Realized Candlestick Wicks*

Yifan Li† Ingmar Nolte‡ Sandra Nolte§ Shifan Yu¶

This Version: January 12, 2024

Abstract

We develop a novel nonparametric estimator of integrated variance that utilizes intraday candlestick information, comprised of the high, low, open, and close prices within short time intervals. The range-return-difference volatility (RRDV) estimator is robust to short-lived extreme return persistence hardly attributable to the diffusion component, such as gradual jumps and flash crashes. By modelling such sharp but continuous price movements following some recent theoretical advances, we demonstrate that RRDV can provide consistent estimates with variances about four times smaller than those obtained with the differenced-return volatility (DV) estimator. Monte Carlo simulations and empirical applications further validate the practical reliability of our proposed estimator with some finite-sample refinements.

JEL Classifications: C14, C22, C58, G14

Keywords: High-Frequency Data, Integrated Variance, Extreme Return Persistence, Drift Burst, Range-Based Volatility Estimation

*We thank Torben Andersen, Tim Bollerslev, Kim Christensen, Wenhao Cui, Dick van Dijk, Dobrislav Dobrev, Roxana Halbleib, David Harvey, Seok Young Hong, Steve Leybourne, Qiyuan Li, Aleksey Kolokolov, Manh Cuong Pham, Roberto Renò, Neil Shephard, Ruixun Zhang (discussant), as well as participants at the Financial Econometrics Conference in Honour of Stephen J. Taylor, QFFE 2023, IAAE 2023, Asian Meeting of the Econometric Society 2023 in both China and Singapore, EcoSta 2023, 3rd International Econometrics PhD Conference, CFE-CMStatistics 2023, and seminars at various institutions, for helpful comments and suggestions. This paper was previously circulated under the title "Nonparametric Range-Based Estimation of Integrated Variance with Episodic Extreme Return Persistence". We would like to acknowledge financial support from the Economic and Social Research Council (ESRC) North West Social Science Doctoral Training Partnership [Grant Number ES/P000665/1]. The usual disclaimer applies.

†Division of Accounting and Finance, Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, United Kingdom. Email: yifan.li@manchester.ac.uk.

‡Department of Accounting and Finance, Lancaster University Management School, Lancaster LA1 4YX, United Kingdom. Email: i.nolte@lancaster.ac.uk.

§Department of Accounting and Finance, Lancaster University Management School, Lancaster LA1 4YX, United Kingdom. Email: s.nolte@lancaster.ac.uk.

¶Corresponding author, Department of Accounting and Finance, Lancaster University Management School, Lancaster LA1 4YX, United Kingdom. Email: s.yu7@lancaster.ac.uk.

1 Introduction

The discussion about intraday periods with extreme high-frequency return persistence was brought back to the fore by the May 2010 "flash crash" in the U.S. stock market (Kirilenko et al., 2017; Menkveld and Yueshen, 2019). The crash originated in E-mini S&P 500 future contracts, and led to an extraordinarily rapid decline by 5-6% and a V-shaped recovery of U.S. equity indices in 30 minutes. It swiftly spread to almost 8,000 individual stocks and exchange traded funds (ETFs), and echoed internationally (CFTC and SEC, 2010). Prices with short-lived locally explosive trends and returns with highly positive autocorrelations exhibit compelling short-horizon predictability. "Gradual jumps" identified by Barndorff-Nielsen et al. (2009) also have similar characteristics. These sharp but "continuous" price movements explain to a large extent the reason for spurious detection of jumps with sparsely sampled data (Christensen et al., 2014; Bajgrowicz et al., 2016), although they have attracted limited attention. Empirical evidence shows such extreme events like mini flash crashes occur more frequently in recent years, which raises widespread concern about market inefficiency and vulnerability (Golub et al., 2012; Laly and Petitjean, 2020; Flora and Renò, 2022). These market glitches are also a threat to the standard theoretical framework, as a temporary violation of the Itô semimartingale assumption and potentially the no-arbitrage principle (Andersen et al., 2023). Two recent influential studies, Christensen et al. (2022) and Andersen et al. (2021), attempt to incorporate the mechanism behind these short-term directional and persistent price movements into the standard Itô semimartingale framework. Christensen et al. (2022) attribute the short-lived explosive trend to a locally unbounded drift, which prevails over volatility and dominates log-returns in the vicinity of explosion points.[1] Andersen et al. (2021) consider these unusual patterns as outcomes of the temporary disequilibrium after ambiguous information arrivals, i.e., the market price over- or under-reacts to information in an inefficient financial market, and deviates temporarily from the true value.

The existence of such events poses new challenges for the estimation of integrated variance (IV), which serves as the cornerstone of statistical inference with high-frequency financial data (Aït-Sahalia and Jacod, 2014). Since the realized volatility (RV) estimator of Andersen and Bollerslev (1998), the increased data availability motivates the development of nonparametric estimation techniques to mitigate the impact of distinctive data characteristics, either in isolation or in combination. A stream of literature focuses on robust IV estimation when the price process has jumps. There are basically two methods to overcome this problem: the bipower and multipower estimators (Barndorff-Nielsen and Shephard, 2004, 2006; Huang and Tauchen, 2005) and truncated estimators (Mancini, 2009), with some combinations thereof (Vetter, 2010; Corsi et al., 2010). Theoretical innovations on this issue continue to emerge afterwards, see, e.g., Andersen et al. (2012) and Jacod and Todorov (2014, 2018). All the aforementioned tools spotlight merely extreme price movements characterized by a discontinuous component, while the distortion of IV measurement by non-trivial periods with sharp

[1] See also Flora and Renò (2022), Laurent et al. (2022), Bellia et al. (2023), Christensen and Kolokolov (2023), and Kolokolov (2023), for some recent theoretical and empirical studies on drift burst.

but continuous price movements has long been ignored. Laurent and Shi (2020) visualize the bias of (original and modified) RV and realized bipower variation (RBPV) in the presence of a nonzero drift coefficient. The differenced-return volatility (DV) estimator of Andersen et al. (2021) is the first IV estimator robust to this type of episodic Itô semimartingale violation.

This paper develops an alternative nonparametric estimator which can exclude excessive return variation induced by short-lived dominant trends and consistently estimate IV from the diffusion component. Motivated by the drift-independent variance estimator of Yang and Zhang (2000), we propose the range-return-difference volatility (RRDV) estimator based on intraday "candlestick" information, specifically the high, low, open, and close prices (HLOCs) within short time intervals. The RRDV estimator is constructed from pairwise differences between high-low ranges and absolute open-close returns. This construction based on range-return differences is designed to remove the contribution to total variation from a locally persistent component, which dominates both ranges and returns in the intervals within such non-trivial episodes. Fig. 1 illustrates some simulated examples of intraday candlestick charts. The candlesticks have long "real bodies" but small or nearly no "wicks", i.e., *marubozu* candlesticks, when the price movements are locally dominated by either discontinuities or short-lived explosive trends. The pairwise offset between ranges and returns offers the wick-based RRDV estimator a built-in robustness to such extreme events.

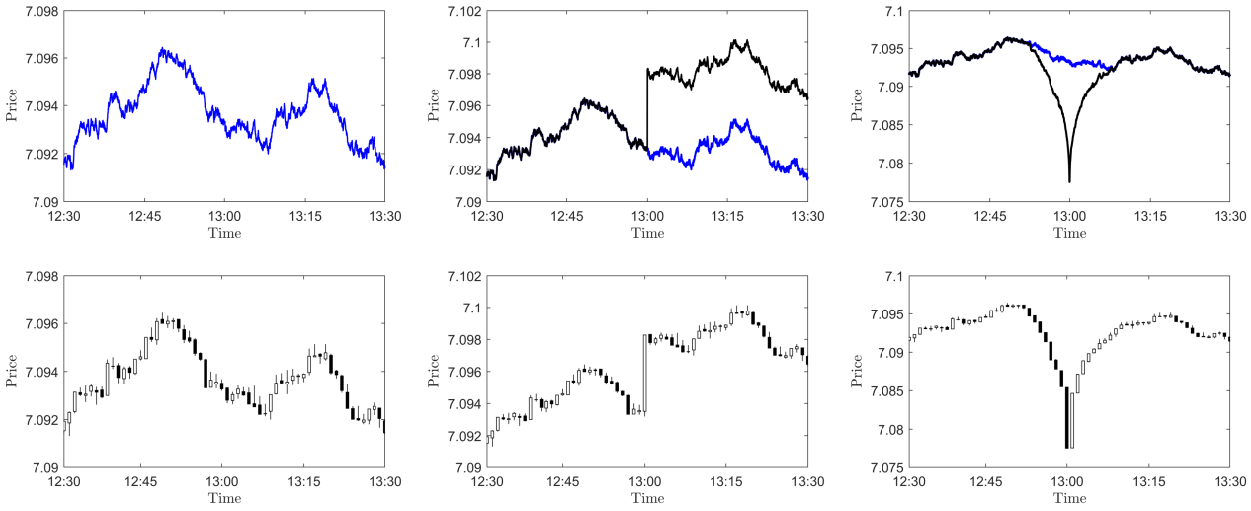

Figure 1: Examples of one-minute intraday candlestick charts for the simulated second-by-second log-prices from different DGPs: continuous (left), discontinuous with a jump (middle), continuous with a V-shaped flash crash (right).

Different from the DV estimator that is constructed from the first-order differenced returns to offset the excessive return drift in consecutive intervals, RRDV utilizes the candlestick information and implements a similar offset independently within each interval. We derive the consistency and asymptotic normality of the RRDV estimator under infill asymptotics, which reveals its ability to provide consistent IV estimates with variances approximately four times smaller than those obtained with DV. Importantly, we demonstrate that the presence of jumps with both finite and infinite activities, as well as the episodes of extreme return persistence as modelled by Christensen et al. (2022) and Andersen et al. (2021), has no impact on the consistency and asymptotic distribution.

Simulation results confirm that our new estimator outperforms selected competitors in scenarios with various specifications of extreme directional price movements, with an effective correction for the finite-sample biases. Our empirical application focuses on the prediction of out-of-sample IV estimates of the SPDR S&P 500 ETF Trust (SPY) with the heterogeneous autoregressive (HAR) model of Corsi (2009). We find that the HAR model based on RRDV estimates can achieve smaller forecast errors for both robust and non-robust IV measures than all selected benchmark models, especially on days with continuous or discontinuous extreme events.

From a technical point of view, our estimator is more closely related to the literature on range-based volatility estimation. Since the classical works of Parkinson (1980) and Garman and Klass (1980), a number of studies in this field show the strength of range-based volatility estimators to improve estimation accuracy by extracting more information from realized price paths than return-based measures, see, e.g., Beckers (1983), Ball and Torous (1984), Rogers and Satchell (1991), Kunitomo (1992), Yang and Zhang (2000), Alizadeh et al. (2002), and Brandt and Diebold (2006). The realized range-based volatility (RRV) estimator introduced by Christensen and Podolskij (2007) is the first nonparametric IV measure constructed from high-frequency intraday ranges, which is then extended by Martens and van Dijk (2007), Christensen et al. (2009), and Christensen and Podolskij (2012). More recently, Li et al. (2022) and Bollerslev et al. (2023) introduce the optimal candlestick-based spot volatility estimators with the linear and nonlinear functional forms, respectively, which benefit from both the statistical superiority of range-based estimation and the broader availability of intraday candlestick charts. As an fundamental tool in technical analysis that predates the rise of high-frequency data, easy access to intraday candlesticks is now widely available in most online trading applications. This accessibility facilitates the straightforward implementation of candlestick-based inference techniques, rendering them a convenient option for general investors.

The remainder of this paper is structured as follows: Section 2 lays out the basic setup and introduces the new candlestick-based IV estimator. Section 3 discusses its asymptotic behavior under two different specifications of episodic extreme return persistence. Section 4 contains finite-sample refinements of our estimator and instructions in practice. Section 5 includes an extensive Monte Carlo study that verifies its asymptotic unbiasedness and illustrates the finite-sample performance. After an empirical application of volatility forecasting in Section 6, we conclude in Section 7. All proofs and additional materials are relegated to the Appendix.

2 Volatility Estimation Based on Range-Return Differences

2.1 Range-Return-Difference Volatility (RRDV) Estimator

For a finite time interval $[0, t]$, e.g., a trading day, we apply an equidistant partition at $0 < \Delta_n < 2\Delta_n < \cdots < n\Delta_n \leq t$ to divide it into $n = \lfloor t/\Delta_n \rfloor$ short time intervals. We denote the i-th interval

by $I_{n,i} = [(i-1)\Delta_n, i\Delta_n]$. The HLOC over the i-th interval can be expressed respectively as

$$H_i = \sup_{t \in I_{n,i}} X_t, \qquad L_i = \inf_{t \in I_{n,i}} X_t, \qquad O_i = X_{(i-1)\Delta_n}, \qquad C_i = X_{i\Delta_n}. \tag{1}$$

The high-low range and open-close return are then denoted by

$$w_i = H_i - L_i, \qquad r_i = C_i - O_i. \tag{2}$$

The range-return-difference volatility (RRDV) estimator based on the differences between ranges and absolute returns is defined as

$$\widehat{V}_{t,n} = \frac{1}{\Lambda_2} \sum_{i=1}^{n} \left(w_i - |r_i| \right)^2, \tag{3}$$

with

$$\Lambda_p = \mathbb{E}\left[\left(\sup_{t,s \in [0,1]} W_t - W_s - |W_1| \right)^p \right], \tag{4}$$

where $W = (W_t)_{t \geq 0}$ is a standard Brownian motion, and $\Lambda_2 = 4 \ln 2 - 2 \approx 0.7726$, specifically.

2.2 Limit Theorems for Continuous Itô Semimartingales

We consider a continuous Itô semimartingale in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \tag{5}$$

where t stands for time, X_0 is $\mathcal{F}_0$-measurable, $\mu = (\mu_t)_{t \geq 0}$ is a locally bounded and predictable process of drift, $\sigma = (\sigma_t)_{t \geq 0}$ is an adapted, càdlàg and strictly positive (almost surely) process of spot volatility, and $W = (W_t)_{t \geq 0}$ is a standard Brownian motion.

Theorem 1 (Consistency)**.** Assume that the efficient price X evolves according to Eq. (5) with all traditional conditions satisfied. Then it holds that as $\Delta_n \to 0$,

$$\widehat{V}_{t,n} \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^2 ds, \tag{6}$$

where $\xrightarrow{\text{u.c.p.}}$ stands for the uniform convergence in probability, i.e., for any processes Z^n, Z we have $Z^n \xrightarrow{\text{u.c.p.}} Z$ if and only if $\sup_{s \leq t} |Z_s^n - Z_s| \xrightarrow{\mathbb{P}} 0$ for all t finite.

Theorem 1 indicates that RRDV is a consistent estimator under infill asymptotics when the efficient prices follow a continuous Itô semimartingale. The result is straightforward to prove with the law of large numbers (LLN) for path-dependent functionals of Itô semimartingales, as summarized in Duembgen and Podolskij (2015). To derive an associated central limit theorem (CLT), we need to impose some regularity conditions on σ:

Assumption 1. σ does not vanish and follows a continuous Itô semimartingale of the form

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\mu}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{v}_s dB_s, \tag{7}$$

where $\tilde{\mu} = (\tilde{\mu}_t)_{t\geq 0}$, $\tilde{\sigma} = (\tilde{\sigma}_t)_{t\geq 0}$, and $\tilde{v} = (\tilde{v}_t)_{t\geq 0}$ are adapted, càdlàg processes, and $B = (B_t)_{t\geq 0}$ is another Brownian motion independent of W.

Remark 1. This assumption rules out possible discontinuities in σ, which is at odds with some empirical evidence, see, e.g., Eraker et al. (2003), Jacod and Todorov (2010), Todorov and Tauchen (2011), and Bandi and Renò (2016). It can be harmlessly relaxed without altering the limit in the next theorem, but needs substantial extra calibration in the proofs. Some relevant discussions can be found in Christensen et al. (2009) and Christensen and Podolskij (2012).

Theorem 2 (Asymptotic normality)**.** Assume that the efficient price X follows a continuous Itô semimartingale in Eq. (5) with Assumption 1 satisfied. Then as $\Delta_n \to 0$, we have

$$\frac{1}{\sqrt{\Delta_n}}\left(\widehat{V}_{t,n} - \int_0^t \sigma_s^2 ds\right) \xrightarrow{\mathcal{L}-\mathrm{s}} \mathcal{MN}\left(0, \Theta \int_0^t \sigma_s^4 ds\right), \tag{8}$$

with the variance factor $\Theta = (\Lambda_4 - \Lambda_2^2)/\Lambda_2^2 \approx 0.7245$, and $\Lambda_4 = 24\ln 2 - 12 - 3\zeta(3) \approx 1.0294$, where $\zeta(3) = \sum_{n=1}^{\infty} n^{-3} \approx 1.2021$ is the Riemann's zeta function. We denote by $\xrightarrow{\mathcal{L}-\mathrm{s}}$ the stable convergence in law, and by $\mathcal{MN}$ a mixed normal distribution, i.e., a normal distribution conditional on the realization of its $\mathcal{F}$-conditional variance, which is a random variable.

Remark 2. Compared with the DV estimator of Andersen et al. (2021) which features a variance factor of 3, the asymptotic variance of our candlestick-based RRDV estimator is about four times smaller under infill asymptotics. This result might seem surprising initially, given that Kolokolov et al. (2023) demonstrate that DV attains the variance lower bound of drift-robust IV estimator based on returns from two adjacent blocks. In essence, the improvement of RRDV over DV originates from the additional information contained in high-frequency intraday ranges, which induces a different limiting statistical experiment. This additional information also leads to the diminished variance of the RRV estimator of Christensen and Podolskij (2007) over RV, which is the return-based minimum-variance unbiased estimator (MVUE). Finally, if the robustness to locally explosive trends is not pursued, then one can construct a variance optimal candlestick-based IV estimator in the spirit of Garman and Klass (1980), Li et al. (2022), and Bollerslev et al. (2023). As this is not the main focus of this paper, we shall leave it for further research.

Remark 3. Similar to the limiting distribution of RRV in Christensen and Podolskij (2007), the variance factor Θ in Theorem 2 is an infill-asymptotic result based on the presumption that the true HLOCs in all short episodes are observed. In practice, the efficiency of range-based estimators depends on the number of observations used to calculate the high-low range within each interval. See a detailed discussion in Section 4.1 about the RRDV estimator constructed from candlesticks formed by a finite number of observations.

For feasible implementation of the asymptotic distribution in Theorem 2, we can estimate the integrated quarticity (IQ) with the following range-return difference quarticity (RRDQ) estimator constructed analogously to RRDV:

$$\widehat{Q}_{t,n} = \frac{n}{\Lambda_4} \sum_{i=1}^{n} (w_i - |r_i|)^4 . \tag{9}$$

With techniques similar to Theorem 1, we can establish the consistency result for RRDQ:

Corollary 1 (Feasible inference)**.** Under the same conditions as in Theorem 1, it holds that

$$\widehat{Q}_{t,n} \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^4 ds. \tag{10}$$

The stable convergence in Theorem 2 implies that

$$\sqrt{\frac{n}{\Theta \widehat{Q}_{t,n}}} \left(\widehat{V}_{t,n} - \int_0^t \sigma_s^2 ds \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{11}$$

Remark 4. Similar to RRDV, the fourth-moment estimator $\widehat{Q}_{t,n}$ is robust to both discontinuities (in Section 2.3) and short-lived locally explosive trends (in Section 3). The proofs are analogous and thus omitted here.

2.3 Jumps

We examine the behavior of our RRDV estimator constructed on a discontinuous Itô semimartingale defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, e.g., with the Grigelionis (1980) representation:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \left(\delta \mathbb{1}_{\{|\delta|\leq 1\}}\right) \star \left(\underline{p} - \underline{q}\right)_t + \left(\delta \mathbb{1}_{\{|\delta|>1\}}\right) \star \underline{p}_t, \tag{12}$$

where $\underline{p} = \underline{p}(dt, dx)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with a compensator $\underline{q} = \underline{q}(dt, dx) = dt \otimes \lambda(dx)$, λ is a σ-finite measure on $\mathbb{R}$, and the function $\delta(\omega, t, x)$ on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ is predictable; see Aït-Sahalia and Jacod (2014) for details regarding the last two integrals.

Assumption 2. There exists a sequence $(\tau_n)_{n\geq 1}$ of stopping times increasing to ∞, and a sequence of deterministic nonnegative functions f_n on $\mathbb{R}$ for each n, which satisfies $|\delta(\omega, t, x)| \wedge 1 \leq f_n$ for all (ω, t, x) with $t \leq \tau_n(\omega)$, and $\int_{\mathbb{R}} |f_n|^r \lambda(dx) < \infty$ for some $r \in [0, 1)$.

Remark 5. The parameter r sets a bound on the degree of jump activity. With some $r \in [0, 1)$, we consider jumps of both finite and infinite activities, but restrict them to be of finite variation, i.e., they are absolutely summable, such that in Eq. (12) we can dispense with the integral with $\underline{p} - \underline{q}$, see Jacod et al. (2019) for more details.

Jumps of order of magnitude Δ_n^{ϖ} for some $\varpi \in [0, 1/2)$ prevail over the diffusion component and induce non-negligible shifts in $X(\omega)$ under infill asymptotics. We denote by $\Delta X_t = X_t - X_{t-}$

the size of discontinuous shift at time t. The robustness of RRDV in the presence of jumps is shown in the next proposition.

Proposition 1 (Jump robustness)**.** Assume that the efficient price X follows a discontinuous Itô semimartingale in Eq. (12) with Assumption 2 satisfied. Then as $\Delta_n \to 0$,

$$\widehat{V}_{t,n} - \int_0^t \sigma_s^2 ds = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right), \tag{13}$$

where the bias induced by jumps is asymptotically negligible and has no impact on the asymptotic distribution in Theorem 2.

