



Realized candlestick wicks

Yifan Li ^a, Ingmar Nolte ^b, Sandra Nolte ^b, Shifan Yu ^{c,*}

^a Alliance Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, United Kingdom

^b Lancaster University Management School, Lancaster LA1 4YX, United Kingdom

^c Oxford-Man Institute of Quantitative Finance, University of Oxford, Eagle House, Walton Well Road, Oxford OX2 6ED, United Kingdom

ARTICLE INFO

JEL classification:

C14
C22
C58
G14

Keywords:

High-frequency data
Integrated variance
Range-based volatility estimation
Drift burst
Extreme price movements

ABSTRACT

We develop a novel nonparametric estimator of integrated variance by summing up the squared wick lengths of intraday candlesticks over a fixed time interval. The proposed wick-based estimator is robust to short-lived extreme price movements, such as gradual jumps and flash crashes. We investigate the asymptotic properties of the proposed estimator, and show that its asymptotic variance is about four times smaller than the state-of-the-art differenced-return volatility (DV) estimator. We also develop a Hausman-type test for the presence of both jumps and episodic extreme price movements. Monte Carlo simulations and empirical applications further validate the practical reliability of our proposed estimator.

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 min. It swiftly spread to almost 8000 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, which resembles the “gradual jumps” identified by Barndorff-Nielsen et al. (2009). 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). Empirical evidence shows that such extreme events, like mini flash crashes, have occurred more frequently in recent years, which raises widespread concern about market inefficiency and vulnerability (Laly and Petitjean, 2020; Flora and Renò, 2024). These market glitches are also a threat to the standard theoretical framework, as they represent temporary violations of the Itô semimartingale assumption and potentially the no-arbitrage principle (Andersen et al., 2025). Two recent influential studies, Christensen et al. (2022) and Andersen et al. (2023), 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. (2023) consider these unusual patterns as outcomes of the temporary

* Corresponding author.

E-mail addresses: yifan.li@manchester.ac.uk (Y. Li), i.nolte@lancaster.ac.uk (I. Nolte), s.nolte@lancaster.ac.uk (S. Nolte), shifan.yu@eng.ox.ac.uk (S. Yu).

¹ See also Kolokolov (2023), Mancini (2023), Bellia et al. (2024), Christensen and Kolokolov (2024), Flora and Renò (2024), and Laurent et al. (2024) for some recent theoretical and empirical studies on drift burst.

<https://doi.org/10.1016/j.jeconom.2025.106014>

Received 8 September 2023; Received in revised form 28 April 2025; Accepted 28 April 2025

Available online 27 May 2025

0304-4076/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

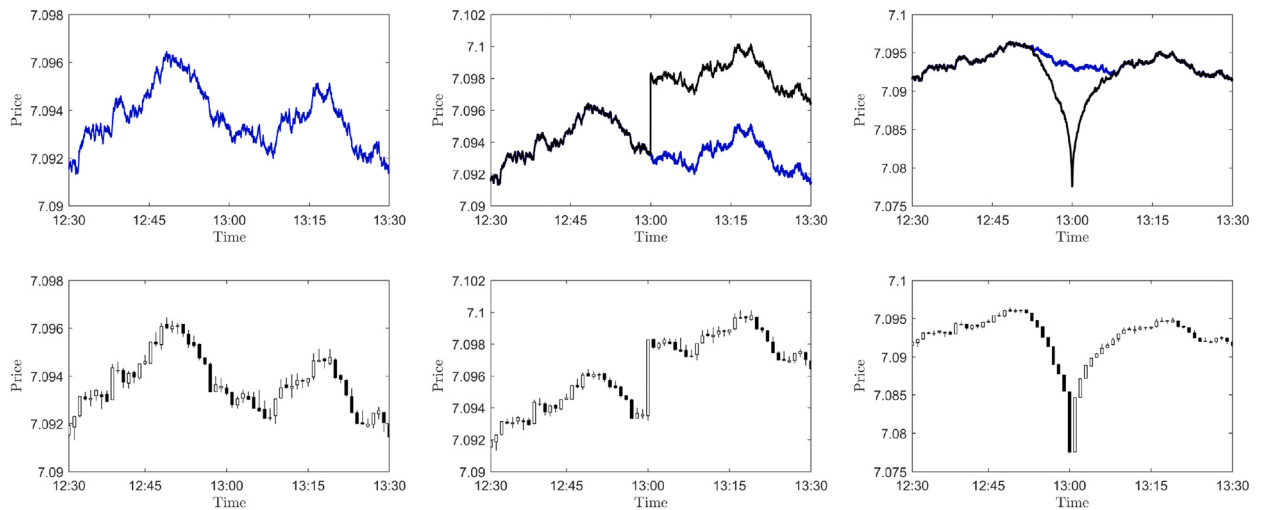


Fig. 1. Examples of one-minute intraday candlestick charts for the simulated second-by-second log-prices from different data generating processes (DGPs): continuous (left), discontinuous with a jump (middle), continuous with a V-shaped flash crash (right). White candlesticks indicate upward movements, and black ones denote downward movements.

