 \frac{1}{S_i^2} dt I_{\text{Am}} + o_P(1).
$$

Also since that $g$ is uniformly continuous, $\forall E \in \mathcal{F}$,

$$
\lim_{n \to \infty} E \left[ g\left( \sqrt{n}(V_0^n - \int_0^1 \sigma_t^2 dt) I_{\text{Am}} \right) \right] = \lim_{n \to \infty} E \left[ g\left( U(n, \phi_{cm}) - \left( \int_0^1 \sigma_t^2 dt + \frac{\beta_n^2}{6} \int_0^1 \frac{1}{S_i^2} dt \right) I_{\text{Am}} \right) \right] = E \left[ g\left( \int_0^1 (2\sigma_t^4)^{1/2} dB_t I_{\text{Am}} \right) \right] \text{ (by (A.12))},
$$

which implies, for any $\epsilon > 0$, there exists $N$, such that $\forall n \geq N$,

$$
E \left[ g(\sqrt{n}(V_0^n - \sigma_t^2) I_{\text{Am}}) I_{\text{Am}} \right] = E \left[ g\left( \int_0^1 (2\sigma_t^4)^{1/2} dB_t I_{\text{Am}} \right) \right] < \epsilon.
$$

Note also that $g$ is bounded, suppose $|g| \leq M_g$. Recall that $P(A_M^c) \to 0$, one can choose $M$ such that $P(A_M^c) < \epsilon/M_g$. 


So for $n \geq N$, 
\[
\left| E[g(\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt])I_E] - E \left[ g \left( \int_0^1 (2\sigma_t^4)^{1/2} dB_t \right) I_E \right] \right| 
\leq \left| E[g(\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt])I_{A_M}I_E] - E \left[ g \left( \int_0^1 (2\sigma_t^4)^{1/2} dB_t I_{A_M} \right) I_{A_M}I_E \right] \right| 
+ 2M_g \ast P(A_M) \leq 3\epsilon 
\]

Hence we’ve proved that for all continuous function $g$ that vanishes outside a compact set, $\forall E \in \mathcal{F}$,
\[
\lim_{n \to \infty} E[g(\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt])] = E \left[ g \left( \int_0^1 (2\sigma_t^4)^{1/2} dB_t \right) \right].
\]

i.e.,
\[
\sqrt{n}[V_0^n - \int_0^1 \sigma_t^2 dt] \to \mathcal{L} - \text{stably} \int_0^1 (2\sigma_t^4)^{1/2} dB_t.
\]

This finishes the proof of Theorem 3. The proof of Theorem 2 is basically contained in the proof above.

### A.4 The Case of General $\mu_t$ and $\sigma_t$

#### Step 1:
For general cases when $\mu_t \neq 0$, if there exists $L_\sigma$, $U_\sigma$, $C_\mu \in (0, \infty)$, such that $L_\sigma \leq \sigma_t \leq U_\sigma$ and $|\mu_t| \leq C_\mu$ for $t \in [0, 1]$, the previous results all hold.

For the simplicity of notation, we consider the log scale. Let $P$ be the probability measure corresponding to the system
\[dX_t = \sigma_t dW_t \]
and $Q$ the probability measure corresponding to the system
\[dX_t = \mu_t dt + \sigma_t dW_t^Q, \]
where $W_t$ and $W_t^Q$ are standard Brownian motions under $P$ and $Q$ respectively.

Note that by the Girsanov Theorem (see, for example, page 164 of [Øksendal (2003)]), for bounded $\sigma_t$ and $\mu_t$ (as stated in the conditions of “Step 1”), $P$ and $Q$ are mutually absolutely continuous.

The following proposition justifies the conclusion of “Step 1”.

---

LI & MYKLAND | Rounding Errors and Volatility Estimation 501
Proposition 1 (Mykland and Zhang (2009)) Suppose that $\zeta_n$ is a sequence of random variables which converges stably to $N(b, a^2)$ under $P$ (meaning that $N(b, a^2) = b + aN(0, 1)$, where $N(0, 1)$ is a standard normal variable independent of $F$, also $a$ and $b$ are $F$ measurable). Then $\zeta_n$ converges stably in law to $b + aN(0, 1)$ under $Q$, where $N(0, 1)$ remains independent of $F$ under $Q$.

Step 2: for locally bounded $\sigma_t$ and $\mu_t$, the stable convergence and the convergence in probability stay valid.

This can be proved by a localization argument which uses essentially the same techniques as in the derivation in the last part of Section A.3. For example, to unbound $\sigma_t$, one considers a sequence of stopping times $\tau_m$ corresponding to a sequence of positive constants $\sigma_m$ which increases to infinity as $m \to \infty$: $\tau_m = \min\{t: \sigma_t^2 \geq \sigma_m^2\}$, and note the fact that the sets $\{\tau_m > T\} \not\in \Omega$.

In particular, the locally bounded assumption is automatically satisfied when $\sigma_t$ and $\mu_t$ are continuous.

A.5 Proof of Theorem

Similar argument as the Proof of Theorem 3 in Li and Mykland (2007) gives the result.

Received 31 December 2009; revised December 30, 2013; accepted January 25, 2014.

REFERENCES