Proposition 1 indicates the robustness of RRDV to both finite-activity and infinite-activity but finite-variation jumps in the efficient price. In each of the intervals with a nonzero ΔX_t, the discontinuous component in Eq. (12) has a higher asymptotic order than the continuous component, and thus dominates both the range and absolute return under infill asymptotics. Consequently, the jumps are mechanically cancelled in the range-return differences. As a result, the contribution from intervals that contain jumps in RRDV is asymptotically negligible and does not affect the consistency and asymptotic normality in Theorems 1 and 2, respectively.

In contrast, the differencing of returns removes the contribution from similar realizations of a locally persistent term in consecutive intervals, but it retains the unexpected increments from the distinctly less persistent jump component. Therefore, the DV estimator relies on an additional truncation method introduced by Mancini (2009) to discard all unexpectedly large differenced returns that may possibly contain jumps.

As pointed out by Andersen et al. (2021), market participants may imperfectly react to the shifts in economic fundamentals, and sometimes induce an short-lived deviation between the efficient and observed prices. This phenomenon, referred to as the "gradual jumps" identified by Barndorff-Nielsen et al. (2009) and Hoffmann et al. (2018), will be discussed in the next section as a typical manifestation of short-lived extreme return persistence.

3 Extreme Return Persistence

3.1 Drift Burst Model

As assumed in Section 2.2, the drift $\mu = (\mu_t)_{t\geq 0}$ is locally bounded, so that we can estimate IV consistently under infill asymptotics, because the drift becomes invisible since $\Delta_n \ll \sqrt{\Delta_n}$ in the limit, i.e., for a fixed time point τ, we have

$$\int_\tau^{\tau+\Delta_n} \mu_s ds = O_{\mathbb{P}}\left(\Delta_n\right) \quad \text{and} \quad \int_\tau^{\tau+\Delta_n} \sigma_s dW_s = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right), \quad \text{as } \Delta_n \to 0. \tag{14}$$

Christensen et al. (2022) point us in a new direction to understand some highly directional price movements over short episodes, in which the unbounded drift prevails over volatility and locally

dominates log-returns in the limit, which is summarized in the following assumption.

Assumption 3 (Drift burst model)**.** The efficient price X is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and assumed to be a continuous semimartingale described by

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + H_t, \tag{15}$$

with

$$H_t = \int_0^t \mu_s^b ds = \int_0^t \frac{c_s^- \mathbb{1}_{\{s<\tau\}} + c_s^+ \mathbb{1}_{\{s>\tau\}}}{|s-\tau|^\alpha} ds, \tag{16}$$

where $\tau \geq 0$, and $1/2 < \alpha < 1$. The coefficients c_t^- and c_t^+ are continuous and twice differentiable deterministic functions. All usual conditions for μ and σ are satisfied.

It is assumed that the bursting drift term μ_t^b has a singularity at the "burst time" τ, and thus explode in the vicinity of τ. The order of magnitude of H_t is given by

$$\int_{\tau-\Delta_n}^{\tau} \frac{c_s^-}{(\tau-s)^\alpha} ds \asymp \int_{\tau}^{\tau+\Delta_n} \frac{c_s^+}{(s-\tau)^\alpha} ds = O_{\mathbb{P}}\left(\Delta_n^{1-\alpha}\right). \tag{17}$$

We allow for different drift explosion coefficients c_t^- and c_t^+ before and after τ, and use the same rate of explosion α on both sides for ease of exposition. We restrict $\alpha < 1$ for the continuity of sample paths. When $\alpha > 1/2$, the volatility is completely swamped by the drift in the vicinity of τ, which induces a short-lived return persistence, and biases the nonparametric volatility estimators constructed from high-frequency intraday returns.[2]

When $1/2 < \alpha < 1$ and $c_\tau^- c_\tau^+ < 0$, the trajectory shows a "V-shape" (or "Λ-shape") in the neighborhood of τ, thanks to a discontinuity in the sign of H_t which locally dominates log-returns (Flora and Renò, 2022). Different drift explosion coefficients c_t^- and c_t^+ can be harmlessly employed to mimic patterns akin to V-shaped flash crashes ($c_\tau^- < 0, c_\tau^+ > 0$) or gradual jumps ($c_\tau^+ = 0$). It will not affect main intuitions.

Proposition 2. Assume that the efficient price X follows a continuous semimartingale in Eq. (15) and Assumption 3 holds with $1/2 < \alpha < 1$. For the RRDV estimator, it holds that

$$\widehat{V}_{t,n} - \int_0^t \sigma_s^2 ds = O_{\mathbb{P}}\left(\Delta_n^{\frac{1}{2\alpha}} \bigvee \sqrt{\Delta_n}\right) = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right), \quad \forall \alpha \in \left(\frac{1}{2}, 1\right), \tag{18}$$

where the bias induced by the drift burst is asymptotically negligible and has no impact on the asymptotic distribution in Theorem 2.

Under Assumption 3, the drift burst component $H = (H_t)_{t\geq 0}$ dominates the price movement in the vicinity of τ, i.e., the "explosion effect zone", while its impact diminishes as t moves away from

[2] We follow Andersen et al. (2021) to consider such episodic Itô semimartingale violations with only an exploding drift. It does not necessarily allow local arbitrage opportunities in the specification of Christensen et al. (2022), which accommodates simultaneous drift and volatility bursts with different rates of explosion.

τ. Both the range and absolute return in an interval within the effect zone near τ are dominated by a common excessive component, that is, the H increment of a higher asymptotic order than $\sqrt{\Delta_n}$. The pairwise offset between ranges and returns naturally nullifies the impact of H. As a result, the presence of drift burst can only induce an estimation bias of order $O_{\mathbb{P}}(\Delta_n^{1/2\alpha})$ for the RRDV estimator, which is asymptotically negligible and does not affect the CLT result in Theorem 2 for all possible $\alpha \in (1/2, 1)$. Over the region not in the vicinity of τ, the invisible H retains the continuous Itô semimartingale assumption with a bias of order $O_{\mathbb{P}}(\sqrt{\Delta_n})$.

The bias result of RRDV in Proposition 2 has a similar form as that of DV in Andersen et al. (2021). We find that the asymptotic order of the RRDV bias under Assumption 3 aligns with the bias order $O_{\mathbb{P}}(\sqrt{\Delta_n})$ under the continuous Itô semimartingale assumption, and its upper bound $\sqrt{\Delta_n}$ does not exceed that of the DV bias. Furthermore, the bias result of RRDV remains unaffected by the rate of drift explosion α and does not depend on any parameter choices, while the DV bias is independent of α only when some conditions for the truncation threshold are satisfied. The result can be extended to the case with stochastically distributed explosion times over $[0, t]$, as illustrated next in Section 3.2.

3.2 Persistent Noise Model

Andersen et al. (2021) introduce an alternative specification for the episodic emergence of extreme directional price movements. They consider these complex price patterns as outcomes of market uncertainty caused by imperfect information and irrational market participants.

Assumption 4 (Persistent noise model)**.** The observed price X is a combination of the efficient price, modeled as a possibly discontinuous Itô semimartingale in Eq. (12), and a component H that accommodates persistent price movements over irregularly spaced episodes:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{0 \le s \le t} \Delta X_s + H_t. \tag{19}$$

We denote by τ_i the first occurrence of the i-th "persistent noise" episode, so that $\tau_1, \tau_2, \ldots, \tau_N \in [0, t)$ form an increasing sequence of stopping times, with N finite almost surely. The persistent noise component is given by

$$H_t = \sum_{i:\tau_i \le t} H_t^{(i)} \mathbb{1}_{\{\epsilon_t^{(i)} \ge 0\}}, \tag{20}$$

with $H_t^{(i)}$ defined as

$$H_t^{(i)} = f^{(i)}(\Delta X_{\tau_i}, \eta_{\tau_i}) g^{(i)}(t), \tag{21}$$

where $\Delta X_{\tau_i} = X_{\tau_i} - X_{\tau_i^-}$ is the efficient price jump at τ_i, η_{τ_i} is an $\mathcal{F}_{\tau_i}$-measurable random variable, $f^{(i)}$ is a continuous and bounded function, and $g^{(i)}$ has one of the functional forms as follows:

$$g_{gj}^{(i)}(t) = \left[1 - \left(\frac{t - \tau_i}{\overline{\tau}_i - \tau_i}\right)^{\beta}\right] \mathbb{1}_{\{t \in [\tau_i, \overline{\tau}_i]\}}, \tag{22}$$

where $0 < \beta < 1/2$, and $\overline{\tau}_i > \tau_i$ is an $\mathcal{F}_{\tau_i}$-measurable random variable, or

$$g_{fc}^{(i)}(t) = c_t^- \left[1 - \left(\frac{\breve{\tau}_i - t}{\breve{\tau}_i - \tau_i}\right)^{\beta^-}\right] \mathbb{1}_{\{t\in[\tau_i,\breve{\tau}_i]\}} + c_t^+ \left[1 - \left(\frac{t - \breve{\tau}_i}{\overline{\tau}_i - \breve{\tau}_i}\right)^{\beta^+}\right] \mathbb{1}_{\{t\in[\breve{\tau}_i,\overline{\tau}_i]\}}, \tag{23}$$

where $0 < \beta^-, \beta^+ < 1/2$, the coefficients c_t^- and c_t^+ are continuous and twice differentiable deterministic function, and $\tau_i < \breve{\tau}_i < \overline{\tau}_i$ are all $\mathcal{F}_{\tau_i}$-measurable random variables. Moreover,

$$\epsilon_t^{(i)} = \sum_{s\in[\tau_i,t]} \Delta\epsilon_s \tag{24}$$

is a finite-activity pure jump process with negative jumps.

Remark 6. Each of the episodes is activated and terminated by the realizations of τ_i and $\overline{\tau}_i$, or randomly ended by $\epsilon_t^{(i)}$ in the middle. The function $f^{(i)}$ captures the initial market reaction to events with ambiguous information that trigger persistent noise episodes, the random variable η_{τ_i} allows for a random response to such events, and the function $g^{(i)}$ describes the price pattern over a temporary disequilibrium after ambiguous information arrives. Assumption 4 allows for two basic forms of $H^{(i)}$ to model market uncertainty in two different scenarios:

I. Market participants underreact (or slowly react) to a shift in fundamentals. In this scenario, there exists $\Delta X_{\tau_i} \neq 0$, the function $g^{(i)}$ takes the form $g_{gj}^{(i)}$ in Eq. (22), and $f^{(i)}(\Delta X_{\tau_i}, \eta_{\tau_i}) = -\eta_{\tau_i}\Delta X_{\tau_i}$ with $\eta_{\tau_i} = 1$ or $\eta_{\tau_i} \in (0,1)$, which partially offsets the efficient price jump at τ_i.

II. Market participants worry about a potential shift in fundamentals. In this scenario, $\Delta X_{\tau_i} = 0$, and the function $g^{(i)}$ takes the form $g_{fc}^{(i)}$ in Eq. (23). The deviation from efficient price is (fully or partially) recovered shortly after due to reverse trades by arbitrageurs, which leads to a V-shaped trajectory with a turning point at a random time $\breve{\tau}_i$.

The scenario I and II correspond to two phenomena observed in financial markets, i.e., gradual jumps and flash crashes, respectively. $H^{(i)}$ in scenario II can be viewed as a stochastic extension of the drift burst model in Assumption 3, with stochastically distributed explosion times over $[0,t]$.

We next state an analogous result to Proposition 2 when the observed prices persistently deviate from the fundamental values due to short-lived market inefficiency.

Proposition 3. Assume that the market price X follows a contaminated Itô semimartingale in Eq. (19) with finite-activity jumps, i.e., $r = 0$, and there exists a persistent noise episode $[\tau, \overline{\tau}] \subset [0,t]$. The function $g^{(1)}$ in the noise component $H_t^{(1)}$ takes either of the two forms in Eqs. (22) and (23). For the RRDV estimator, it holds that

$$\widehat{V}_{t,n} - \int_0^t \sigma_s^2 ds = \begin{cases} O_{\mathbb{P}}\left(\Delta_n^{\frac{1}{2(1-\beta)}} \bigvee \sqrt{\Delta_n}\right) = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right), & \forall \beta \in (0,1/2), \quad \text{when } g^{(1)} = g_{gj}^{(1)}, \\ O_{\mathbb{P}}\left(\Delta_n^{\frac{1}{2(1-\beta^-\wedge\beta^+)}} \bigvee \sqrt{\Delta_n}\right) = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right), & \forall \beta^\pm \in (0,1/2), \quad \text{when } g^{(1)} = g_{fc}^{(1)}, \end{cases} \tag{25}$$

where the bias induced by the persistent noise is asymptotically negligible and has no impact on the asymptotic distribution in Theorem 2.

In Proposition 3, our discussion is confined to a simplified scenario featuring a single persistent noise episode within the interval $[0, t]$. Similar to the result in Proposition 2 with a drift burst, the estimation bias is of order $O_{\mathbb{P}}(\Delta_n^{\frac{1}{2(1-\beta)}})$ with a gradual jump and $O_{\mathbb{P}}(\Delta_n^{\frac{1}{2(1-\beta^-\wedge\beta^+)}})$ with a V-shaped flash crash, respectively, over the "explosion effect zone" where the role of H is no smaller than the diffusion component. The results obtained in the simplified scenario can be straightforwardly extended to more general cases involving a finite number of such episodes with non-overlapping effect zones. Simulation results in Section 5 shows that the RRDV estimator remains unbiased in the presence of a gradual jump with an intermittent small flash crash.

Over each short interval in the vicinity of $\check{\tau}$, the persistent noise H adds the same amount to both range and absolute return. The only exception is the so-called "V" interval in a flash crash, i.e., the interval which accommodates the reversal point $\check{\tau}$. The range-return difference in the "V" interval is an $O_{\mathbb{P}}(\Delta_n^{\beta^-\wedge\beta^+})$ variable, while its impact is negligible when n approaches infinity. However, a steep "V-shape" could deteriorate the finite-sample performance of RRDV when the interval length is far away from infinitesimal in practice, which shall be discussed in Section 4.2.

4 Finite-Sample Refinements

4.1 Finite-Sample Bias I: Discretization Errors

Variance estimators constructed from high-low ranges often exhibit a systematic downward bias in practice. This issue was initially identified by Garman and Klass (1980), Beckers (1983), and Rogers and Satchell (1991) in the context of daily variance estimation. Range-based IV estimation with intraday observations faces a similar challenge. The source of this downward bias is the difference between the discretized range calculated from the discrete observations available in practice and the true range originating from a continuous-time process, which is referred to as discretization error. More specifically, since the discretized minimum (maximum) is obtained from a smaller set, it will be greater (smaller) than the continuous minimum (maximum). Consequently, the discretized range includes a negative discretization error. In other words, the scaling factor Λ_2 derived from a standard Brownian motion leads to an over-scaling of the sum of squared discretized range-return differences. Therefore, it is advisable to replace it with a discretized scaling factor based on discretely observed Brownian motion.

To formalize this idea, we introduce some additional notation: We denote by N the number of observations in each interval $((i-1)\Delta_n, i\Delta_n]$, and assume there are totally $Nn+1$ equidistant price observations available over $[0, t]$. We denote the discretized high-low range over the i-th interval by

$$w_{i,N} = H_{i,N} - L_{i,N} = \sup_{j\in\{0,1,\ldots,N\}} X_{(i-1)\Delta_n+j\Delta_n/N} - \inf_{j\in\{0,1,\ldots,N\}} X_{(i-1)\Delta_n+j\Delta_n/N}, \tag{26}$$

and then define the discretized RRDV estimator as

$$\widehat{V}_{t,n,N} = \frac{1}{\Lambda_{2,N}} \sum_{i=1}^{n} \left(w_{i,N} - |r_i|\right)^2, \tag{27}$$

where $\Lambda_{p,N}$ is the counterpart to Λ_p in Eq. (4) when the standard Brownian motion is discretely observed at $t = i/N$ for $i = 0, 1, \ldots, N$ over a unit interval:

$$\Lambda_{p,N} = \mathbb{E}\left[\left(\sup_{i,j\in\{0,1,\ldots,N\}} W_{i/N} - W_{j/N} - |W_1|\right)^p\right]. \tag{28}$$

To investigate how the discretization error $\Lambda_{p,N} - \Lambda_p$ evolves across different N, we present the asymptotic expansions for $\Lambda_{2,N}$ and $\Lambda_{4,N}$ as follows:

Proposition 4. Assume $W_N = (W_{t,N})_{t\in[0,1]} = (W_{i/N})_{i\in\{0,1,\ldots,N\}}$ is an embedded random walk equidistantly spaced at $N+1$ points over $[0,1]$. The following results hold for $\Lambda_{p,N}$ in Eq. (28):

$$\Lambda_{2,N} = \Lambda_2 + \frac{4}{\pi}\zeta\left(\frac{1}{2}\right)\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right), \tag{29}$$

$$\Lambda_{4,N} = \Lambda_4 + \left(\frac{48}{\pi} - 4\pi\right)\zeta\left(\frac{1}{2}\right)\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right), \tag{30}$$

as $N \to \infty$, where $\zeta(1/2) \approx -1.4604$.

Remark 7. The asymptotic expansions in Proposition 4 are based on the results in Asmussen et al. (1995), who derive the asymptotic results for Euler discretization errors of one-dimensional reflected Brownian motions, with details summarized in Lemma B.4 in Appendix B.7.

Proposition 4 indicates that both $\Lambda_{2,N}$ and $\Lambda_{4,N}$, as well as the variance factor $\Theta_N = (\Lambda_{4,N} - \Lambda_{2,N}^2)/\Lambda_{2,N}^2$, can be approximated when N is sufficiently large. This fact inspires us to provide practitioners with polynomial approximations for all factors when a finite N is applied in practice. Fig. 2 compares the approximated and simulated values of $\Lambda_{2,N}$ and Θ_N with $11 \le N \le 2000$, which confirms the precision of polynomial approximations. Further details on the approximations and practical guidance can be found in Appendix C.2.

The consistency and asymptotic normality of the discretized RRDV estimator in Eq. (27) are summarized in the next corollary. Similar to the discretized RRV in Christensen and Podolskij (2007), the CLT result holds for arbitrary N converging to some integer larger than 1.

Corollary 2. Under the same conditions as in Theorem 1, it holds that as $\Delta_n \to 0$,

$$\widehat{V}_{t,n,N} \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^2 ds. \tag{31}$$

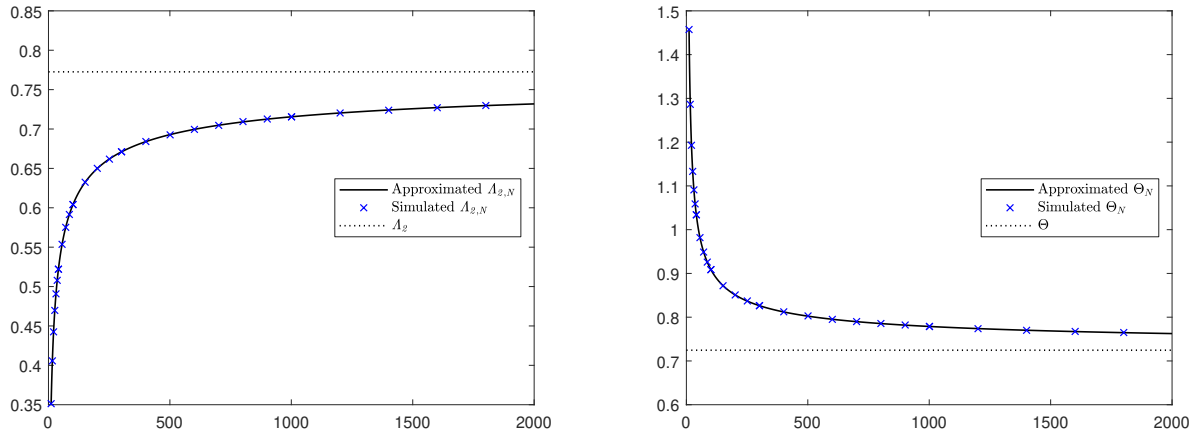

Figure 2: Comparison of approximated and simulated values of $\Lambda_{2,N}$ and Θ_N with $11 \leq N \leq 2000$.

Moreover, if Assumption 1 holds and $N \to c \in \mathbb{N}^{>1}$, it holds that

$$\frac{1}{\sqrt{\Delta_n}}\left(\widehat{V}_{t,n,N} - \int_0^t \sigma_s^2 ds\right) \xrightarrow{\mathcal{L}-\mathrm{s}} \mathcal{MN}\left(0, \Theta_c \int_0^t \sigma_s^4 ds\right), \tag{32}$$

where $\Theta_c = (\Lambda_{4,c} - \Lambda_{2,c}^2)/\Lambda_{2,c}^2$. Finally, the stable convergence implies that

$$\sqrt{\frac{n}{\Theta_N \widehat{Q}_{t,n,N}}}\left(\widehat{V}_{t,n,N} - \int_0^t \sigma_s^2 ds\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \tag{33}$$

where the discretized RRDQ estimator $\widehat{Q}_{t,n,N}$ is given by

$$\widehat{Q}_{t,n,N} = \frac{n}{\Lambda_{4,N}} \sum_{i=1}^{n} (w_{i,N} - |r_i|)^4 \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^4 ds. \tag{34}$$

The discretized RRDV is undefined when $N \to 1$ because $\Lambda_{2,1} = 0$, unlike RRV that reduces to the standard RV when there are only open and close prices available for all intervals. Furthermore, compared with the DV estimator, the asymptotic variance of the discretized RRDV becomes smaller even when only five observations (including open and close) are available in each interval. Simulation results in Section 5.3 demonstrate that the discretized RRDV based on half-a-minute observations (with 1, 2, 3, 5-minute candlestick intervals) can still produce reliable IV estimates, with only a mild increase in finite-sample variance.

The effective correction for discretization errors ensures the reliability of RRDV constructed from discretized candlestick information, i.e., the HLOCs obtained from sparsely or "not-too-finely" sampled price observations within each candlestick interval. The fact that the market microstructure noise becomes more pronounced with higher sampling frequencies has inspired the widespread use of volatility estimators based on sparsely sampled data (Aït-Sahalia et al., 2005). The utilization of discretized HLOCs provides our estimator with a similar robustness to market microstructure

noise, without introducing additional complexity for implementation.[3] However, a comprehensive investigation of the asymptotic and finite-sample behavior of range-based estimators constructed from ultra high-frequency or "finely" sampled data with noise contamination requires a more explicit assumption for the noise structure, where the literature is still far from reaching a consensus (Bollerslev et al., 2023). The extension of RRDV in this direction is left for future research.

4.2 Finite-Sample Bias II: V-Shapes

Both models in Section 3 can be employed to mimic flash crashes. As commented after Proposition 3, the reversal point $\breve{\tau}_i$ (or the explosion time τ in the drift burst model) has no impact on RRDV in the limit, but could deteriorate its finite-sample performance when the interval size is not sufficiently small. An example in Fig. 3 illustrates the candlestick patterns around a V-shaped flash crash. In this example, the candlestick in the 5-minute "V" interval has a long lower wick, i.e., the so-called hammer pattern, which potentially introduces a positive bias in the RRDV estimate. This bias becomes particularly pronounced in cases where the V-shape is steep. To mitigate the V-shape bias in finite samples, we augment the RRDV estimator with a truncation threshold for the range-return differences. This augmentation has no impact on the asymptotic results in Sections 2 and 3, but improves the finite-sample robustness of RRDV to different interval lengths.

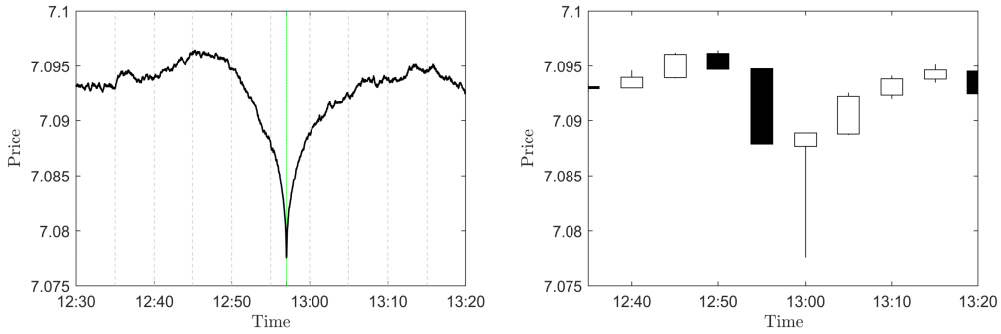

Figure 3: 5-minute candlesticks around a simulated V-shaped flash crash.

We employ the classical truncation threshold $\varphi = \zeta \Delta_n^{\varpi}$ with $\varpi \in (0, 1/2)$, initially introduced by Mancini (2009), and set the threshold parameters by using a data-adaptive method. Andersen et al. (2021) use the truncation threshold for both absolute returns in the truncated realized volatility (TRV) and absolute differenced returns in DV, with

$$\zeta = C_\zeta \sqrt{\mathrm{MedRV}_{t,n}}, \tag{35}$$

where $\mathrm{MedRV}_{t,n}$ is the median RV estimator of Andersen et al. (2012), i.e.,

$$\mathrm{MedRV}_{t,n} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{n}{n-2} \right) \sum_{i=2}^{n-1} \mathrm{median}\left(|r_{i-1}|, |r_i|, |r_{i+1}| \right)^2, \tag{36}$$

[3] Some simulation results are reported in Appendix C.5.

and $C_\zeta^{\mathrm{DV}} = \sqrt{2}C_\zeta^{\mathrm{TRV}} = 3\sqrt{2}$, motivated by the ratio $\sqrt{2}$ between standard deviations of (absolute) differenced and undifferenced i.i.d. Brownian returns. For the RRDV estimator, we truncate the range-return differences with the threshold of the same form, i.e.,

$$\widehat{V}_{t,n,N} = \frac{1}{\Lambda_{2,N}} \sum_{i=1}^{n} (w_{i,N} - |r_i|)^2 \, \mathbb{1}_{\{w_{i,N} - |r_i| \leq \zeta^{\mathrm{RRDV}} \Delta_n^{\varpi}\}}, \tag{37}$$

where the parameter ζ^{RRDV} is given by Eq. (35) with $C_\zeta^{\mathrm{RRDV}} = 2$, which is approximately the same quantile (99.7%) of range-return differences from a standard Brownian motion.

5 Monte Carlo Simulations

This section contains a Monte Carlo study to examine both the asymptotic unbiasedness and the finite-sample performance of the RRDV estimator, which corresponds to the results developed in Sections 2 and 3.

5.1 Simulation Design

We simulate a Heston model for the efficient price process X (Heston, 1993):

$$\begin{aligned} dX_t &= \mu dt + \sigma_t dW_{1,t} + dJ_t, \quad t \in [0,1], \\ d\sigma_t^2 &= \kappa \left(\theta - \sigma_t^2\right) dt + \eta \sigma_t dW_{2,t}, \end{aligned} \tag{38}$$

where W_1 and W_2 are standard Brownian motions with $\mathrm{Corr}(W_{1,t}, W_{2,t}) = \rho$, and J is a compound Poisson process, i.e.,

$$J_t = \sum_{i=1}^{N_t} Z_i, \tag{39}$$

where N is a Poisson process with rate λ, and Z_i follows a normal distribution $\mathcal{N}(0, \varsigma^2)$. We start with the initial price $X_0 = \ln 1200$, and take the Heston parameters as follows:

$$\begin{aligned} &\mu = 0.05/252, \quad \kappa = 5/252, \quad \theta = 0.0225/252, \quad \eta = 0.4/252, \\ &\rho = -\sqrt{0.5}, \quad \lambda = 1/5, \quad \varsigma = 0.9\%. \end{aligned} \tag{40}$$

The volatility parameters satisfy the Feller's condition $2\kappa\theta \geq \eta^2$ which ensures the positivity of σ. The process J simulated with $\lambda = 1/5$ corresponds to one jump per week, and generates around 6.5% of the daily quadratic variation on average.

In this section, we firstly examine the unbiasedness of RRDV in "continuous time": We simulate half-millisecond (0.0005-second) price observations for 2000 days, and construct RRDV on candlestick information in 1, 5, 10, 30-second and 1, 2, 3, 5-minute intervals, respectively. Then we evaluate its finite-sample performance: We simulate one-second and 30-second observations for 10000 days, and construct the candlesticks on 1, 2, 3, and 5-minute intervals, respectively. All simulated observations

are equidistantly distributed in $[0, 1]$ which consists of 6.5 hours of trading.

We follow the persistent noise model of Andersen et al. (2021) to simulate three different patterns of episodic extreme return persistence:[4]

- ***Gradual Jump***: We insert a shift in fundamentals $\Delta X_\tau = 2.5\%$ at $\tau = 0.5$ for all days. For the persistent noise component in Eq. (21), we let $i \in \{1\}$, $\tau_1 = \tau$, $f^{(1)}(\Delta X_\tau, \eta_\tau) = -\eta_\tau \Delta X_\tau$ with $\eta_\tau = 1$, $g^{(1)}$ take the form $g^{(1)}_{gj}$ in Eq. (22), and $(\tau, \overline{\tau}) = (0.5, 0.59)$.
- ***Flash Crash***: We let $i \in \{1\}$, $\tau_1 = \tau$, $f^{(1)}(\Delta X_\tau, \eta_\tau) = -\eta_\tau$ with $\eta_\tau = 2\%$, $g^{(1)}$ take the form $g^{(1)}_{fc}$ in Eq. (23), $(\tau, \breve{\tau}, \overline{\tau}) = (0.41, 0.49, 0.57)$, and $c^{\pm} = 1$.[5]
- ***Gradual Jump + Flash Crash***: We consider two overlapped persistent noise episodes: $i \in \{1, 2\}$. We insert a shift in fundamentals $\Delta X_{\tau_1} = 2.5\%$ at $\tau_1 = 0.5$. We let $f^{(1)}(\Delta X_{\tau_1}, \eta_{\tau_1}) = -\eta_{\tau_1} \Delta X_\tau$ with $\eta_{\tau_1} = 1$, $g^{(1)}$ take the form $g^{(1)}_{gj}$ in Eq. (22), and $(\tau_1, \overline{\tau}_1) = (0.5, 0.65)$. For the intermittent (small) flash crash, we assume $f^{(2)}(\Delta X_{\tau_2}, \eta_{\tau_2}) = -\eta_{\tau_2}$ with $\eta_{\tau_2} = 0.75\%$, $g^{(2)}$ takes the form $g^{(2)}_{fc}$ in Eq. (23), and $c^{\pm} = 1$. We assume that the waiting time between τ_1 and τ_2 follows an exponential distribution with rate parameter $\lambda^{\text{Exp}} = 15$, and $(\breve{\tau}_2, \overline{\tau}_2) = (\tau_2 + 0.04, \tau_2 + 0.08)$.

For each scenario, we consider three different choices of parameter $\beta = \beta^{\pm} \in \{0.45, 0.35, 0.25\}$, which controls the steepness of short-lived directional price movement. For example, a smaller β in g_{gj} leads to a steeper gradual jump in observed prices, which is closer to the discontinuous shift in efficient prices, and corresponds to a less sticky expectation of market participants. Fig. 4 shows examples of simulated price paths e^X, efficient (in blue) and observed (in black), for all three scenarios with $\beta = \beta^{\pm} = 0.45$.

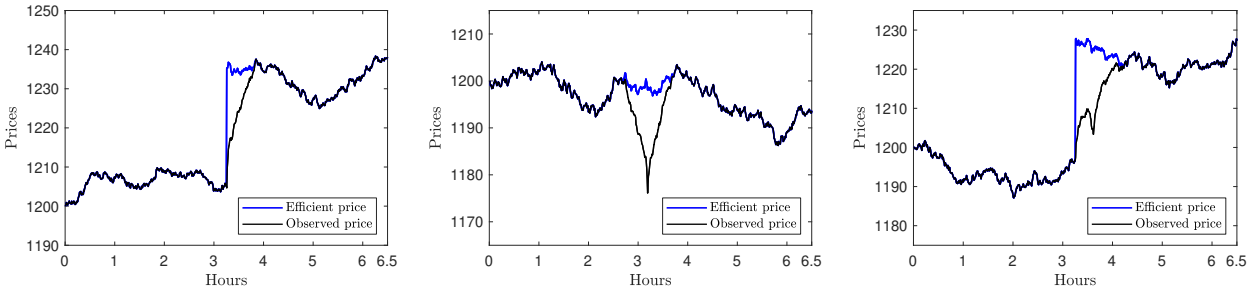

Figure 4: Simulated price paths e^X, efficient (in blue) and observed (in black), with (i) a gradual jump (left), (ii) a flash crash (middle), and (iii) a gradual jump with an intermittent small flash crash (right), respectively. The efficient price X is simulated with the Heston model in Eq. (38). The deviation between efficient and observed prices is simulated with the persistent noise model in Assumption 4, with all parameters listed above ($\beta = \beta^{\pm} = 0.45$).

[4] As shown in Eqs. (56) and (57) in Andersen et al. (2021), there exists an asymptotic correspondence between the two models of episodic extreme return persistence in Section 3, and they are equivalent with identical asymptotic analyses when $\beta = 1 - \alpha$. The simulation results with the drift burst model in Eq. (15) indicate the same qualitative conclusions.

[5] For flash crashes simulated with $g^{(i)} = g^{(i)}_{fc}$, we stale the observation on $\breve{\tau}_i$ to avoid an unnecessary "jump" on $\breve{\tau}_i$. For example, when $(\tau, \breve{\tau}, \overline{\tau}) = (0.41, 0.49, 0.57) = (9594, 11466, 13338)$ seconds, we truncate the tick-by-tick H increments between 11465 and 11467 seconds.

5.2 Asymptotic Unbiasedness

Table 1 reports the relative biases (%) of RRDV in "continuous time" with the simulated half-millisecond observations for 2000 days. In Panel A, we find that the biases have fairly small size when there exists no episodic extreme return persistence. The bias results with $H = 0$ confirm the consistency of our estimator (Section 2.2) and its robustness to discontinuities (Section 2.3). The existence of gradual jumps leads to only small biases of RRDV constructed from candlesticks with all selected interval lengths, and the biases shrink sufficiently when the number of intervals (resp. the length of intervals) becomes larger (resp. smaller). The V-shaped flash crashes also lead to only negligible biases with small intervals. These bias results in "continuous time" show compelling evidence for the asymptotic unbiasedness of RRDV in the presence of short-lived explosive trends, as shown in Proposition 2 and Proposition 3.[6] Fig. 5 collects the histograms and QQ plots of estimation errors in some of the Monte Carlo trials, which indicates the close-to-normality of the estimation errors of RRDV in all scenarios.

Table 1: Monte Carlo bias results (%)

Panel A: No "V" Bias Correction

		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-0.01	-0.15	-0.15	-0.16	-0.19	-0.23	-0.26	-0.15	-0.16	-0.20
5	0.01	-0.47	-0.40	-0.38	-0.72	-0.73	-0.62	-0.61	-0.53	-0.37
10	-0.02	-0.75	-0.77	-0.63	-0.61	0.25	1.72	-0.81	-0.69	-0.49
30	-0.04	-1.89	-1.52	-1.24	-1.38	0.27	2.90	-1.36	-1.06	-0.42
60	0.11	-2.75	-2.25	-1.69	-2.76	-0.91	2.18	-2.27	-0.11	1.23
120	0.12	-3.73	-3.06	-2.44	10.74	22.69	36.21	-0.73	1.11	4.11
180	0.14	-4.58	-3.94	-3.17	10.90	23.50	37.34	-0.70	1.54	5.10
300	0.13	-5.62	-4.88	-3.45	14.38	28.38	43.38	2.45	5.22	5.01

Panel B: With "V" Bias Correction

		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-0.80	-0.92	-0.87	-0.80	-0.98	-0.93	-0.90	-0.91	-0.90	-0.87
5	-0.92	-1.18	-1.05	-0.94	-1.34	-1.19	-1.03	-1.22	-1.24	-1.07
10	-0.93	-1.44	-1.27	-1.04	-1.70	-1.60	1.44	-1.48	-1.55	-2.18
30	-0.98	-2.12	-1.86	-1.51	-2.86	-2.54	2.10	-2.03	-2.35	-2.22
60	-0.99	-2.95	-2.52	-2.17	-3.19	-3.34	-3.00	-2.52	-2.65	-2.86
120	-1.26	-3.93	-3.10	-2.59	-6.32	-5.49	-4.34	-3.55	-4.27	-3.92
180	-1.35	-4.53	-3.89	-2.95	-7.60	-6.33	-5.19	-2.99	-4.48	-4.87
300	-1.66	-5.61	-5.07	-4.14	-7.41	-7.84	-6.65	-1.96	-2.81	-3.86

Relative bias (%) of RRDV constructed from 1, 5, 10, 30, 60, 120, 180, and 300-second candlesticks for 2000 days. Panel B reports the relative biases of RRDV with a truncation threshold applied for the range-return differences in all intervals, i.e., $2\sqrt{\Delta_n \mathrm{MedRV}_{t,n}}$, see details in Section 4.2. The discretization errors are corrected following the steps in Section 4.1. The DGP is the Heston model in Eq. (38), and we follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence.

[6] The bias results of some other IV estimators are presented in Appendix C.3.

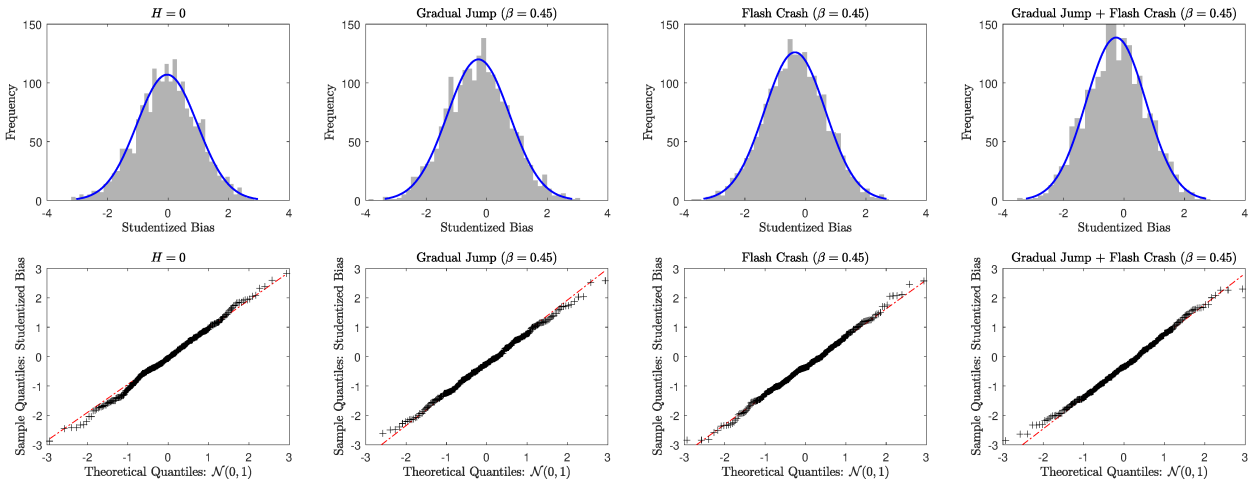

Figure 5: Histograms and QQ plots of studentized estimation errors of one-second RRDV. The estimation errors of RRDV are studentized by the estimates of asymptotic variance in Theorem 2, which involves the RRQV estimator in Eq. (9). The discretization errors are corrected following the steps in Section 4.1. The sample size for QQ plots is 300.

We notice that the reversal point of the V-shaped flash crash contributes to a positive bias when the intervals are relatively large, and the bias has a larger size when the V-shape is steeper, i.e., with a smaller β. For instance, the relative bias of 5-minute RRDV is 43.38% in the presence of a V-shaped flash crash with $\beta = 0.25$. Panel B of Table 1 reports the relative biases of RRDV with a truncation threshold applied for the range-return differences in all intervals, i.e., $2\sqrt{\Delta_n \mathrm{MedRV}_{t,n}}$, as elaborated in Section 4.2.[7] Estimating IV with the range-return differences truncated by some right-tail extreme quantile slightly worsen the bias results in "continuous time", but it avoids the significant V-shape bias in finite samples.

5.3 Finite-Sample Performance

To evaluate the finite-sample performance of RRDV, we limit the number of observations available in each interval to construct intraday candlesticks: We simulate both second-by-second and half-minute observations, and collect HLOCs in each of the 1, 2, 3, and 5-minute intervals. The RRDV estimators based on HLOCs obtained from 1-second and 30-second data are labeled as "RRDV" and "RRDV*", respectively, in the table of results, where the later corresponds to sparsely or "not-too-finely" sampled data. For the comparative analysis, we consider the truncated realized volatility (TRV) estimator of Mancini (2009):

$$\mathrm{TRV}_{t,n} = \sum_{i=1}^{n} r_i^2 \mathbb{1}_{\{|r_i| \leq \zeta^{\mathrm{TRV}} \Delta_n^{\varpi}\}}, \tag{41}$$

[7] Theoretically, the truncation threshold should have a higher order than $\sqrt{\Delta_n}$, i.e., $\varpi < 1/2$, while setting $\varpi = 1/2$ for a fixed interval length makes no difference in practice.

and the general family of DV estimators in Andersen et al. (2021):

$$\mathrm{DV}_{1-m,t,n} = \frac{\mathrm{DV}_{t,n}^{(1)} + \mathrm{DV}_{t,n}^{(2)} + \cdots + \mathrm{DV}_{t,n}^{(m)}}{m}, \tag{42}$$

where

$$\mathrm{DV}_{t,n}^{(m)} = \frac{1}{2} \sum_{i=m+1}^{n} (r_i - r_{i-m})^2 \, \mathbb{1}_{\{|r_i - r_{i-m}| \leq \zeta^{\mathrm{DV}} \Delta_n^{\varpi}\}}. \tag{43}$$

The choice of truncation parameters for TRV, DV and RRDV follows the instructions in Section 4.2, with $(C_\zeta^{\mathrm{TRV}}, C_\zeta^{\mathrm{DV}}, C_\zeta^{\mathrm{RRDV}}, \varpi) = (3, 3\sqrt{2}, 2, 1/2)$.[8] Their finite-sample performances are assessed and compared via the root-mean-square error (RMSE), i.e.,

$$\mathrm{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(\widehat{V}_{1,n} - \int_0^1 \sigma_t^2 dt \right)^2}, \quad \text{with } M = 10000. \tag{44}$$

In Panel A of Table 2, we present the RMSEs of all selected IV estimators in the absence of short-lived extreme return persistence. The small RMSEs ($< 2 \times 10^5$) indicate the robustness of all estimators to discontinuities. Compared with other estimators, RRDV has the smallest RMSE results with all selected sampling frequencies, which is consistent with the smallest variance in asymptotic theory. RRDV* based on HLOCs obtained from half-minute data has the largest RMSE with one-minute intervals, while it starts to achieve smaller RMSEs than DV when the length of intervals is extended to two minutes, i.e., there are at lease five observations (open and close included) available in each interval, which is in line with our numerical results in Section 4.1. Panel B, C and D in Table 2 report the RMSE results in the presence of gradual jumps or/and flash crashes. When there exist local explosive trends, the TRV estimator has larger RMSEs than RRDV and two DV estimators in all scenarios, and the difference becomes more pronounced for "stickier" (less steep) deviations between observed prices and efficient prices, i.e., with larger β's, and for lower sampling frequencies. For the candlestick-based estimators, RRDV* can achieve smaller RMSEs than DV in all cases when the interval length is three minutes or longer, and RRDV based on HLOCs taken from second-by-second observations outperforms all other IV measures across all relevant scenarios.

6 Empirical Analysis

In this section, we use the RRDV estimator as the basis for volatility forecasting under the popular heterogeneous autoregressive (HAR) framework for the SPDR S&P 500 ETF Trust (SPY), which is the best-recognized and oldest U.S. listed ETF and by far the most widely traded S&P 500 ETF.

[8] We also consider alternative parameter choices $(C_\zeta^{\mathrm{TRV}}, C_\zeta^{\mathrm{DV}}) = (4, 4\sqrt{2})$ used for comparison in Andersen et al. (2021). We find that the less aggressive threshold choices will not change the qualitative results and even worsen the finite-sample performance of both estimators when there exists excessive return drift, see Table C.6 in Appendix C.4, which is consistent with the Monte Carlo results in Andersen et al. (2021).

Table 2: Monte Carlo RMSE results

Panel A: $H = 0$

Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	**0.44**	1.09	0.74	0.84	0.76
2 min	**0.58**	1.02	1.00	1.14	1.03
3 min	**0.69**	1.08	1.22	1.39	1.24
5 min	**0.86**	1.23	1.55	1.77	1.59

Panel B: Gradual Jump

	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	**0.52**	1.14	1.11	0.80	0.70	**0.50**	1.12	0.96	0.79	0.70	**0.48**	1.11	0.82	0.79	0.70
2 min	**0.69**	1.07	1.95	1.15	1.15	**0.67**	1.06	1.61	1.18	1.07	**0.64**	1.03	1.31	1.11	1.01
3 min	**0.84**	1.18	2.47	1.47	1.34	**0.80**	1.14	2.00	1.38	1.23	**0.78**	1.13	1.64	1.36	1.21
5 min	**1.04**	1.36	3.59	2.14	2.21	**1.02**	1.33	2.86	1.89	1.74	**0.98**	1.30	2.21	1.78	1.61

Panel C: Flash Crash

	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	**0.52**	1.19	1.74	0.78	0.70	**0.53**	1.16	1.54	0.83	0.75	**0.53**	1.14	1.24	0.82	0.72
2 min	**0.85**	1.21	2.93	1.27	1.12	**0.78**	1.15	2.34	1.27	1.11	**0.72**	1.09	1.77	1.23	1.07
3 min	**1.00**	1.30	4.34	1.55	1.55	**0.93**	1.27	3.47	1.49	1.50	**0.86**	1.19	2.61	1.44	1.42
5 min	**1.18**	1.69	7.00	2.11	2.86	**1.16**	1.59	5.60	2.02	2.68	**1.10**	1.42	4.09	1.98	2.43

Panel D: Gradual Jump with an Intermittent Flash Crash

	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	**0.51**	1.14	1.28	0.81	0.74	**0.53**	1.15	1.23	0.79	0.71	**0.50**	1.11	1.03	0.79	0.71
2 min	**0.72**	1.12	2.40	1.26	1.55	**0.66**	1.05	2.28	1.27	1.36	**0.68**	1.06	1.59	1.15	1.12
3 min	**0.87**	1.24	3.14	1.72	2.07	**0.78**	1.11	2.89	1.53	1.60	**0.72**	1.14	2.32	1.44	1.51
5 min	**0.95**	1.32	3.71	2.41	2.76	**1.05**	1.40	3.49	2.09	2.30	**0.99**	1.31	2.45	1.81	1.80

RMSE (multiplied by 10^5) of different IV estimators. RRDV and RRDV* are constructed from intraday candlestick information, which is obtained from second-by-second and half-minute observations, respectively. The choice of truncation parameters for TRV, DV and RRDV follows the instructions in Section 4.2, with $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}, C_\zeta^{\text{RRDV}}, \varpi) = (3, 3\sqrt{2}, 2, 1/2)$. The discretization errors of RRDV are corrected following the steps in Section 4.1. The DGP is the Heston model in Eq. (38). We follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence.

6.1 Data

We obtain all high-frequency transaction records of SPY from the daily Trade and Quote (TAQ) dataset, with the sample period ranging from January 2, 2014 to December 31, 2021. The tick-by-tick transactions are timestamped in milliseconds until mid-2015 and in microseconds since then.[9] As is standard in empirical research with TAQ data, we use the filters as in Barndorff-Nielsen et al. (2009) to eliminate clear data errors, remove all transactions in the original record that are later corrected, cancelled or otherwise invalidated. In addition, we remove all trading days with an early market closure, and restrict our sample to transactions between 9:41:00 – 16:00:00 Eastern Time (ET) for all individual stocks. The final sample comprises of 1998 days.

6.2 Heterogeneous Autoregressive (HAR) Model

The HAR model of Corsi (2009) is designed to parsimoniously capture the dependence structures of return volatility in different horizons, and therefore aims to approximate its long memory that has been extensively confirmed by empirical literature. Renowned for its consistent and remarkable predictive performance, the HAR model serves as the predominant benchmark in volatility forecasting research. In this section, we denote some selected IV measure at day t by $\widehat{V}_t$, and introduce the following moving averages of daily volatility measures as:

$$\widehat{V}_{w,t} = \frac{1}{5}\sum_{i=1}^{5}\widehat{V}_{t-i+1} \qquad \text{and} \qquad \widehat{V}_{m,t} = \frac{1}{22}\sum_{i=1}^{22}\widehat{V}_{t-i+1}, \tag{45}$$

where $\widehat{V}_{w,t}$ represents the one-week average and $\widehat{V}_{m,t}$ denotes the one-month average of daily IV estimates, respectively. The standard one-day-ahead HAR model exhibits the following structure:

$$\widehat{V}_t = \omega + \beta_d\widehat{V}_{t-1} + \beta_w\widehat{V}_{w,t-1} + \beta_m\widehat{V}_{m,t-1} + \varepsilon_t, \tag{46}$$

which can be easily estimated via ordinary least squares (OLS). As demonstrated by numerous empirical applications in the literature, the implementation of a more refined volatility measure on the right-hand side (RHS) can better exploit the information, and has the potential to significantly improve the predictive accuracy of the HAR model for the left-hand side (LHS) target variable.

In addition to the implementation of better volatility measures in the standard HAR model, a constructive modification of the HAR model structure can also contribute to improved forecasting outcomes. In this section, we consider two important extensions of the original HAR-RV model for the comparative study. One is the HARQ model of Bollerslev et al. (2016). With the motivation that the persistence of RV is affected by the temporal variation in its measurement errors, the HARQ model allows for a time-varying coefficient for the previous day's RV on the RHS, and the coefficient

[9] We use the SAS code from Holden and Jacobsen (2014) to extract all tick-by-tick transaction records matched with relevant ask/bid quotes from the daily TAQ dataset of WRDS.

depends on the heteroskedasticity in the error, which is captured by the realized quarticity (RQ):[10]

$$\widehat{V}_t = \omega + \left(\beta_d + \beta_q \sqrt{\mathrm{RQ}_{t-1}}\right) \mathrm{RV}_{t-1} + \beta_w \mathrm{RV}_{w,t-1} + \beta_m \mathrm{RV}_{m,t-1} + \varepsilon_t, \quad \text{where } \mathrm{RQ}_t = \frac{n}{3} \sum_{i=1}^{n} r_{i,t}^4. \tag{47}$$

Inspired by the realized semivariance (RS) introduced by Barndorff-Nielsen et al. (2010), the semivariance HAR (SHAR) model of Patton and Sheppard (2015) stands out as another important HAR-RV modification:

$$\widehat{V}_t = \omega + \beta_d^- \mathrm{RS}_{t-1}^- + \beta_d^+ \mathrm{RS}_{t-1}^+ + \beta_w \mathrm{RV}_{w,t-1} + \beta_m \mathrm{RV}_{m,t-1} + \varepsilon_t, \tag{48}$$

where the RS measures are given by

$$\mathrm{RS}_t^- = \sum_{i=1}^{n} r_{i,t}^2 \mathbb{1}_{\{r_{i,t}<0\}} \quad \text{and} \quad \mathrm{RS}_t^+ = \sum_{i=1}^{n} r_{i,t}^2 \mathbb{1}_{\{r_{i,t}>0\}}. \tag{49}$$

The intuition that "good" and "bad" volatilities are not created equal motivates the decomposition of the original RV into separate up and downside RS measures. The empirical results in Patton and Sheppard (2015) demonstrate that this decomposition leads to more accurate volatility forecasts, with the "bad" volatility predominantly driving the short-run changes in the future.

6.3 Empirical Results

In this section, we estimate the standard HAR model in Eq. (46) with various IV measures on the RHS, namely RV, RBPV, TRV, DV, and our RRDV estimator, with an initial in-sample period of the first 1000 days, and forecast one-day-ahead out-of-sample RV, DV, and RRDV. Moreover, we estimate both the HARQ model in Eq. (47) and the SHAR model in Eq. (48) for the comparison of forecasts. We repeat this procedure of in-sample estimation and out-of-sample forecasting in both a rolling-window (RW) and an expanding-window (EW) fashion, respectively.

All return-based IV measures are constructed from log-returns over 5-minute intervals. For the construction of RRDV based on 5-minute candlesticks, we obtain the corresponding HLOCs from the transaction data, either at the tick level or under previous-tick equidistant sampling. In particular, for the discretized RRDV based on HLOCs from equistantly sampled data, we correct the discretization errors following the steps in Section 4.1.

We evaluate the out-of-sample forecasting performance via two widely used loss functions, i.e., the mean squared error (MSE) and the quasi-likelihood (QLIKE) function:

$$\mathrm{MSE}(\theta, h) = (\theta - h)^2 \quad \text{and} \quad \mathrm{QLIKE}(\theta, h) = \frac{\theta}{h} - \ln\left(\frac{\theta}{h}\right) - 1, \tag{50}$$

[10] Following Bollerslev et al. (2016), the "insanity filter" of Swanson and White (1997) is applied: For each rolling or expanding window, the minimum, maximum, and average of in-sample estimates are re-calculated. All one-step-ahead out-of-sample forecasts that are greater (smaller) than the maximum (minimum) in-sample value will be replaced by the in-sample mean.

where θ and h represent the actual value and the forecast of the target variable, respectively.

Table 3: Daily out-of-sample 5-minute HAR volatility forecasts

	RV		DV		RRDV: tick data	
	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE
Panel A: HAR RW Forecasts						
HAR-RV	2.39	0.41	1.57	0.47	1.19	0.30
HARQ	2.35	0.36	1.46	0.40	1.48	0.37
SHAR	2.29	0.39	1.51	0.45	1.04	0.30
HAR-RBPV	2.36	0.41	1.57	0.45	1.20	0.30
HAR-TRV	2.30	0.39	1.51	0.44	1.15	0.28
HAR-DV	2.55	0.40	1.68	0.45	1.32	0.29
HAR-DV$_{1-3}$	2.45	0.39	1.62	0.44	1.24	0.28
HAR-RRDV: tick data	2.17	0.37	1.42	0.42	1.05	0.27
HAR-RRDV: 1-second data	2.16	0.37	1.41	0.42	1.06	0.27
HAR-RRDV: 30-second data	2.09	0.37	1.36	0.42	0.96	0.27
HAR-RRDV: 1-minute data	2.03	0.36	1.31	0.41	0.89	0.27
Panel B: HAR EW Forecasts						
HAR-RV	2.29	0.37	1.51	0.42	1.15	0.29
HARQ	2.27	0.33	1.43	0.38	1.12	0.32
SHAR	2.18	0.35	1.45	0.42	1.02	0.30
HAR-RBPV	2.27	0.36	1.51	0.42	1.16	0.28
HAR-TRV	2.22	0.34	1.45	0.39	1.11	0.26
HAR-DV	2.45	0.35	1.62	0.39	1.27	0.26
HAR-DV$_{1-3}$	2.35	0.34	1.55	0.39	1.19	0.26
HAR-RRDV: tick data	2.10	0.33	1.38	0.38	1.03	0.26
HAR-RRDV: 1-second data	2.09	0.33	1.37	0.37	1.03	0.25
HAR-RRDV: 30-second data	2.03	0.33	1.31	0.37	0.94	0.25
HAR-RRDV: 1-minute data	1.96	0.32	1.26	0.36	0.88	0.25

MSE ($\times 10^8$) and QLIKE of daily out-of-sample volatility forecasts for the SPDR S&P 500 ETF Trust (SPY). The HAR model is re-estimated via OLS in rolling windows and expanding windows, respectively. The fixed (resp. initial) in-sample period for RW (resp. EW) estimation is the first 1000 days. All return-based IV measures are constructed from 5-minute intervals. RRDVs are also constructed from 5-minute candlesticks, in which the HLOCs are obtained from the transaction data either at the tick level or under previous-tick equidistant sampling. The choice of truncation parameters for TRV, DV and RRDV follows the instructions in Section 4.2, with $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}, C_\zeta^{\text{RRDV}}, \varpi) = (3, 3\sqrt{2}, 2, 1/2)$. For RRDVs based on equidistantly sampled observations, the discretization errors are corrected following the steps in Section 4.1.

Table 3 reports the MSE and QLIKE results for one-day-ahead out-of-sample forecasts of three different target volatility measures. Among the standard and modified HAR-RV models, both the HARQ and SHAR models can achieve smaller MSE and QLIKE results than the original HAR-RV model, which demonstrates that the consideration of either the measurement errors in RV or the volatility asymmetry helps to exploit concealed information due to aggregation, and results in more accurate forecasts. Compared with the HAR models augmented with other volatility estimators, the HAR-RRDV model tends to obtain substantially diminished values of both loss functions, and the number of observations available in each candlestick interval seems relatively irrelevant to its predictive capability. Furthermore, the symmetric MSE function penalizes outliers heavily, and is therefore sensitive to excessively misinformative forecasts. The MSE results in Table 3 suggest that the HAR-RRDV model can effectively reduce the occurrence of extremely inaccurate forecasts in both left and right tails.

To further explore the reason for the reduced forecast errors of HAR-RRDV, we partition the entire out-of-sample period into complimentary subsets of days based on two criteria, respectively: (i) days with and without jumps, as well as (ii) days exhibiting episodes of extreme return persistence or not. The presence of discontinuities and persistent noise is identified using the nonparametric tests of Aït-Sahalia et al. (2012) and Andersen et al. (2021), respectively.[11]

Table 4: MSEs on days with/without discontinuities and episodic extreme return persistence

	RV				DV				RRDV: tick data			
	Jumps		Per. Noise		Jumps		Per. Noise		Jumps		Per. Noise	
	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Panel A: HAR RW Forecasts												
HAR-RV	7.24	0.24	11.01	1.52	4.61	0.23	6.51	1.08	3.24	0.28	2.93	1.01
HARQ	7.05	0.27	10.73	1.51	4.16	0.26	6.34	0.97	4.13	0.31	2.86	1.34
SHAR	6.90	0.25	10.12	1.50	4.40	0.23	5.73	1.09	3.06	0.16	2.50	0.90
HAR-RBPV	7.17	0.23	10.92	1.50	4.61	0.22	6.45	1.07	3.36	0.24	2.89	1.03
HAR-TRV	6.93	0.25	10.81	1.45	4.39	0.24	6.38	1.02	3.10	0.28	2.84	0.98
HAR-DV	7.76	0.24	10.80	1.72	4.96	0.23	6.37	1.21	3.70	0.26	2.82	1.17
HAR-DV$_{1-3}$	7.42	0.25	10.76	1.61	4.75	0.24	6.34	1.14	3.45	0.26	2.84	1.08
HAR-RRDV: tick data	6.46	0.27	9.93	1.39	4.07	0.25	5.81	0.98	2.73	0.31	2.49	0.91
HAR-RRDV: 1-second data	6.45	0.26	9.77	1.39	4.06	0.24	5.70	0.98	2.68	0.33	2.43	0.92
HAR-RRDV: 30-second data	6.21	0.28	9.63	1.33	3.86	0.25	5.60	0.93	2.42	0.32	2.37	0.82
HAR-RRDV: 1-minute data	6.07	0.23	9.72	1.25	3.75	0.23	5.65	0.87	2.25	0.29	2.40	0.74
Panel B: HAR EW Forecasts												
HAR-RV	6.98	0.22	11.35	1.38	4.45	0.20	6.73	0.98	3.12	0.27	3.03	0.96
HARQ	6.83	0.25	10.99	1.39	4.15	0.23	6.49	0.93	2.99	0.29	2.91	0.94
SHAR	6.08	0.23	10.39	1.36	4.25	0.21	5.95	1.00	2.96	0.16	2.60	0.86
HAR-RBPV	6.94	0.21	11.22	1.38	4.47	0.20	6.65	0.99	3.24	0.24	2.97	0.97
HAR-TRV	6.72	0.23	11.12	1.32	4.27	0.21	6.58	0.94	3.00	0.27	2.93	0.93
HAR-DV	7.50	0.23	11.11	1.59	4.81	0.21	6.58	1.12	3.56	0.26	2.91	1.11
HAR-DV$_{1-3}$	7.16	0.25	11.12	1.47	4.59	0.21	6.57	1.05	3.31	0.26	2.94	1.02
HAR-RRDV: tick data	6.28	0.25	10.16	1.29	3.97	0.23	5.96	0.92	2.65	0.31	2.55	0.87
HAR-RRDV: 1-second data	6.25	0.25	10.00	1.30	3.94	0.23	5.86	0.91	2.60	0.33	2.49	0.88
HAR-RRDV: 30-second data	6.03	0.26	9.83	1.24	3.76	0.23	5.75	0.87	2.35	0.31	2.43	0.79
HAR-RRDV: 1-minute data	5.90	0.22	10.08	1.15	3.65	0.21	5.90	0.80	2.21	0.29	2.50	0.71

MSE ($\times 10^8$) of daily out-of-sample volatility forecasts for the SPDR S&P 500 ETF Trust (SPY) on the days with or without jumps and persistent noise. Jumps and persistent noise are identified with the nonparametric tests of Aït-Sahalia et al. (2012) and Andersen et al. (2021), respectively. The HAR model is re-estimated via OLS in rolling windows and expanding windows, respectively. The fixed (resp. initial) in-sample period for RW (resp. EW) estimation is the first 1000 days. All return-based IV measures are constructed from 5-minute intervals. RRDVs are also constructed from 5-minute candlesticks, in which the HLOCs are obtained from the transaction data either at the tick level or under previous-tick equidistant sampling. The choice of truncation parameters for TRV, DV and RRDV follows the instructions in Section 4.2, with $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}, C_\zeta^{\text{RRDV}}, \varpi) = (3, 3\sqrt{2}, 2, 1/2)$. For RRDVs based on equidistantly sampled observations, the discretization errors are corrected following the steps in Section 4.1.

In Table 4, the MSE results for all selected HAR models are presented within these classifications. There is a notable reduction in MSEs across all selected models on days without discontinuities or excessive return drift. This observation indicates that the presence of extreme events potentially

[11] For the test statistic of Aït-Sahalia et al. (2012), we select the pre-averaging window $k_n = \lfloor \sqrt{n} \rfloor$ and the truncation level $C = 5$. To identify the presence of persistent noise, we construct the test statistic $T_t^n(2)$ of Andersen et al. (2021) from one-minute pre-averaged and winsorized returns. The selected critical values for those two tests are -1.645 and 1.645, respectively.

distorts the estimation of dependence structures in volatility and consequently leads to uniformly worsened forecasts. Among the three target IV measures on the LHS, we find that the one-day-ahead forecasts of RRDV exhibit superior accuracy when there exist either "discontinuous" or "continuous" extreme events, with all chosen RHS variables. Meanwhile, the RV predictions are more vulnerable to both discontinuities and short-lived explosive trends, resulting in substantially larger forecast errors. For each of the target variables on the LHS, the HAR-RRDV model demonstrates the least vulnerability to extreme events and generates the most accurate one-day-ahead forecasts. These collective observations suggest that the robustness of our RRDV estimator in the presence of extreme price movements contributes to the predictive capability when it is integrated within some standard framework of volatility forecasting.

7 Conclusions

Motivated by both the statistical superiority of range-based estimation and the broader availability of intraday candlesticks for general investors, we introduce a novel nonparametric candlestick-based estimator for integrated variance (IV), namely the range-return-difference volatility (RRDV) estimator. The RRDV estimator is designed to mitigate the impact of short-lived explosive trends that locally dominate price movements, such as gradual jumps and flash crashes. By modeling these "continuous" extreme events from two perspectives: (i) a locally unbounded drift component (Christensen et al., 2022), and (ii) sticky expectations of market participants (Andersen et al., 2021), we demonstrate that RRDV can consistently estimate IV with variances about four times smaller than those obtained with the differenced-return volatility (DV) estimator introduced by Andersen et al. (2021). Our simulation results underscore the reliable finite-sample performance of RRDV across various relevant scenarios. An empirical illustration of volatility forecasting shows that the HAR-RRDV model can effectively reduce the occurrence of extremely misleading forecasts and improve forecasting accuracy according to standard out-of-sample loss functions.

References

Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

Aït-Sahalia, Y., Jacod, J., and Li, J. (2012). Testing for jumps in noisy high frequency data. *Journal of Econometrics*, 168(2):207–222.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, 18(2):351–416.

Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57(3):1047–1091.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Andersen, T. G., Li, Y., Todorov, V., and Zhou, B. (2021). Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*, forthcoming.

Andersen, T. G., Todorov, V., and Zhou, B. (2023). Real-time detection of local no-arbitrage violations. Working Paper.

Asmussen, S., Glynn, P., and Pitman, J. (1995). Discretization error in simulation of one-dimensional reflecting Brownian motion. *Annals of Applied Probability*, 5(4):875–896.

Bajgrowicz, P., Scaillet, O., and Treccani, A. (2016). Jumps in high-frequency data: Spurious detections, dynamics, and news. *Management Science*, 62(8):2198–2217.

Ball, C. A. and Torous, W. N. (1984). The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. *Journal of Business*, 57(1):97–112.

Bandi, F. M. and Renò, R. (2016). Price and volatility co-jumps. *Journal of Financial Economics*, 119(1):107–146.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, 12(3):1–32.

Barndorff-Nielsen, O. E., Kinnebrock, S., and Shephard, N. (2010). Measuring downside risk – realised semivariance. in T. Bollerslev, J. Russell, and M. Watson, eds., *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, New York: Oxford University Press, 2010.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.

Beckers, S. (1983). Variances of security price returns based on high, low, and closing prices. *Journal of Business*, 56(1):97–112.

Bellia, M., Christensen, K., Kolokolov, A., Pelizzon, L., and Renò, R. (2023). Do designated market makers provide liquidity during extreme price movements? Working Paper.

Bollerslev, T., Li, J., and Li, Q. (2023). Optimal nonparametric range-based volatility estimation. *Journal of Econometrics*, forthcoming.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192(1):1–18.

Brandt, M. W. and Diebold, F. X. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business*, 79(1):61–74.

CFTC and SEC (2010). Findings regarding the market events of May 6, 2010: Report of the staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. Available at https://www.sec.gov/files/marketevents-report.pdf.

Christensen, K. and Kolokolov, A. (2023). An unbounded intensity model for point process. Working Paper.

Christensen, K., Oomen, R., and Renò, R. (2022). The drift burst hypothesis. *Journal of Econometrics*, 227(2):461–497.

Christensen, K., Oomen, R. C., and Podolskij, M. (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114(3):576–599.

Christensen, K. and Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349.

Christensen, K. and Podolskij, M. (2012). Asymptotic theory of range-based multipower variation. *Journal of Financial Econometrics*, 10(3):417–456.

Christensen, K., Podolskij, M., and Vetter, M. (2009). Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance and Stochastics*, 13(2):239–268.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Corsi, F., Pirino, D., and Renò, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2):276–288.

Dieker, A. and Lagos, G. (2017). On the Euler discretization error of Brownian motion about random times. Working Paper.

Duembgen, M. and Podolskij, M. (2015). High-frequency asymptotics for path-dependent functionals of Itô semimartingales. *Stochastic Processes and their Applications*, 125(4):1195–1217.

Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance*, 58(3):1269–1300.

Feller, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *Annals of Mathematical Statistics*, 22(3):427–432.

Flora, M. and Renò, R. (2022). V-shapes. Working Paper.

Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1):67–78.

Golub, A., Keane, J., and Poon, S.-H. (2012). High frequency trading and mini flash crashes. Working Paper.

Grigelionis, B. (1980). A martingale approach to the statistical problems of point processes. *Scandinavian Journal of Statistics*, 7(4):190–196.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Hoffmann, M., Vetter, M., and Dette, H. (2018). Nonparametric inference of gradual changes in the jump behaviour of time-continuous processes. *Stochastic Processes and their Applications*, 128(11):3679–3723.

Holden, C. W. and Jacobsen, S. (2014). Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *Journal of Finance*, 69(4):1747–1785.

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4):456–499.

Jacod, J., Li, Y., and Zheng, X. (2019). Estimating the integrated volatility with tick observations. *Journal of Econometrics*, 208(1):80–100.

Jacod, J. and Todorov, V. (2010). Do price and volatility jump together? *Annals of Applied Probability*, 20(4):1425–1469.

Jacod, J. and Todorov, V. (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *Annals of Statistics*, 42(3):1029–1069.

Jacod, J. and Todorov, V. (2018). Limit theorems for integrated local empirical characteristic exponents from noisy high-frequency data with application to volatility and jump activity estimation. *Annals of Applied Probability*, 28(1):511–576.

Kirilenko, A., Kyle, A. S., Samadi, M., and Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *Journal of Finance*, 72(3):967–998.

Kolokolov, A. (2023). Cryptocrashes. Working Paper.

Kolokolov, A., Renò, R., and Zoi, P. (2023). BUMVU estimators. Working Paper.

Kunitomo, N. (1992). Improving the Parkinson method of estimating security price volatilities. *Journal of Business*, 65(2):295–302.

Laly, F. and Petitjean, M. (2020). Mini flash crashes: Review, taxonomy and policy responses. *Bulletin of Economic Research*, 72(3):251–271.

Laurent, S., Renò, R., and Shi, S. (2022). Realized drift. Working Paper.

Laurent, S. and Shi, S. (2020). Volatility estimation and jump detection for drift-diffusion processes. *Journal of Econometrics*, 217(2):259–290.

Li, J., Wang, D., and Zhang, Q. (2022). Reading the candlesticks: An OK estimator for volatility. *Review of Economics and Statistics*, forthcoming.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.

Martens, M. and van Dijk, D. (2007). Measuring volatility with the realized range. *Journal of Econometrics*, 138(1):181–207.

Meilijson, I. (2011). The Garman-Klass volatility estimator revisited. *Revstat*, 9(3):199–212.

Menkveld, A. J. and Yueshen, B. Z. (2019). The flash crash: A cautionary tale about highly fragmented markets. *Management Science*, 65(10):4470–4488.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65.

Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3):683–697.

Rogers, L. C. G. and Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability*, 1(4):504–512.

Swanson, N. R. and White, H. (1997). Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. *International Journal of Forecasting*, 13(4):439–461.

Todorov, V. and Tauchen, G. (2011). Volatility jumps. *Journal of Business & Economic Statistics*, 29(3):356–371.

Vetter, M. (2010). Limit theorems for bipower variation of semimartingales. *Stochastic Processes and their Applications*, 120(1):22–38.

Yang, D. and Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, 73(3):477–492.

Appendix A Normalized High, Low and Close

For the standard Brownian motion starting at zero, i.e., $W = (W_t)_{t\geq 0}$, in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, we denote the normalized high, low, and close as, respectively,

$$u = \sup_{0\leq t\leq 1} W_t, \qquad d = \inf_{0\leq t\leq 1} W_t, \qquad c = W_1. \tag{A.1}$$

For the range of a standard Brownian motion, i.e., $\omega = u - d$, its probability distribution was firstly proposed by Feller (1951) and its moment generating function was then derived by Parkinson (1980), i.e., for the r-th moment:

$$\mathbb{E}[\omega^r] = \frac{4}{\sqrt{\pi}}\left(1 - \frac{4}{2^r}\right) 2^{\frac{r}{2}} \Gamma\left(\frac{r+1}{2}\right) \zeta(r-1), \tag{A.2}$$

where $\Gamma(x)$ and $\zeta(x)$ are the Gamma and Riemann's zeta functions, respectively. In particular, we have

$$\begin{aligned} \mathbb{E}[\omega] &= 2\sqrt{\frac{2}{\pi}} \approx 1.5958, \qquad & \mathbb{E}[\omega^2] &= 4\ln 2 \approx 2.7726, \\ \mathbb{E}[\omega^3] &= \frac{2}{3}\sqrt{2\pi^3} \approx 5.2499, \qquad & \mathbb{E}[\omega^4] &= 9\zeta(3) \approx 10.8185. \end{aligned} \tag{A.3}$$

Also, Garman and Klass (1980) reveals the following fourth moments of the normalized high, low, close via the generating functions $\mathbb{E}[u^p d^q c^r]$:[12]

$$\mathbb{E}[u^4] = \mathbb{E}[d^4] = \mathbb{E}[c^4] = 3, \qquad \mathbb{E}[u^2c^2] = \mathbb{E}[d^2c^2] = 2, \tag{A.4}$$

$$\mathbb{E}[u^3c] = \mathbb{E}[d^3c] = 2.25, \qquad \mathbb{E}[uc^3] = \mathbb{E}[dc^3] = 1.5, \tag{A.5}$$

$$\mathbb{E}[u^2dc] = \mathbb{E}[ud^2c] = \frac{9}{4} - 2\ln 2 - \frac{7}{8}\zeta(3) \approx -0.1881, \tag{A.6}$$

$$\mathbb{E}[u^2d^2] = 3 - 4\ln 2 \approx 0.2274, \tag{A.7}$$

$$\mathbb{E}[udc^2] = 2 - 2\ln 2 - \frac{7}{8}\zeta(3) \approx -0.4381, \tag{A.8}$$

$$\mathbb{E}[ud^3] = \mathbb{E}[u^3d] = 3 - 3\ln 2 - \frac{9}{8}\zeta(3) \approx -0.4318, \tag{A.9}$$

where $\zeta(3) = \sum_{n=1}^{\infty} n^{-3} \approx 1.2021$. It is straightforward that

$$\mathbb{E}[\omega^2c^2] = \mathbb{E}[(u-d)^2c^2] = \mathbb{E}[u^2c^2] + \mathbb{E}[d^2c^2] - 2\mathbb{E}[udc^2] = 4\ln 2 + \frac{7}{4}\zeta(3) \approx 4.8762. \tag{A.10}$$

When we substitute the normalized close c in above moments with its absolute value $|c|$, it is obvious that the values in Eqs. (A.5) and (A.6) do not follow from Garman and Klass (1980). Different from the Garman–Klass triple (u, d, c), Meilijson (2011) considers $(\tilde{u}, \tilde{d}, |c|)$ where $(\tilde{u}, \tilde{d}) = (u, d)$ if

[12] See Appendix C in Garman and Klass (1980).

$c \geq 0$ while $(\tilde{u}, \tilde{d}) = -(d, u)$ if $c < 0$, and derives the second and fourth moments as follows:

$$\mathbb{E}[\tilde{u}^2] = \frac{7}{4}, \quad \mathbb{E}[\tilde{d}^2] = \frac{1}{4}, \quad \mathbb{E}[\tilde{u}|c|] = \frac{5}{4}, \quad \mathbb{E}[\tilde{d}|c|] = -\frac{1}{4}, \quad \mathbb{E}[\tilde{u}\tilde{d}] = 1 - 2\ln 2 \approx -0.3863, \tag{A.11}$$

$$\mathbb{E}[\tilde{u}^4] = \frac{93}{16}, \quad \mathbb{E}[\tilde{d}^4] = \frac{3}{16}, \quad \mathbb{E}[\tilde{u}^2|c|^2] = \frac{31}{8}, \quad \mathbb{E}[\tilde{d}^2|c|^2] = \frac{1}{8}, \tag{A.12}$$

$$\mathbb{E}[\tilde{u}^3|c|] = \frac{147}{32}, \quad \mathbb{E}[\tilde{d}^3|c|] = -\frac{3}{32}, \quad \mathbb{E}[\tilde{u}|c|^3] = \frac{27}{8}, \quad \mathbb{E}[\tilde{d}|c|^3] = -\frac{3}{8}, \tag{A.13}$$

$$\mathbb{E}[\tilde{u}^2\tilde{d}^2] = \mathbb{E}[u^2 d^2] = 3 - 4\ln 2 \approx 0.2274, \tag{A.14}$$

$$\mathbb{E}[\tilde{u}\tilde{d}|c|^2] = \mathbb{E}[udc^2] = 2 - 2\ln 2 - \frac{7}{8}\zeta(3) \approx -0.4381, \tag{A.15}$$

$$\mathbb{E}[\tilde{u}^3\tilde{d}] + \mathbb{E}[\tilde{u}\tilde{d}^3] = \mathbb{E}[ud(u^2 + d^2)] = 6 - 6\ln 2 - \frac{9}{4}\zeta(3) \approx -0.8635, \tag{A.16}$$

$$\mathbb{E}[\tilde{u}^2\tilde{d}|c|] + \mathbb{E}[\tilde{u}\tilde{d}^2|c|] = \mathbb{E}[udc(u + d)] = \frac{9}{2} - 4\ln 2 - \frac{7}{4}\zeta(3) \approx -0.3762, \tag{A.17}$$

$$\mathbb{E}[\tilde{u}\tilde{d}^2|c|] = \frac{1}{16}\zeta(3) - 2\ln 2 + \frac{47}{32} \approx 0.1576. \tag{A.18}$$

We can use the above results to obtain the following second and fourth moments of $(\omega, |c|)$:

$$\mathbb{E}[\omega|c|] = \mathbb{E}[(\tilde{u} - \tilde{d})|c|] = \mathbb{E}[\tilde{u}|c|] - \mathbb{E}[\tilde{d}|c|] = \frac{3}{2}, \tag{A.19}$$

$$\mathbb{E}[\omega|c|^3] = \mathbb{E}[(\tilde{u} - \tilde{d})|c|^3] = \mathbb{E}[\tilde{u}|c|^3] - \mathbb{E}[\tilde{d}|c|^3] = \frac{15}{4}, \tag{A.20}$$

$$\begin{aligned}
\mathbb{E}[\omega^3|c|] &= \mathbb{E}[(\tilde{u} - \tilde{d})^3|c|] \\
&= \mathbb{E}[(\tilde{u}^3 - \tilde{d}^3 - 3\tilde{u}^2\tilde{d} + 3\tilde{u}\tilde{d}^2)|c|] \\
&= \mathbb{E}[\tilde{u}^3|c|] - \mathbb{E}[\tilde{d}^3|c|] - 3\mathbb{E}[\tilde{u}^2\tilde{d}|c|] + 3\mathbb{E}[\tilde{u}\tilde{d}^2|c|] \\
&= \frac{147}{32} + \frac{3}{32} - 3\left(\frac{9}{2} - 4\ln 2 - \frac{7}{4}\zeta(3)\right) + 6\left(\frac{1}{16}\zeta(3) - 2\ln 2 + \frac{47}{32}\right) \\
&= \frac{45}{8}\zeta(3) \approx 6.7616.
\end{aligned} \tag{A.21}$$

To calculate the third moments of $(\omega, |c|)$, we derive the analytical expressions for $\mathbb{E}[\tilde{u}^2|c|]$, $\mathbb{E}[\tilde{d}^2|c|]$, $\mathbb{E}[\tilde{u}|c|^2]$, $\mathbb{E}[\tilde{d}|c|^2]$, and $\mathbb{E}[\tilde{u}\tilde{d}|c|]$, which are not available in the literature. For the first four quantities, we obtain the results by integrating the joint densities in Meilijson (2011), i.e.,

$$f_{\tilde{u},|c|}(a, x) = 4(2a - x)\phi(2a - x), \quad 0 < x < a, \tag{A.22}$$

$$f_{\tilde{d},|c|}(b, x) = 4(x - 2b)\phi(x - 2b), \quad b < 0 < x, \tag{A.23}$$

where $\phi(z) = (2\pi)^{-1/2}e^{-z^2/2}$ is the PDF of $\mathcal{N}(0, 1)$:

$$\mathbb{E}[\tilde{u}^2|c|] = \frac{17}{3\sqrt{2\pi}} \approx 2.2607, \quad \mathbb{E}[\tilde{d}^2|c|] = \frac{1}{3\sqrt{2\pi}} \approx 0.1330, \tag{A.24}$$

$$\mathbb{E}[\tilde{u}|c|^2] = \frac{7}{3}\sqrt{\frac{2}{\pi}} \approx 1.8617, \quad \mathbb{E}[\tilde{d}|c|^2] = -\frac{1}{3}\sqrt{\frac{2}{\pi}} \approx -0.2660. \tag{A.25}$$

There is one more moment needed, i.e., $\mathbb{E}[\tilde{u}\tilde{d}|c|]$. We start with the infinitesimal event $A = \{W_1 \in (x, x+dx), W_t \in (b, a), \forall t \in [0,1]\}$, where $b < \min\{x, 0\} \leq 0 \leq \max\{x, 0\} < a$, and its probability $\mathbb{P}(A) = Q(a, b, x)dx$, where

$$Q(a,b,x) = \sum_{j=-\infty}^{\infty} \{\phi(x - 2j(a-b)) - \phi(x - 2b - 2j(a-b))\}. \tag{A.26}$$

The joint density of $(\tilde{u}, \tilde{d}, |c|)$ is then given by $f_{\tilde{u},\tilde{d},|c|}(a,b,x) = -2\partial^2 Q(a,b,x)/\partial a \partial b$, restricted to $b < 0 < x < a$, which is also an infinite series.[13] The summand with $j = 0$ takes value 0 because both two ϕ functions are independent of at least one of a and b, as similar to the second term in the summand with $j = 1$. The required moment can be obtained by solving the triple integral:

$$\begin{aligned}
\mathbb{E}[\tilde{u}\tilde{d}|c|] &= -2\int_0^{\infty}\int_0^{a}\int_{-\infty}^{0} abx \frac{\partial^2 Q(a,b,x)}{\partial a \partial b} dbdxda \\
&= -2\sum_{j\in\mathbb{Z}\setminus\{0\}} \int_0^{\infty} ada \int_0^{a} xdx \int_{-\infty}^{0} \frac{\partial}{\partial a} b \left[\frac{\partial}{\partial b}\phi(x - 2j(a-b)) - \frac{\partial}{\partial b}\phi(x - 2b - 2j(a-b))\mathbb{1}_{\{j\neq 1\}}\right] db
\end{aligned} \tag{A.27}$$

We integrate each summand in three univariate steps. The first step will integrate over $b \in (-\infty, 0)$ the product of b and mixed second derivative $\partial^2\phi(x + Ma + Kb)/\partial a \partial b$:

$$\begin{aligned}
\int_{-\infty}^{0} \frac{\partial}{\partial a} b \frac{\partial}{\partial b}\phi(x + Ma + Kb)db &= \frac{\partial}{\partial a}\int_{-\infty}^{0} b\frac{\partial}{\partial b}\phi(x + Ma + Kb)db \\
&= \frac{\partial}{\partial a}\int_{-\infty}^{0} bd\phi(x + Ma + Kb) \\
&= \frac{\partial}{\partial a}\left[b\phi(x + Ma + Kb)\right]_{-\infty}^{0} - \frac{\partial}{\partial a}\int_{-\infty}^{0}\phi(x + Ma + Kb)db \\
&= -\int_{-\infty}^{0}\frac{\partial}{\partial a}\phi(x + Ma + Kb)db \\
&= -M\int_{-\infty}^{0}\phi'(x + Ma + Kb)db \\
&= -\frac{M}{K}\left[\phi(x + Ma + Kb)\right]_{-\infty}^{0} \\
&= -\frac{M}{K}\phi(x + Ma).
\end{aligned} \tag{A.28}$$

[13] See more details in the Appendix of Meilijson (2011).

Then we multiply the above result by x and integrate it over $x \in (0, a)$:

$$\begin{aligned}
&\int_0^a x dx \int_{-\infty}^0 \frac{\partial}{\partial a} b \frac{\partial}{\partial b} \phi(x + Ma + Kb) db \\
=& -\frac{M}{K} \int_0^a x\phi(x + Ma) dx \\
=& -\frac{M}{K} \int_{Ma}^{(M+1)a} y\phi(y) dy + \frac{M^2 a}{K} \int_{Ma}^{(M+1)a} \phi(y) dy \\
=& \frac{M}{K} \int_{Ma}^{(M+1)a} \phi'(y) dy + \frac{M^2 a}{K} \int_{Ma}^{(M+1)a} \phi(y) dy \qquad \text{because } \phi'(z) = -z\phi(z) \\
=& \frac{M}{K} \left[\phi((M+1)a) - \phi(Ma)\right] + \frac{M^2 a}{K} \left[\Phi((M+1)a) - \Phi(Ma)\right] \\
\text{or } =& \frac{M}{K} \left[\phi((M+1)a) - \phi(Ma)\right] - \frac{M^2 a}{K} \left[\Phi^*((M+1)a) - \Phi^*(Ma)\right],
\end{aligned} \tag{A.29}$$

where $\Phi(z) = \int_{-\infty}^z \phi(t) dt = 0.5(1 + \operatorname{erf} z/\sqrt{2})$ is the CDF of $\mathcal{N}(0, 1)$, and $\Phi^*(z) = 1 - \Phi(z) = 0.5(1 - \operatorname{erf} z/\sqrt{2})$ is the survival function. Finally, this expression is multiplied by a and integrated over $a \in (0, \infty)$. We use the results

$$\int_0^\infty a\phi(aA) da = \int_0^\infty a\phi(-aA) da = \frac{1}{\sqrt{2\pi} A^2}, \tag{A.30}$$

$$\int_0^\infty a^2 \Phi(-aA) da = \int_0^\infty a^2 \Phi^*(aA) da = \frac{1}{3A^3} \sqrt{\frac{2}{\pi}}, \quad \text{with } A > 0, \tag{A.31}$$

to calculate the triple integral of $abx\partial^2 \phi(x + Ma + Kb)/\partial a \partial b$. When $M \in \mathbb{Z}^+$, we have

$$\begin{aligned}
&\int_0^\infty a da \int_0^a x dx \int_{-\infty}^0 \frac{\partial}{\partial a} b \frac{\partial}{\partial b} \phi(x + Ma + Kb) db \\
=& \frac{M}{K} \int_0^\infty a\phi((M+1)a) da - \frac{M}{K} \int_0^\infty a\phi(Ma) da + \frac{M^2}{K} \int_0^\infty a^2 \Phi((M+1)a) da - \frac{M^2}{K} \int_0^\infty a^2 \Phi(Ma) da \\
=& \frac{1}{\sqrt{2\pi}} \frac{M}{K} \left[\frac{1}{(M+1)^2} - \frac{1}{M^2}\right] - \frac{1}{3} \sqrt{\frac{2}{\pi}} \frac{M^2}{K} \left[\frac{1}{(M+1)^3} - \frac{1}{M^3}\right] = \mathcal{G}(M, K).
\end{aligned} \tag{A.32}$$

When $M \in \mathbb{Z}^- \backslash \{-1\}$, we have

$$\begin{aligned}
&\int_0^\infty a da \int_0^a x dx \int_{-\infty}^0 \frac{\partial}{\partial a} b \frac{\partial}{\partial b} \phi(x + Ma + Kb) db \\
=& \frac{M}{K} \int_0^\infty a\phi((M+1)a) da - \frac{M}{K} \int_0^\infty a\phi(Ma) da - \frac{M^2}{K} \int_0^\infty a^2 \Phi^*((M+1)a) da + \frac{M^2}{K} \int_0^\infty a^2 \Phi^*(Ma) da \\
=& \frac{1}{\sqrt{2\pi}} \frac{M}{K} \left[\frac{1}{(M+1)^2} - \frac{1}{M^2}\right] - \frac{1}{3} \sqrt{\frac{2}{\pi}} \frac{M^2}{K} \left[\frac{1}{(M+1)^3} - \frac{1}{M^3}\right] = \mathcal{G}(M, K).
\end{aligned} \tag{A.33}$$

We now transfer each summand into a rational function of j, by letting M take $-2j$, and K take $2j$ or $2(j-1)$. For summands with $j \in \mathbb{Z}\backslash\{0,1\}$, we have

$$\begin{aligned}
&\int_0^\infty ada \int_0^a xdx \int_{-\infty}^0 \frac{\partial}{\partial a} b \left[\frac{\partial}{\partial b}\phi(x-2j(a-b)) - \frac{\partial}{\partial b}\phi(x-2b-2j(a-b)) \right] db \\
&= \mathcal{G}(-2j,2j) - \mathcal{G}(-2j,2(j-1)) \\
&= -\frac{1}{\sqrt{2\pi}}\left(1-\frac{j}{j-1}\right)\left[\frac{1}{(1-2j)^2} - \frac{1}{(2j)^2}\right] - \frac{1}{3}\left(2j - \frac{2j^2}{j-1}\right)\sqrt{\frac{2}{\pi}}\left[\frac{1}{(1-2j)^3} + \frac{1}{(2j)^3}\right] \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{j-1}\left[\frac{1}{(1-2j)^2} - \frac{1}{(2j)^2}\right] + \frac{2}{3}\frac{j}{j-1}\sqrt{\frac{2}{\pi}}\left[\frac{1}{(1-2j)^3} + \frac{1}{(2j)^3}\right],
\end{aligned} \tag{A.34}$$

and the infinite series

$$\sum_{j\in\mathbb{Z}\backslash\{0,1\}} \int_0^\infty ada \int_0^a xdx \int_{-\infty}^0 \frac{\partial}{\partial a} b \left[\frac{\partial}{\partial b}\phi(x-2j(a-b)) - \frac{\partial}{\partial b}\phi(x-2b-2j(a-b)) \right] db = \frac{14\pi^2-138}{72\sqrt{2\pi}}. \tag{A.35}$$

For summand with $j=1$, we have

$$\int_0^\infty ada \int_0^a xdx \int_{-\infty}^0 \frac{\partial}{\partial a} b \frac{\partial}{\partial b}\phi(x-2a+2b)db = \mathcal{G}(-2,2) = \frac{5}{12\sqrt{2\pi}}. \tag{A.36}$$

Therefore, the joint moment is calculated by Eq. (A.27):

$$\mathbb{E}[\tilde{u}\tilde{d}|c|] = -2\left(\frac{5}{12\sqrt{2\pi}} + \frac{14\pi^2-138}{72\sqrt{2\pi}}\right) = \frac{54-7\pi^2}{18\sqrt{2\pi}} \approx -0.3344. \tag{A.37}$$

Now we can use the results in Eqs. (A.24), (A.25) and (A.37) to calculate the third moments of $(\omega, |c|)$:

$$\begin{aligned}
\mathbb{E}[\omega^2|c|] &= \mathbb{E}[\tilde{u}^2|c|] + \mathbb{E}[\tilde{d}^2|c|] - 2\mathbb{E}[\tilde{u}\tilde{d}|c|] \\
&= \frac{17}{3\sqrt{2\pi}} + \frac{1}{3\sqrt{2\pi}} - 2\times\frac{54-7\pi^2}{18\sqrt{2\pi}} \\
&= \frac{7}{9}\sqrt{\frac{\pi^3}{2}} \approx 3.0624,
\end{aligned} \tag{A.38}$$

$$\mathbb{E}[\omega|c|^2] = \mathbb{E}[\tilde{u}|c|^2] - \mathbb{E}[\tilde{d}|c|^2] = \frac{8}{3}\sqrt{\frac{2}{\pi}} \approx 2.1277. \tag{A.39}$$

Moreover, direct calculation of

$$\mathbb{E}[|c|^r] = 2\int_0^\infty \frac{x^r}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \tag{A.40}$$

shows the following moments:

$$\mathbb{E}[|c|] = \sqrt{\frac{2}{\pi}} \approx 0.7979, \quad \mathbb{E}[|c|^2] = 1, \quad \mathbb{E}[|c|^3] = 2\sqrt{\frac{2}{\pi}} \approx 1.5958, \quad \mathbb{E}[|c|^4] = 3. \tag{A.41}$$

For the convenience of further discussion, we denote the moments of $(\omega, |c|)$ by

$$\lambda_{p,r} = \mathbb{E}[\omega^p |c|^r]. \tag{A.42}$$

All $\lambda_{p,r}$ that will be used in Appendix B have been calculated in this section, and are summarized in Table A.1.

Table A.1: Analytical values of $\lambda_{p,r} = \mathbb{E}[\omega^p |c|^r]$

p \ r	0	1	2	3	4
0	1	$\sqrt{\frac{2}{\pi}}$	1	$2\sqrt{\frac{2}{\pi}}$	3
1	$2\sqrt{\frac{2}{\pi}}$	$\frac{3}{2}$	$\frac{8}{3}\sqrt{\frac{2}{\pi}}$	$\frac{15}{4}$	–
2	$4\ln 2$	$\frac{7}{9}\sqrt{\frac{\pi^3}{2}}$	$4\ln 2 + \frac{7}{4}\zeta(3)$	–	–
3	$\frac{2}{3}\sqrt{2\pi^3}$	$\frac{45}{8}\zeta(3)$	–	–	–
4	$9\zeta(3)$	–	–	–	–

This table reports the analytical values of joint moments of $(\omega, |c|)$, i.e., $\lambda_{p,r} = \mathbb{E}[\omega^p |c|^r]$ with $p, r \in \mathbb{N}$ and $0 \leq p + r \leq 4$, where ω (resp. c) is defined as the high-low range (resp. open-close return) of a standard Brownian motion within an unit interval.

Appendix B Proofs of Theorems and Propositions

B.1 Proof of Theorem 1

Proof. It is straightforward to prove Theorem 1 with the LLN for path-dependent functionals of continuous Itô semimartingales, as summarized in Duembgen and Podolskij (2015). Here we start with some notation. We denote by $C([0,1])$ the space of continuous real valued functions on the interval $[0,1]$, and by $\|\cdot\|_\infty$ the supremum norm on $C([0,1])$. A function $f : C([0,1]) \to \mathbb{R}$ is said to have polynomial growth if $|f(x)| \leq C(1 + \|x\|_\infty^p)$ for some $C, p > 0$.

Definition B.1 (Local uniform continuity)**.** The function $f : C([0,1]) \to \mathbb{R}$ is locally uniformly continuous if for all $x \in C([0,1])$, there exists a closed ball of radius $K > 0$ centred at 0, i.e., $B_{\leq K}(0)=\{x \in C([0,1]); \|x\|_\infty \leq K\}$,[14] such that for every $\epsilon > 0$, there exists $\delta > 0$, for $x, y \in B_{\leq K}(0)$, $\|x - y\|_\infty \leq \delta$, we have $|f(x) - f(y)| \leq \epsilon$. This locally uniform continuity assumption is satisfied whenever $|f(x) - f(y)| \leq C\|x - y\|_\infty^p$ for all $x, y \in C([0,1])$ and some $C, p > 0$.

Lemma B.1 (Theorem 2.1, Duembgen and Podolskij, 2015)**.** Assume that the efficient price X follows a continuous Itô semimartingale in Eq. (5) with all traditional conditions satisfied. Given a function $g : C([0,1]) \to \mathbb{R}$ and a vanishing sequence Δ_n, for the sequence of processes

$$\widehat{V}_{t,n}(g) = \Delta_n \sum_{i=1}^{n} g\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right), \tag{B.1}$$

with $d_i^n(X) = \left\{X_{(i-1+s)\Delta_n} - X_{(i-1)\Delta_n}; s \in [0,1]\right\}$, if g is locally uniformly continuous and has polynomial growth, it holds that

$$\widehat{V}_{t,n}(g) \xrightarrow{\text{u.c.p.}} V_t(g) = \int_0^t \rho_{\sigma_\tau}(g) d\tau, \tag{B.2}$$

as $n \to \infty$, where $\rho_z(g) = \mathbb{E}\left[g(\{zW_s; s \in [0,1]\})\right]$ whenever it is finite.

It is obvious that the new HLOC estimator $\widehat{V}_{t,n}$ in Eq. (3) can be written in the form of Eq. (B.1) with a specific path-dependent function $g : C([0,1]) \to \mathbb{R}$ of the scaled incremental process:

$$\begin{aligned} g\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right) &= \frac{1}{4\ln 2 - 2}\left\{\sup_{0\leq s\leq 1}\frac{d_i^n(X)}{\sqrt{\Delta_n}} - \inf_{0\leq s\leq 1}\frac{d_i^n(X)}{\sqrt{\Delta_n}} - \frac{\left|X_{i\Delta_n} - X_{(i-1)\Delta_n}\right|}{\sqrt{\Delta_n}}\right\}^2 \\ &= \frac{1}{4\ln 2 - 2}\left\{f_1\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right) - f_2\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right)\right\}^2, \end{aligned} \tag{B.3}$$

where

$$f_1(x) = \sup_{0\leq s\leq 1} x(s) - \inf_{0\leq s\leq 1} x(s) \quad \text{and} \quad f_2(x) = |x(1) - x(0)|. \tag{B.4}$$

[14] The notion of locally uniform continuity is slightly different from the usual one that requires uniform continuity on neighbourhoods or compact sets, see more details in Remark 2.1 in Duembgen and Podolskij (2015).

The function $g(x)$ is therefore a linear combination of polynomials of the range $f_1(x)$ and a finite power variation $f_2(x)$, as well as the cross term $f_1(x)f_2(x)$. This path-dependent function has polynomial growth, and is thus locally uniformly continuous. Then the LLN in Lemma B.1 readily applies with

$$
\begin{aligned}
\int_0^t \rho_{\sigma_\tau}(g)d\tau &= \int_0^t \mathbb{E}\left[g(\{\sigma_\tau W_s; s\in[0,1]\})\right]d\tau \\
&= \frac{1}{4\ln 2-2}\int_0^t \mathbb{E}\left[\left(f_1(\{\sigma_\tau W_s; s\in[0,1]\}) - f_2(\{\sigma_\tau W_s; s\in[0,1]\})\right)^2\right]d\tau \\
&= \frac{1}{4\ln 2-2}\int_0^t \mathbb{E}\left[\left(\sup_{0\le s\le 1}\sigma_\tau W_s - \inf_{0\le s\le 1}\sigma_\tau W_s - \sigma_\tau\left|W_1 - W_0\right|\right)^2\right]d\tau \\
&= \frac{1}{4\ln 2-2}\int_0^t \mathbb{E}\Bigg[\left(\sup_{0\le s\le 1}\sigma_\tau W_s - \inf_{0\le s\le 1}\sigma_\tau W_s\right)^2 + \sigma_\tau^2 W_1^2 \\
&\quad - 2\left(\sup_{0\le s\le 1}\sigma_\tau W_s - \inf_{0\le s\le 1}\sigma_\tau W_s\right)\sigma_\tau\left|W_1\right|\Bigg]d\tau \\
&= \frac{1}{4\ln 2-2}\int_0^t \sigma_\tau^2\,\mathbb{E}\left[\omega^2 + c^2 - 2\omega|c|\right]d\tau \\
&= \frac{1}{4\ln 2-2}\int_0^t \sigma_\tau^2\left(\lambda_{2,0}+\lambda_{0,2}-2\lambda_{1,1}\right)d\tau \\
&= \frac{1}{4\ln 2-2}\int_0^t \sigma_\tau^2\left(4\ln 2+1-2\times\frac{3}{2}\right)d\tau \\
&= \int_0^t \sigma_\tau^2 d\tau,
\end{aligned}
\tag{B.5}
$$

where ω, c, and $\lambda_{p,r} = \mathbb{E}[\omega^p|c|^r]$ are defined in Appendix A. This completes the proof.

□

B.2 Proof of Theorem 2

Proof. We denote by $f'_y(x)$ the Gâteaux derivative of f at point x in the direction of y, i.e.,

$$
f'_y(x) = \lim_{h\to 0}\frac{f(x+hy)-f(x)}{h}. \tag{B.6}
$$

Lemma B.2 (Theorem 2.2, Duembgen and Podolskij, 2015)**.** Assume that the conditions of Lemma B.1 hold and Assumption 1 is satisfied. If $g'_y(x)$ for some $\|y\|_\infty \le 1$ is (i) locally uniformly continuous, and (ii) has polynomial growth, it follows that as $n\to\infty$,

$$
\frac{1}{\sqrt{\Delta_n}}\left(\widehat{V}_{t,n}(g) - V_t(g)\right) \xrightarrow{\mathcal{L}-\mathrm{s}} U_t(g), \tag{B.7}
$$

where $U_t(g) = \int_0^t u_\tau^{(1)} d\tau + \int_0^t u_\tau^{(2)} dW_\tau + \int_0^t u_\tau^{(3)} dW'_\tau$ with

$$\begin{aligned} u_\tau^{(1)} &= \mu_\tau \rho_{\sigma_\tau}^{(2)}(g') + \frac{1}{2}\tilde{\sigma}_\tau \rho_{\sigma_\tau}^{(3)}(g') - \frac{1}{2}\tilde{\sigma}_\tau \rho_{\sigma_\tau}^{(2)}(g'), \\ u_\tau^{(2)} &= \rho_{\sigma_\tau}^{(1)}(g), \\ u_\tau^{(3)} &= \sqrt{\rho_{\sigma_\tau}(g^2) - \rho_{\sigma_\tau}^2(g) - (\rho_{\sigma_\tau}^{(1)}(g))^2}, \end{aligned} \tag{B.8}$$

and, for $z \in \mathbb{R}$ and $G(x, y) = g'_y(x)$,

$$\begin{aligned} \rho_z^{(1)}(g) &= \mathbb{E}\left[g(\{zW_s; s \in [0,1]\})W_1\right], \\ \rho_z^{(2)}(g') &= \mathbb{E}\left[G(\{zW_s; s \in [0,1]\}, \{s; s \in [0,1]\})\right], \\ \rho_z^{(3)}(g') &= \mathbb{E}\left[G(\{zW_s; s \in [0,1]\}, \{W_s^2; s \in [0,1]\})\right]. \end{aligned} \tag{B.9}$$

The process $W' = (W'_t)_{t\geq 0}$ is a Brownian motion defined on an extension of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which is independent of $\mathcal{F}$. This is especially the case when g is an even function, i.e., $g(x) = g(-x)$ for all $x \in C([0,1])$, where it holds that

$$\rho_z^{(1)}(g) = \rho_z^{(2)}(g') = \rho_z^{(3)}(g') = 0, \tag{B.10}$$

for all $z \in \mathbb{R}$, since $W \overset{\mathcal{L}}{=} -W$ and expectations of odd functionals of W are 0, and hence we have

$$U_t(g) = \int_0^t \sqrt{\rho_{\sigma_\tau}(g^2) - \rho_{\sigma_\tau}^2(g)}\, dW'_\tau. \tag{B.11}$$

which is an $\mathcal{F}$-conditional Gaussian martingale with mean 0.

As mentioned in Appendix B.1, the path-dependent function $g : C([0,1]) \to \mathbb{R}$ in Eq. (B.3) is a linear combination of f_1^2, f_2^2, and $f_1 f_2$. Even though the stable CLT for $f_2^2(d_i^n(X)/\sqrt{\Delta_n})$ is easily deduced from Lemma B.2 (cf. Example 1 in Section 3, Duembgen and Podolskij, 2015), the result of g cannot be obtained straightforwardly because the range is not Gâteaux differentiable in general.

However, we may replace the Gâteaux derivative by an alternative form for functions which are not Gâteaux differentiable. We consider a range-based functional $\xi(x) = f(f_1(x))$ as an example, where $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function, such that both f and f' have polynomial growth (cf. Example 3 in Section 3, Duembgen and Podolskij, 2015), by applying the following lemma from Christensen and Podolskij (2007):

Lemma B.3. Given two continuous functions $x, y \in C([0,1])$, assume t^* is the only point in $[0,1]$ where the maximum of x is achieved, i.e., $t^* = \operatorname{argmax}_{0\leq s\leq 1} x(s)$. Then it holds that

$$\lim_{h\to 0} \frac{\sup_{0\leq s\leq 1}(x(s) + hy(s)) - \sup_{0\leq s\leq 1} x(s)}{h} = y(t^*). \tag{B.12}$$

In the proofs $x(t)$ plays the role of the Brownian motion, which attains its maximum (resp. minimum) at a unique point almost surely. Let $\bar{t} = \operatorname{argmax}_{0\leq s\leq 1} W_s$ and $\underline{t} = \operatorname{argmin}_{0\leq s\leq 1} W_s$. Then Lemma B.2

remains valid when σ is everywhere invertible (Christensen and Podolskij, 2012) with

$$\begin{aligned}
\rho_z^{(1)}(\xi) &= \mathbb{E}\left[f\left(z\left(\sup_{0\le s\le 1} W_s - \inf_{0\le s\le 1} W_s\right)\right) W_1\right], \\
\rho_z^{(2)}(\xi') &= \mathbb{E}\left[f'\left(z\left(\sup_{0\le s\le 1} W_s - \inf_{0\le s\le 1} W_s\right)\right) (\bar{t}-\underline{t})\right], \\
\rho_z^{(3)}(\xi') &= \mathbb{E}\left[f'\left(z\left(\sup_{0\le s\le 1} W_s - \inf_{0\le s\le 1} W_s\right)\right) (W_{\bar{t}}^2 - W_{\underline{t}}^2)\right],
\end{aligned} \tag{B.13}$$

which extends the asymptotic theory in Lemma B.2 to general functions of the range.

Moreover, the derivative of the cross term $f_1 f_2$ is a linear combination of two separate components which include f_1' and f_2' respectively. It means that for the path-dependent function $g : C([0,1]) \to \mathbb{R}$ in Eq. (B.3) we can obtain the closed-form $\rho_z^{(1)}(g)$, $\rho_z^{(2)}(g')$, and $\rho_z^{(3)}(g')$ for all $z \in \mathbb{R}$, and Eq. (B.10) holds when g is an even function. Therefore, the stable CLT in Lemma B.2 holds with the limiting process $U_t(g)$ in Eq. (B.11), where the squared integrand is given by

$$\begin{aligned}
\rho_{\sigma_\tau}(g^2) - \rho_{\sigma_\tau}^2(g) &= \mathbb{E}\left[g^2(\{\sigma_\tau W_s; s\in[0,1]\})\right] - \left(\mathbb{E}\left[g(\{\sigma_\tau W_s; s\in[0,1]\})\right]\right)^2 \\
&= \frac{1}{(4\ln 2 - 2)^2}\left\{\mathbb{E}\left[\left(f_1(\{\sigma_\tau W_s; s\in[0,1]\}) - f_2(\{\sigma_\tau W_s; s\in[0,1]\})\right)^4\right]\right. \\
&\quad \left. - \left(\mathbb{E}\left[\left(f_1(\{\sigma_\tau W_s; s\in[0,1]\}) - f_2(\{\sigma_\tau W_s; s\in[0,1]\})\right)^2\right]\right)^2\right\} \\
&= \frac{1}{(4\ln 2 - 2)^2}\left\{\mathbb{E}\left[\left(\sup_{0\le s\le 1}\sigma_\tau W_s - \inf_{0\le s\le 1}\sigma_\tau W_s - \sigma_\tau |W_1|\right)^4\right]\right. \\
&\quad \left. - \left(\mathbb{E}\left[\left(\sup_{0\le s\le 1}\sigma_\tau W_s - \inf_{0\le s\le 1}\sigma_\tau W_s - \sigma_\tau |W_1|\right)^2\right]\right)^2\right\} \\
&= \frac{\sigma_\tau^4}{(4\ln 2 - 2)^2}\left\{\mathbb{E}\left[(\omega - |c|)^4\right] - \left(\mathbb{E}\left[(\omega - |c|)^2\right]\right)^2\right\} \\
&= \frac{\sigma_\tau^4}{(4\ln 2 - 2)^2}\left\{\mathbb{E}\left[\omega^4 - 4\omega^3|c| + 2\omega^2 c^2 - 4\omega|c|^3 + c^4\right] - \left(\mathbb{E}\left[\omega^2 + c^2 - 2\omega|c|\right]\right)^2\right\} \\
&= \frac{\sigma_\tau^4}{(4\ln 2 - 2)^2}\left\{\lambda_{4,0} - 4\lambda_{3,1} + 6\lambda_{2,2} - 4\lambda_{1,3} + \lambda_{0,4} - (\lambda_{2,0} + \lambda_{0,2} - 2\lambda_{1,1})^2\right\} \\
&= \frac{40\ln 2 - 16(\ln 2)^2 - 3\zeta(3) - 16}{(4\ln 2 - 2)^2}\sigma_\tau^4 \approx 0.7245\sigma_\tau^4.
\end{aligned} \tag{B.14}$$

This completes the proof.

□

B.3 Proof of Corollary 1

The proof is analogous to that of Theorem 1. The RRDQ estimator $\widehat{Q}_{t,n}$ in Eq. (9) can be written in the form of Eq. (B.1) with a locally uniformly continuous function $g^2 : C([0,1]) \to \mathbb{R}$ of the scaled

incremental process:

$$g^2\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right) = \frac{1}{\Lambda_4}\left\{f_1\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right) - f_2\left(\frac{d_i^n(X)}{\sqrt{\Delta_n}}\right)\right\}^4. \tag{B.15}$$

Then the LLN in Lemma B.1 readily applies with

$$\begin{aligned}
\int_0^t \rho_{\sigma_\tau}(g^2)d\tau &= \int_0^t \mathbb{E}\left[g^2(\{\sigma_\tau W_s; s\in[0,1]\})\right]d\tau \\
&= \frac{1}{\Lambda_4}\int_0^t \mathbb{E}\left[(f_1(\{\sigma_\tau W_s; s\in[0,1]\}) - f_2(\{\sigma_\tau W_s; s\in[0,1]\}))^4\right]d\tau \\
&= \frac{1}{\Lambda_4}\int_0^t \mathbb{E}\left[\left(\sup_{0\le s\le 1}\sigma_\tau W_s - \inf_{0\le s\le 1}\sigma_\tau W_s - \sigma_\tau|W_1|\right)^4\right]d\tau \\
&= \frac{1}{\Lambda_4}\int_0^t \sigma_\tau^4\,\mathbb{E}\left[\omega^4 - 4\omega^3|c| + 2\omega^2c^2 - 4\omega|c|^3 + c^4\right]d\tau \\
&= \frac{1}{\Lambda_4}\int_0^t \sigma_\tau^4\left(\lambda_{4,0} - 4\lambda_{3,1} + 6\lambda_{2,2} - 4\lambda_{1,3} + \lambda_{0,4}\right)d\tau \\
&= \frac{1}{\Lambda_4}\int_0^t \sigma_\tau^4\left\{9\zeta(3) - 4\times\frac{45}{8}\zeta(3) + 6\left(4\ln 2 + \frac{7}{4}\zeta(3)\right) - 4\times\frac{15}{4} + 3\right\}d\tau \\
&= \frac{1}{\Lambda_4}\int_0^t \sigma_\tau^4\left(24\ln 2 - 12 - 3\zeta(3)\right)d\tau \\
&= \int_0^t \sigma_\tau^4 d\tau.
\end{aligned} \tag{B.16}$$

This completes the proof.

B.4 Proof of Proposition 1

Proof. We define the set

$$\Gamma_n = \{1\le i\le n : X \text{ is discontinuous in } I_{n,i}\}, \quad \text{with } k_n = |\Gamma_n|, \tag{B.17}$$

where $|A|$ stands for the cardinality of set A. The absolute summability of $\Delta X_s = X_s - X_{s^-}$ for all $s\in[0,t]$ implies that the number of "visible" realizations of the discontinuous component, i.e., ΔX_s of a larger order of magnitude than $\sqrt{\Delta_n}$, is an $O_{\mathbb{P}}(\Delta_n^\gamma)$ random variable, where

$$\gamma = -\sup_{0\le s\le t}\left\{0\le\varpi<\frac{1}{2} : \Delta X_s \asymp \Delta_n^\varpi\right\}, \tag{B.18}$$

which corresponds to the smallest order of non-negligible jumps on $[0,t]$. It is obvious that k_n is bounded by the number of non-negligible jumps over $[0,t]$, such that we have $k_n = O_{\mathbb{P}}(\Delta_n^\gamma)$.

We decompose the RRDV over $[0,t]$ into two complementary parts:

$$\widehat{V}_{t,n} = \frac{1}{\Lambda_2}\sum_{i\in\Gamma_n}(w_i-|r_i|)^2 + \frac{1}{\Lambda_2}\sum_{i\in\Gamma'_n}(w_i-|r_i|)^2 = \widehat{V}^{(1)}_{t,k_n} + \widehat{V}^{(2)}_{t,n-k_n}. \tag{B.19}$$

For $\widehat{V}^{(1)}_{t,k_n}$, we have

$$w_i = \sup_{\tau,\tau'\in I_{n,i}}\left|\int_\tau^{\tau'}\sigma_s dW_s + \sum_{s\in I_{n,i}}\Delta X_s\right| \le \sup_{\tau,\tau'\in I_{n,i}}\left|\int_\tau^{\tau'}\sigma_s dW_s\right| + \left|\sum_{s\in I_{n,i}}\Delta X_s\right|, \tag{B.20}$$

$$|r_i| = \left|\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s dW_s + \sum_{s\in I_{n,i}}\Delta X_s\right| \ge \left|\sum_{s\in I_{n,i}}\Delta X_s\right| - \left|\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s dW_s\right|, \tag{B.21}$$

and thus

$$\begin{aligned} w_i - |r_i| &\le \sup_{\tau,\tau'\in I_{n,i}}\left|\int_\tau^{\tau'}\sigma_s dW_s\right| + \left|\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s dW_s\right| \\ &= \sigma_{(i-1)\Delta_n}\sqrt{\Delta_n}\left(\sup_{\tau,\tau'\in[0,1]}|W_\tau - W_{\tau'}| + |W_1|\right) + o_{\mathbb{P}}\left(\sqrt{\Delta_n}\right) \quad \text{(Euler discretization)} \\ &= O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right). \end{aligned} \tag{B.22}$$

By adding up the squares of range-return differences in all k_n intervals, we have

$$\widehat{V}^{(1)}_{t,k_n} = O_{\mathbb{P}}\left(k_n\Delta_n\right) = O_{\mathbb{P}}\left(\Delta_n^{\gamma+1}\right). \tag{B.23}$$

For the sum of IV over all k_n intervals, we also have

$$\sum_{i\in\Gamma_n}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s^2 ds = O_{\mathbb{P}}\left(k_n\Delta_n\right) = O_{\mathbb{P}}\left(\Delta_n^{\gamma+1}\right). \tag{B.24}$$

With the triangle inequality, the absolute bias satisfies

$$\left|\widehat{V}^{(1)}_{t,k_n} - \sum_{i\in\Gamma_n}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s^2 ds\right| \le \widehat{V}^{(1)}_{t,k_n} + \sum_{i\in\Gamma_n}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s^2 ds = O_{\mathbb{P}}\left(\Delta_n^{\gamma+1}\right). \tag{B.25}$$

For $\widehat{V}^{(2)}_{t,n-k_n}$, it holds naturally that

$$\widehat{V}^{(2)}_{t,n-k_n} - \sum_{i\in\Gamma'_n}\int_{(i-1)\Delta_n}^{i\Delta_n}\sigma_s^2 ds = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right). \tag{B.26}$$

The results for $\widehat{V}^{(1)}_{t,k_n}$ and $\widehat{V}^{(2)}_{t,n-k_n}$ in Eqs. (B.25) and (B.26) imply the result in the theorem.

□

B.5 Proof of Proposition 2

Proof. We follow Andersen et al. (2021) to assume $\tau = 0$ for simplicity:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + H_t, \quad \text{where } H_t = \int_0^t \frac{c_s^+}{s^\alpha} ds, \quad \frac{1}{2} < \alpha < 1. \tag{B.27}$$

The increment of H_t over the i-th interval is given by

$$H_{i\Delta_n} - H_{(i-1)\Delta_n} = \int_{(i-1)\Delta_n}^{i\Delta_n} \frac{c_s^+}{s^\alpha} ds = C_i \Delta_n^{1-\alpha} f(i;\alpha), \tag{B.28}$$

for some constant C_i, where $f(x;\theta) = x^{1-\theta} - (x-1)^{1-\theta}$ is a monotonically decreasing function over $[1,\infty)$ with $f(1;\theta) = 1$ and $\lim_{x\to\infty} f(x;\theta) = 0$ for all $0 < \theta < 1$. There exists an unique integer K_n defined as

$$K_n = \max_i \left\{ i \in \mathbb{Z}^+, 1 \le i \le n : f(i;\alpha) \asymp \Delta_n^{\alpha - \frac{1}{2}} \right\}. \tag{B.29}$$

The mean value theorem indicates

$$f(K_n;\alpha) = K_n^{1-\alpha} - (K_n - 1)^{1-\alpha} = (1-\alpha)(K_n - \varepsilon)^{-\alpha}, \tag{B.30}$$

for some $\varepsilon \in (0,1)$. It is therefore satisfied that

$$K_n \asymp \Delta_n^{\frac{1}{2\alpha} - 1}. \tag{B.31}$$

The role of H is no smaller than the diffusion component over the first K_n intervals, while it starts to be swamped by volatility from the (K_n+1)-th interval because its contribution vanishes in the limit. Depending on the asymptotic order of H, we decompose the RRDV over $[0,t]$ into two complementary parts:

$$\widehat{V}_{t,n} = \frac{1}{\Lambda_2} \sum_{i=1}^{K_n} (w_i - |r_i|)^2 + \frac{1}{\Lambda_2} \sum_{i=K_n+1}^{n} (w_i - |r_i|)^2 = \widehat{V}_{[0,K_n\Delta_n]} + \widehat{V}_{[K_n\Delta_n,t]}. \tag{B.32}$$

For all $1 \le i \le K_n$, we have

$$w_i = \sup_{\tau,\tau' \in I_{n,i}} \left| \int_\tau^{\tau'} \sigma_s dW_s + H_{\tau'} - H_\tau \right| \le \sup_{\tau,\tau' \in I_{n,i}} \left| \int_\tau^{\tau'} \sigma_s dW_s \right| + \left| H_{i\Delta_n} - H_{(i-1)\Delta_n} \right|, \tag{B.33}$$

$$|r_i| = \left| \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s + H_{i\Delta_n} - H_{(i-1)\Delta_n} \right| \ge \left| H_{i\Delta_n} - H_{(i-1)\Delta_n} \right| - \left| \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s \right|, \tag{B.34}$$

and thus

$$w_i - |r_i| \leq \sup_{\tau,\tau' \in I_{n,i}} \left| \int_{\tau}^{\tau'} \sigma_s dW_s \right| + \left| \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s \right| = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right). \tag{B.35}$$

By adding up all squared range-return differences in the first K_n intervals, we have

$$\widehat{V}_{[0,K_n\Delta_n]} = O_{\mathbb{P}}\left(K_n\Delta_n\right) = O_{\mathbb{P}}\left(\Delta_n^{\frac{1}{2\alpha}}\right). \tag{B.36}$$

For the IV over the period which accommodates the first K_n intervals, it holds that

$$\int_0^{K_n\Delta_n} \sigma_s^2 ds = O_{\mathbb{P}}\left(K_n\Delta_n\right) = O_{\mathbb{P}}\left(\Delta_n^{\frac{1}{2\alpha}}\right). \tag{B.37}$$

With the triangle inequality, the absolute bias satisfies

$$\left| \widehat{V}_{[0,K_n\Delta_n]} - \int_0^{K_n\Delta_n} \sigma_s^2 ds \right| \leq \widehat{V}_{[0,K_n\Delta_n]} + \int_0^{K_n\Delta_n} \sigma_s^2 ds = O_{\mathbb{P}}\left(\Delta_n^{\frac{1}{2\alpha}}\right). \tag{B.38}$$

For $\widehat{V}_{[K_n\Delta_n,t]}$, it holds naturally that

$$\widehat{V}_{[K_n\Delta_n,t]} - \int_{K_n\Delta_n}^{t} \sigma_s^2 ds = O_{\mathbb{P}}\left(\sqrt{\Delta_n}\right). \tag{B.39}$$

Since $K_n\Delta_n \to 0$ under infill asymptotics, the RRDV over $[0,t]$ is equivalent to $\widehat{V}_{[K_n\Delta_n,t]}$. The bias results in Eqs. (B.38) and (B.39) show that the bias of RRDV due to drift burst is asymptotically negligible and has no impact on the asymtotic distribution in Theorem 2. This completes the proof. □

B.6 Proof of Proposition 3

Proof. As an analogous result to Proposition 2, the bias of RRDV in the presence of persistent noise can be proved following the same steps with similar simplifying assumptions: There exists one persistent noise episode $[0,1]$, which is triggered by some ambiguous information arriving at time 0, and the funtion $g^{(1)}$ takes the form $g_{gj}^{(1)}$ in Eq. (22). The process $\epsilon_t^{(1)}$ in H_t only introduces extra randomness to the duration of persistent noise episode, which shall be harmlessly ignored.

As shown in Eqs. (56) and (57) in Andersen et al. (2021), there exists an asymptotic correspondence between the two models of episodic extreme return persistence, and they are equivalent with identical asymptotic analyses if we let $\beta = 1-\alpha$. The increment of H_t on the i-th interval is

$$H_{i\Delta_n} - H_{(i-1)\Delta_n} = f^{(1)}(\Delta X_0, \eta)\left[(i-1)^{\beta} - i^{\beta}\right]\Delta_n^{\beta} = \eta\Delta X_0\left[i^{\beta} - (i-1)^{\beta}\right]\Delta_n^{\beta}, \tag{B.40}$$

where $f^{(1)}(\Delta X_0, \eta) = -\eta\Delta X_0$ with $\eta \in (0,1]$. With $\beta = 1-\alpha \in (0,1/2)$, the above persistent noise increment is equivalent to the drift burst increment in Eq. (B.28). Jumps with $r=0$ induce the bias of order $O_{\mathbb{P}}(\Delta_n^{\gamma+1}) = O_{\mathbb{P}}(\Delta_n)$ with $\gamma = 0$, which has no impact on the bias result of RRDV. The

proof from here can proceed following the same steps as the proof of Theorem 5. When the function $g^{(1)} = g^{(1)}_{fc}$, the asymptotic effect of H depends on the smaller of the two parameters β^- and β^+. □

B.7 Proof of Proposition 4

Proof. We obtain the asymptotic expansions of $\Lambda_{2,N}$ and $\Lambda_{4,N}$ by specializing the general results in Asmussen et al. (1995) and Dieker and Lagos (2017). Lemma B.4 demonstrates the asymptotic distribution of the Euler discretization error of one-dimensional reflected Brownian motion.

Lemma B.4. We denote a reflected Brownian motion by $\overline{W} = \Gamma W$ with the reflection mapping

$$\Gamma X_t = X_t - \left(\inf_{0 \le s \le t} X_s \wedge 0 \right), \tag{B.41}$$

where W is a standard Brownian motion that starts from 0. Let $\overline{W}_{t,N}$ be the embedded reflected Brownian motion observed at N discrete points, i.e., at $t_i = it/N$ for $i = 1, 2, \ldots, N$. The Euler discretization error of $\overline{W}_1$, i.e., $\varepsilon_{1,N} = \overline{W}_1 - \overline{W}_{1,N}$, has a weak convergence to a nonzero limit:

$$\sqrt{N}\varepsilon_{1,N} \xrightarrow{\mathcal{L}} \Upsilon, \quad \text{with } \Upsilon = \min_{k \in \mathbb{Z}} R_{U+k}, \tag{B.42}$$

as $N \to \infty$, where $R = (R_t)_{t \ge 0}$ is a two-sided Bessel process of order 3, U is a uniformly distributed random variable on $(0, 1)$ which is independent of R. The scaled discretization error $\sqrt{N}\varepsilon_{1,N}$ is asymptotically independent of W, with the $\mathbb{R} \times C([0,1])$-valued random pair $(\sqrt{N}\varepsilon_{1,N}, W) \xrightarrow{\mathcal{L}} (\Upsilon, W)$, where Υ is independent of W.

Lemma B.5. The Euler discretization error of $\overline{W}$ satisfies

$$\varepsilon_{1,N} = \inf_{i \in \{0,1,\ldots,N\}} W_{i/N} - \inf_{0 \le t \le 1} W_t. \tag{B.43}$$

Given a function $g : \mathbb{R} \to \mathbb{R}$ whose first derivative g' exists at $\inf_{0 \le t \le 1} W_t$ and is non-zero valued, the delta method implies that as $N \to \infty$,

$$\sqrt{N}\left(g\left(\inf_{i \in \{0,1,\ldots,N\}} W_{i/N} \right) - g\left(\inf_{0 \le t \le 1} W_t \right) \right) \xrightarrow{\mathcal{L}} g'\left(\inf_{0 \le t \le 1} W_t \right) \Upsilon. \tag{B.44}$$

Therefore, the expected functional values of discretized infimum can be approximated with the polynomial expansion as follows:

$$\mathbb{E}\left[g\left(\inf_{i \in \{0,1,\ldots,N\}} W_{i/N} \right) \right] = \mathbb{E}\left[g\left(\inf_{0 \le t \le 1} W_t \right) \right] + \mathbb{E}\left[g'\left(\inf_{0 \le t \le 1} W_t \right) \right] \mathbb{E}[\Upsilon] \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}} \right), \tag{B.45}$$

where

$$\mathbb{E}[\Upsilon] = -\frac{\zeta(1/2)}{\sqrt{2\pi}} \approx 0.5826. \tag{B.46}$$

The results above enable us to derive asymptotic expansions of the moments of $(\omega, |c|)$, i.e., $\lambda_{p,r} = \mathbb{E}[\omega^p |c|^r]$, whose analytical values are summarized in Table A.1, with $p, r \in \mathbb{N}$ and $0 \leq p+r \leq 4$.

Corollary B.1. For the moments of $(\omega, |c|)$ derived in Appendix A, i.e., $\lambda_{p,r} = \mathbb{E}[\omega^p |c|^r]$, we have the following asymptotic result:

$$\lambda_{p,r,N} = \lambda_{p,r} + M_{p,r} \frac{\zeta(1/2)}{\sqrt{2\pi}} \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right), \quad \text{with } M_{p,r} = 2p\lambda_{p-1,r}. \tag{B.47}$$

Proof. It is intuitively clear that the random variable $\varepsilon_{1,N}$ in Lemma B.4, for N large, is solely determined by the behavior of W in a neighborhood of its minimizer $\underline{t}$, i.e., the almost surely unique random time $\underline{t} \in [0,1]$ at which W attains its minimum value $\inf_{0\leq t\leq 1} W_t$ over the unit interval.

The results in Lemma B.5 are also convenient to switch from infima to suprema, with

$$\varepsilon_{1,N} = \sup_{0\leq t\leq 1} W_t - \sup_{i\in\{0,1,\ldots,N\}} W_{i/N}, \tag{B.48}$$

which is a direct result from sign reversion.

Because the Brownian motion is space-homogeneous and symmetric, it holds that

$$\begin{aligned} \omega_N - \omega &= \left(\sup_{i\in\{0,1,\ldots,N\}} W_{i/N} - \inf_{i\in\{0,1,\ldots,N\}} W_{i/N}\right) - \left(\sup_{0\leq t\leq 1} W_t - \inf_{0\leq t\leq 1} W_t\right) \\ &= \left(\sup_{i\in\{0,1,\ldots,N\}} W_{i/N} - \sup_{0\leq t\leq 1} W_t\right) - \left(\inf_{i\in\{0,1,\ldots,N\}} W_{i/N} - \inf_{0\leq t\leq 1} W_t\right) \\ &\overset{\mathcal{L}}{=} -2\varepsilon_{1,N}, \end{aligned} \tag{B.49}$$

and therefore

$$\sqrt{N}\left(\omega_N^p - \omega^p\right) \xrightarrow{\mathcal{L}} -2p\omega^{p-1}\Upsilon, \quad \text{as } N \to \infty, \tag{B.50}$$

by the delta method. The equivalence in distribution "$\overset{\mathcal{L}}{=}$" in Eq. (B.49) holds because $\sup_{0\leq t\leq 1} W_t - \sup_{i\in\{0,1,\ldots,N\}} W_{i/N}$ and $\inf_{i\in\{0,1,\ldots,N\}} W_{i/N} - \inf_{0\leq t\leq 1} W_t$ are asymptotically i.i.d..

For the Euler discretization error of the moment $\lambda_{p,r}$, we have

$$\begin{aligned} \lambda_{p,r,N} - \lambda_{p,r} &= \mathbb{E}[(\omega_N^p - \omega^p)|c|^r] \\ &= -2p\mathbb{E}[\omega^{p-1}|c|^r]\mathbb{E}[\Upsilon]\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right) \\ &= 2p\lambda_{p-1,r}\frac{\zeta(1/2)}{\sqrt{2\pi}}\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right). \end{aligned} \tag{B.51}$$

This completes the proof of Corollary B.1.

□

Table B.1 lists the values of "bias factor" $M_{p,r}$ for all $1 \le p \le 4$ and $0 \le r \le 3$.

Table B.1: Bias factor $M_{p,r}$ for discrete moment $\lambda_{p,r,N}$

p \ r	0	1	2	3
1	2	$2\sqrt{\frac{2}{\pi}}$	2	$4\sqrt{\frac{2}{\pi}}$
2	$8\sqrt{\frac{2}{\pi}}$	6	$\frac{32}{3}\sqrt{\frac{2}{\pi}}$	–
3	$24\ln 2$	$\frac{14}{3}\sqrt{\frac{\pi^3}{2}}$	–	–
4	$\frac{16}{3}\sqrt{2\pi^3}$	–	–	–

This table lists the "bias factor" $M_{p,r}$ used in asymptotic expansions of $\lambda_{p,r,N}$ with $1 \le p \le 4$ and $0 \le r \le 3$. In the polynomial expansion for $\lambda_{p,r,N}$, the coefficient for $N^{-1/2}$ is $M_{p,r}\zeta(1/2)/\sqrt{2\pi}$.

It is now straightforward to obtain the asymptotic expansions of $\Lambda_{2,N}$ and $\Lambda_{4,N}$ in Proposition 4:

$$\begin{aligned}
\Lambda_{2,N} &= \lambda_{2,0,N} + \lambda_{0,2} - 2\lambda_{1,1,N} \\
&= \Lambda_2 + (M_{2,0} - 2M_{1,1})\frac{\zeta(1/2)}{\sqrt{2\pi}}\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right) \\
&= \Lambda_2 + 4\sqrt{\frac{2}{\pi}}\frac{\zeta(1/2)}{\sqrt{2\pi}}\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right) \\
&= \Lambda_2 + \frac{4}{\pi}\zeta\left(\frac{1}{2}\right)\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right),
\end{aligned} \tag{B.52}$$

$$\begin{aligned}
\Lambda_{4,N} &= \lambda_{4,0,N} - 4\lambda_{3,1,N} + 6\lambda_{2,2,N} - 4\lambda_{1,3,N} + \lambda_{0,4} \\
&= \Lambda_4 + (M_{4,0} - 4M_{3,1} + 6M_{2,2} - 4M_{1,3})\frac{\zeta(1/2)}{\sqrt{2\pi}}\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right) \\
&= \Lambda_4 + \left(\frac{16}{3}\sqrt{2\pi^3} - \frac{56}{3}\sqrt{\frac{\pi^3}{2}} + 48\sqrt{\frac{2}{\pi}}\right)\frac{\zeta(1/2)}{\sqrt{2\pi}}\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right) \\
&= \Lambda_4 + \left(\frac{48}{\pi} - 4\pi\right)\zeta\left(\frac{1}{2}\right)\frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).
\end{aligned} \tag{B.53}$$

This completes the proof.

□

B.8 Proof of Corollary 2

Proof. It follows the same steps as the proofs of Theorem 1, 2, and 3. □

Appendix C Supplementary Materials

C.1 Simulation Scheme

The numerical results for $(\Lambda_{2,N}, \Lambda_{4,N}, \Theta_N)'$ with different N are calculated from a large number of simulated paths of the standard Brownian motion. Each replication generates a sequence of $N+1$ equidistant observations at $t_i = i/N$ for $i = 0, 1, \ldots, N$. As N is allowed to span any natural number except 0 and 1, we adopt the following simulation scheme to efficiently utilize our computational resources:

i. For $N \in \{2, 3, 4, \ldots, 10\}$, we simulate 10^9 replications of $W_{i/N}$ for $i = 0, 1, \ldots, N$.
ii. For $N \in \{11, 12, 13, \ldots, 200\}$, we simulate 10^8 replications of $W_{i/N}$ for $i = 0, 1, \ldots, N$.
iii. For $N \in \{201, 202, 203, \ldots, 2000, 2005, 2010, \ldots, 5000, 5010, 5020, \ldots, 10^4, 10^5, 10^6, 10^7\}$, we simulate 10^7 replications of $W_{i/N}$ for $i = 0, 1, \ldots, N$.

C.2 Discretized Factors

Table C.1 reports the estimation results for the polynomial equation

$$Y_N = \sum_{i=0}^{k} \beta_i N^{-i/2} + \epsilon_N, \tag{C.1}$$

where $Y_N = (\Lambda_{2,N}, \Lambda_{4,N}, \Theta_N)'$ collects the simulated values for all three factors with N ranging from 11 to 10^8. The intercepts $\beta_0 = (\Lambda_2, \Lambda_4, \Theta)'$ and the first two coefficients in the vector β_1 are obtained from the analytical results in Proposition 4. We find that a cubic (resp. quartic) approximation works very well across all $N \geq 11$ for $\Lambda_{2,N}$ (resp. $\Lambda_{4,N}$ or Θ_N), as indicated by the root-mean-square error (RMSE) and R^2.

Table C.1: Polynomial regression results for discrete factors

Coefficients	$\Lambda_{2,N}$	$\Lambda_{4,N}$	Θ_N
β_0	Λ_2	Λ_4	Θ
β_1	$4\zeta(1/2)/\pi$	$(48/\pi - 4\pi)\zeta(1/2)$	1.6618
β_2	1.7429	6.8076	1.7371
β_3	-0.6999	-6.3635	1.0395
β_4	–	2.8711	5.4477
RMSE $\times 10^{-4}$	0.6088	1.6358	1.3555
R^2	0.9998	0.9999	1.0000

This table reports the estimated coefficients for the polynomial regression model $Y_N = \sum_{i=0}^{k} \beta_i N^{-i/2} + \epsilon_N$ with weighted least squares. $Y_N = (\Lambda_{2,N}, \Lambda_{4,N}, \Theta_N)'$ collects the simulated values for all three factors with N ranging from 11 to 10^8. The intercepts β_0 and the first two coefficients in the vector β_1 are from the analytical results in Proposition 4.

We next provide a practical instruction on the selection for all factors with different $N \in \mathbb{N}^{>1}$:

- Use the simulated values in Table C.2 for $N \in \{2, 3, \ldots, 10\}$.
- Use the polynomial approximation with coefficients listed in Table C.1 for all $N \geq 11$.

Table C.2 shows the simulated values for $(\Lambda_{2,N}, \Lambda_{4,N}, \Theta_N)'$ with N ranging from 2 to 10, which achieve the highest level of precision within our simulation schemes in Appendix C.1.

Table C.2: Simulated values for discrete factors

N	$\Lambda_{2,N}$	$\Lambda_{4,N}$	Θ_N
2	0.0908	0.0567	5.8696
3	0.1486	0.0945	3.2809
4	0.1926	0.1304	2.5170
5	0.2277	0.1631	2.1457
6	0.2567	0.1926	1.9224
7	0.2812	0.2192	1.7712
8	0.3023	0.2432	1.6616
9	0.3206	0.2650	1.5777
10	0.3368	0.2849	1.5110

Simulated values for $\Lambda_{2,N}$, $\Lambda_{4,N}$, and Θ_N with N ranging from 2 to 10. For the detailed simulation scheme, see Appendix C.1.

C.3 Monte Carlo Bias Results of Other Estimators

In addition to the Monte Carlo bias results in Section 5.2, Table C.6 reports the relative bias (%) in "continuous time" of the TRV estimator of Mancini (2009) and the DV estimator of Andersen et al. (2021). The choices of truncation parameters for TRV and DV are in line with Section 5.3.

C.4 Monte Carlo RMSE Results of Other Estimators

In addition to the comparison of finite-sample performances among RRDV and the main competitors TRV and DV, we also consider two traditional IV estimators, i.e., RV and RBPV, and also TRV and DV with less aggressive choices of truncation threshold. The RMSE results are shown in Table C.6.

Table C.3: Monte Carlo bias results (%): Truncated realized volatility (TRV)

Panel A: $C_\zeta^{\text{TRV}} = 4$										
		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-0.14	0.51	0.56	0.56	0.85	0.89	0.94	0.58	0.72	0.83
5	-0.18	2.12	2.02	1.80	3.46	3.19	2.88	2.55	2.72	2.60
10	-0.16	3.66	3.36	2.83	6.05	5.39	4.74	4.56	4.58	4.23
30	-0.19	8.83	7.63	5.92	15.07	12.58	9.90	9.96	10.08	9.36
60	-0.27	15.09	12.55	9.14	24.87	19.93	15.29	17.21	15.29	15.83
120	-0.35	27.92	22.02	16.02	42.97	35.15	25.04	34.72	29.33	21.38
180	-0.28	32.40	25.29	16.85	61.86	49.50	35.73	30.59	29.20	21.88
300	-0.39	43.35	31.71	19.63	96.99	78.29	56.19	42.98	31.45	28.53
Panel B: $C_\zeta^{\text{TRV}} = 3$										
		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-3.06	-2.45	-2.35	-2.25	-2.19	-2.15	-2.05	-2.44	-2.27	-2.10
5	-3.09	-1.02	-1.05	-1.08	0.05	0.03	0.07	-0.73	-0.55	-0.59
10	-3.08	0.29	0.16	-0.17	2.30	1.69	1.27	0.93	0.95	0.75
30	-3.12	4.43	3.56	2.19	9.24	7.24	5.54	5.56	5.27	4.60
60	-3.20	9.36	7.02	4.32	17.76	14.74	11.22	11.10	9.64	8.89
120	-3.29	17.76	13.51	9.09	29.81	22.18	14.27	21.30	19.47	13.72
180	-3.37	22.29	16.48	10.47	44.94	34.00	23.32	23.20	19.37	16.88
300	-3.68	32.41	23.52	14.47	72.39	54.66	37.25	26.24	23.27	24.39

Relative bias (%) of the truncated realized volatility (TRV) estimator of Mancini (2009) constructed from 1, 5, 10, 30, 60, 120, 180, and 300-second intervals for 2000 days. The truncation threshold for returns in all intervals is $C_\zeta^{\text{TRV}}\sqrt{\Delta_n \text{MedRV}_{t,n}}$, with $C_\zeta^{\text{TRV}} = 4$ or 3. The DGP is the Heston model in Eq. (38), and we follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence.

Table C.4: Monte Carlo bias results (%): Differenced-return volatility (DV)

Panel A: $C_\zeta^{\text{DV}} = 4\sqrt{2}$										
		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-0.15	-0.10	-0.09	-0.08	-0.04	-0.09	-0.04	-0.15	-0.11	-0.01
5	-0.22	-0.08	-0.14	-0.01	-0.01	0.09	0.15	0.12	0.16	0.11
10	-0.18	-0.24	-0.23	-0.09	0.53	0.35	0.67	0.46	0.16	0.13
30	-0.24	0.08	-0.23	-0.04	0.37	0.58	1.26	1.65	1.03	0.87
60	-0.39	1.30	-0.41	-0.59	1.59	2.89	4.62	4.60	1.55	1.86
120	-0.50	6.58	6.34	0.76	9.70	9.81	8.61	15.33	12.94	2.99
180	-0.47	8.80	1.21	-1.02	14.43	11.35	9.74	16.16	11.74	2.16
300	-1.02	23.35	8.80	-0.33	30.23	23.24	19.18	40.15	17.62	8.58
Panel B: $C_\zeta^{\text{DV}} = 3\sqrt{2}$										
		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-3.05	-2.88	-2.81	-2.65	-2.86	-2.83	-2.67	-2.95	-2.90	-2.64
5	-3.12	-2.63	-2.48	-2.22	-2.37	-2.10	-1.75	-2.62	-2.45	-2.21
10	-3.10	-2.47	-2.28	-2.07	-1.76	-1.77	-1.52	-2.14	-1.97	-2.02
30	-3.14	-1.88	-1.69	-1.59	-0.82	-0.70	-0.27	-1.11	-1.26	-1.33
60	-3.32	-1.17	-1.11	-1.45	0.83	1.58	2.17	-0.08	-0.80	-0.72
120	-3.53	1.68	1.79	-0.24	5.51	5.10	3.92	5.21	4.33	1.07
180	-3.56	0.81	-1.65	-2.28	4.49	2.85	2.19	5.49	3.10	0.41
300	-4.15	6.29	-0.81	-2.87	7.30	5.31	4.77	14.24	8.92	3.81

Relative bias (%) of the differenced-return volatility (DV) estimator of Andersen et al. (2021) constructed from 1, 5, 10, 30, 60, 120, 180, and 300-second intervals for 2000 days. The truncation threshold for all first-order differenced returns is $C_\zeta^{\text{DV}}\sqrt{\Delta_n \text{MedRV}_{t,n}}$, with $C_\zeta^{\text{DV}} = 4\sqrt{2}$ or $3\sqrt{2}$. The DGP is the Heston model in Eq. (38), and we follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence.

Table C.5: Monte Carlo bias results (%): Generalized differenced-return volatility (DV_{1-3})

Panel A: $C_\zeta^{DV} = 4\sqrt{2}$										
		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-0.15	-0.07	-0.08	-0.06	-0.01	-0.03	0.03	-0.07	-0.05	0.02
5	-0.23	-0.02	-0.03	0.09	0.09	0.12	0.25	0.29	0.37	0.32
10	-0.23	0.12	-0.11	0.23	0.54	0.52	0.76	0.65	0.72	0.81
30	-0.39	0.68	0.71	0.90	1.40	1.42	1.93	2.58	2.63	2.81
60	-0.67	2.21	1.34	1.16	4.32	5.35	6.48	6.65	4.03	5.95
120	-1.18	14.01	9.98	5.26	12.02	10.99	9.50	17.83	19.46	8.14
180	-1.44	12.72	4.62	1.89	24.31	21.09	17.93	26.73	16.28	7.08
300	-2.27	31.54	11.95	3.89	57.17	48.30	38.20	52.08	30.39	14.81
Panel B: $C_\zeta^{DV} = 3\sqrt{2}$										
		Gradual Jump			Flash Crash			Gradual Jump with an Intermittent Flash Crash		
Interval (sec)	$H = 0$	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
1	-3.06	-2.89	-2.80	-2.64	-2.84	-2.80	-2.66	-2.90	-2.82	-2.64
5	-3.13	-2.60	-2.48	-2.22	-2.33	-2.12	-1.74	-2.48	-2.25	-2.17
10	-3.14	-2.39	-2.15	-1.98	-1.77	-1.79	-1.60	-2.03	-1.88	-1.74
30	-3.31	-1.51	-1.33	-1.30	-0.67	-0.53	-0.21	-0.39	-0.29	-0.31
60	-3.62	-0.46	-0.61	-0.94	1.45	1.75	1.74	1.88	1.32	1.81
120	-4.14	5.46	3.46	1.92	5.53	4.83	3.53	9.79	10.25	4.78
180	-4.45	3.92	1.89	-0.04	9.35	7.50	6.52	12.95	9.92	4.50
300	-5.53	11.92	4.57	0.87	24.49	20.74	17.37	27.81	10.57	10.49

Relative bias (%) of the generalized DV_{1-3} estimator of Andersen et al. (2021) constructed from 1, 5, 10, 30, 60, 120, 180, and 300-second intervals for 2000 days. The truncation threshold for all first-, second-, and third-order differenced returns is $C_\zeta^{DV}\sqrt{\Delta_n \text{MedRV}_{t,n}}$, with $C_\zeta^{DV} = 4\sqrt{2}$ or $3\sqrt{2}$. The DGP is the Heston model in Eq. (38), and we follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence.

Table C.6: Monte Carlo RMSE results of other estimators

Panel A: $H = 0$					
Interval	RV	RBPV	TRV	DV	DV_{1-3}
1 min	6.59	0.92	0.64	0.79	0.69
2 min	6.65	1.30	0.91	1.11	0.98
3 min	6.66	1.63	1.12	1.39	1.21
5 min	6.73	2.10	1.47	1.81	1.61

Panel B: Gradual Jump	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RV	RBPV	TRV	DV	DV_{1-3}	RV	RBPV	TRV	DV	DV_{1-3}	RV	RBPV	TRV	DV	DV_{1-3}
1 min	8.32	3.85	1.54	0.83	0.76	10.53	4.51	1.35	0.80	0.72	14.81	5.24	1.11	0.80	0.72
2 min	9.48	6.81	2.81	1.45	1.73	11.67	7.68	2.32	1.36	1.47	15.82	8.62	1.85	1.18	1.17
3 min	13.08	8.79	3.23	1.87	1.95	16.70	9.14	2.65	1.59	1.44	22.62	9.13	1.99	1.41	1.29
5 min	17.61	12.80	4.35	3.41	3.87	21.81	12.55	3.39	2.66	2.55	28.07	11.86	2.46	2.09	1.99

Panel C: Flash Crash	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RV	RBPV	TRV	DV	DV_{1-3}	RV	RBPV	TRV	DV	DV_{1-3}	RV	RBPV	TRV	DV	DV_{1-3}
1 min	8.18	5.16	2.37	0.80	0.82	8.49	5.72	1.95	0.85	0.88	8.42	5.76	1.55	0.91	0.94
2 min	8.63	5.51	4.05	1.48	1.51	8.30	4.77	3.39	1.50	1.48	7.58	3.67	2.60	1.45	1.40
3 min	9.94	8.42	5.75	2.10	2.62	9.36	7.30	4.78	2.00	2.43	8.30	5.61	3.65	1.88	2.19
5 min	12.64	13.28	9.01	3.84	5.62	11.49	11.17	7.46	3.44	4.99	9.65	8.28	5.58	3.04	4.12

Panel D: Gradual Jump with an Intermittent Flash Crash	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RV	RBPV	TRV	DV	DV_{1-3}	RV	RBPV	TRV	DV	DV_{1-3}	RV	RBPV	TRV	DV	DV_{1-3}
1 min	7.98	3.28	1.73	0.90	0.90	9.21	3.89	1.51	0.82	0.81	13.06	4.83	1.40	0.84	0.85
2 min	9.07	5.89	3.45	1.65	2.09	10.13	6.41	2.56	1.50	1.87	14.27	8.08	2.57	1.53	1.72
3 min	11.22	7.30	3.51	2.24	2.72	13.74	7.83	2.91	2.04	2.04	19.83	9.09	2.82	1.85	2.07
5 min	15.40	12.12	5.94	5.60	6.93	17.13	9.23	3.34	3.06	2.76	23.22	10.35	2.65	2.25	2.36

RMSE (multiplied by 10^5) of different IV estimators: RV, RBPV, and also TRV and DV with less aggressive truncation thresholds. The choice of truncation parameters for TRV and DV follows the instructions in Section 4.2, with $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}, \varpi) = (4, 4\sqrt{2}, 1/2)$. The DGP is the Heston model in Eq. (38), and we follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence.

C.5 Monte Carlo RMSE Results with Market Microstructure Noise

In order to examine the impact of market microstructure noise on the finite-sample performance of RRDV, we augment the Heston model in Eq. (38) with an additive heterogeneous Gaussian noise term, which is in line with the Monte Carlo simulation in Christensen et al. (2022):

$$Y_i = X_i + \epsilon_i, \quad \epsilon_i \sim \text{i.i.d.}\, \mathcal{N}(0, \tilde{\sigma}_i^2), \quad \text{where } \tilde{\sigma}_i = \gamma \sqrt{\frac{\sigma_{t_i}^2}{n}}. \tag{C.2}$$

We set the noise-to-volatility ratio $\gamma = 0.5$, which corresponds to a medium contamination level (Christensen et al., 2014).[15]

Besides the additive noise, we consider the rounding errors on the price level, i.e., let the observed prices $e^{Y_i} = e^{X_i + \epsilon_i}$ be further rounded to cents. The observed logarithmic prices are given as

$$Y_i = \ln\left(\left[\frac{e^{X_i+\epsilon_i}}{0.01}\right] \times 0.01\right), \tag{C.3}$$

where the function $[x]$ rounds a number x to the nearest integer.

Table C.7 reports the RMSEs of all selected IV estimators when there exists the heterogeneous Gaussian noise in Eq. (C.2). For RRDV based on candlestick information obtained from one-second data, it exhibits noticeably elevated RMSEs compared to the noise-free case when the candlestick window is small (1 minute), and performs worse than its competitors. This observation confirms that the range-based estimators are comparatively more susceptible to noise contamination. However, RRDV regains its superiority as the interval extends slightly to 2 minutes. For the "sparse" RRDV* based on HLOCs from half-minute data, the RMSE results do not show significant differences between the noise-free and noisy cases. As discussed in Section 4.1, the implementation of effective discretization error correction facilitates the construction of RRDV on sparsely sampled observations, and thus enhance its robustness to market microstructure noise.

[15] A larger noise-to-volatility ratio, e.g., $\gamma = 1$, will not change the qualitative results.

Table C.7: Monte Carlo RMSE results with market microstructure noise

Panel A: $H = 0$

Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	1.29	1.18	0.71	0.81	**0.73**
2 min	**0.98**	1.05	0.99	1.14	1.02
3 min	**0.92**	1.11	1.21	1.38	1.24
5 min	**0.96**	1.24	1.55	1.77	1.58

Panel B: Gradual Jump

	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	1.07	1.12	1.17	0.80	**0.70**	1.13	1.12	1.02	0.80	**0.71**	1.15	1.14	0.87	0.80	**0.71**
2 min	**0.80**	1.05	1.98	1.17	1.18	**0.83**	1.03	1.66	1.21	1.09	**0.88**	1.05	1.34	1.11	1.02
3 min	**0.78**	1.14	2.51	1.49	1.38	**0.81**	1.13	2.05	1.38	1.23	**0.84**	1.12	1.67	1.36	1.22
5 min	**0.93**	1.33	3.61	2.19	2.26	**0.94**	1.30	2.87	1.86	1.73	**0.94**	1.28	2.21	1.80	1.60

Panel C: Flash Crash

	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	1.05	1.13	1.83	0.82	**0.74**	1.05	1.12	1.60	0.84	**0.75**	1.07	1.12	1.30	0.86	**0.75**
2 min	**0.66**	1.13	2.99	1.32	1.16	**0.71**	1.09	2.39	1.29	1.12	**0.76**	1.04	1.79	1.26	1.10
3 min	**0.74**	1.29	4.37	1.55	1.56	**0.74**	1.20	3.50	1.48	1.50	**0.76**	1.15	2.62	1.44	1.42
5 min	**1.04**	1.73	7.03	2.13	2.90	**0.96**	1.57	5.55	2.03	2.70	**0.94**	1.39	4.06	1.94	2.40

Panel D: Gradual Jump with an Intermittent Flash Crash

	$\beta = 0.45$					$\beta = 0.35$					$\beta = 0.25$				
Interval	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}	RRDV	RRDV*	TRV	DV	DV_{1-3}
1 min	1.12	1.11	1.33	0.82	**0.75**	1.12	1.12	1.28	0.80	**0.76**	1.09	1.15	1.05	0.80	**0.73**
2 min	**0.84**	1.06	2.35	1.30	1.59	**0.83**	1.18	1.99	1.25	1.28	**0.93**	1.05	1.79	1.20	1.26
3 min	**0.80**	1.16	2.85	1.76	1.98	**0.81**	1.20	2.38	1.45	1.46	**0.85**	1.11	2.08	1.40	1.38
5 min	**1.02**	1.28	3.59	2.55	2.43	**1.02**	1.40	3.24	2.10	2.32	**0.99**	1.32	2.57	1.88	1.93

RMSE (multiplied by 10^5) of different IV estimators. RRDV and RRDV* are constructed from intraday candlestick information, which is obtained from second-by-second and half-minute observations, respectively. The choice of truncation parameters for TRV, DV and RRDV follows the instructions in Section 4.2, with $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}, C_\zeta^{\text{RRDV}}, \varpi) = (3, 3\sqrt{2}, 2, 1/2)$. The discretization errors of RRDV are corrected following the steps in Section 4.1. The DGP is the Heston model in Eq. (38). We simulate the additive heterogeneous Gaussian noise in Eq. (C.2), and follow the persistent noise model of Andersen et al. (2021) to simulate the three different patterns of episodic extreme return persistence. All simulated prices are further rounded to cents.