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), or some combinations thereof (Vetter, 2010; Corsi et al., 2010) and further innovations (Andersen et al., 2012; 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 but continuous price movements has long been ignored. Laurent and Shi (2020) quantify the bias of (original and modified) RV and realized bipower variation (BV) in the presence of a nonzero drift coefficient. Laurent et al. (2024) derive the asymptotic properties of some realized measures in the presence of short-lived explosive trends. The differenced-return volatility (DV) estimator of Andersen et al. (2023) provides the first IV estimator robust to this type of episodic Itô semimartingale violation.²

This paper introduces a novel nonparametric estimator of IV that is designed to be robust to the presence of short-lived extreme price movements, including jumps and explosive trends. Inspired by the drift-independent variance estimator of Rogers and Satchell (1991), we develop the wick-based volatility (WV) estimator that utilizes intraday “candlestick” information, i.e., the open, high, low, and close prices (OHLC) within each short time interval. The WV estimator is defined as the sum of squared wick lengths of all intraday candlesticks over a fixed time interval, where the wick length of each candlestick is measured by the difference between the high-low price range and the absolute open-close return. Intuitively, the range-return differencing mechanically offsets extreme price movements that dominate both the range and return of the same candlestick, rendering the WV estimator robust to short-lived price irregularities within that interval. 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 large discontinuities or short-lived explosive trends. The pairwise offset between ranges and returns offers the WV estimator a built-in robustness to such extreme events.

We derive the consistency and asymptotic normality of the WV estimator under infill asymptotics, which reveals that the asymptotic variance of WV is approximately four time smaller than that of DV. Importantly, we demonstrate that the presence of jumps with both finite and infinite activities, as well as the episodes of short-lived explosive trends as modeled by Christensen et al. (2022) and Andersen et al. (2023), have no impact on the consistency and asymptotic distribution of WV. This omnibus robustness of WV is a consequence of its self-differencing design, which is achieved by the DV estimator through both truncation and an adjacent-differencing strategy under slightly more restrictive conditions. Simulation results confirm that our new estimator outperforms state-of-the-art competitors in scenarios with various specifications of extreme price movements. Our empirical application focuses

² Throughout this paper, “drift robustness” refers to the invariance of asymptotic distribution when the short-lived explosive trends are present. Laurent et al. (2024) show that while the standard RV remains consistent in the presence of explosive trends, its asymptotic distribution may change if the explosive drift term dominates the variance, see Section 3.1 for further discussion.

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 WV estimates can achieve smaller forecast errors for both robust and non-robust IV proxies than all selected benchmark models, especially when the previous day is identified with extreme price movements.

Our estimator is closely related to the literature on range-based volatility estimation. Since the seminal works of [Parkinson \(1980\)](#) and [Garman and Klass \(1980\)](#), a large body of literature documents the substantial precision gain of range-based volatility estimators relative to the ones based only on returns, 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 \(2007\)](#) and [Martens and van Dijk \(2007\)](#) is the first nonparametric IV measure constructed from high-frequency intraday ranges, which is then extended by [Christensen et al. \(2009\)](#) and [Christensen and Podolskij \(2012\)](#). More recently, [Li et al. \(2024\)](#) and [Bollerslev et al. \(2024\)](#) introduce the optimal candlestick-based spot volatility estimators with and without a linear functional form restriction, respectively, which exploit the broad availability of intraday candlestick charts in practice. As a fundamental tool in technical analysis that predates the rise of high-frequency data, intraday candlesticks are now readily accessible through most online trading platforms and various public or commercial data providers (e.g., Yahoo Finance and Bloomberg). This widespread availability facilitates the straightforward implementation of candlestick-based inference techniques. We contribute to this stream of literature by offering a practical and cost-effective IV estimator for general investors who might otherwise struggle to benefit from expensive tick-level data.

The remainder of this paper is structured as follows: Section 2 introduces the new wick-based IV estimator and establishes the asymptotic properties. Section 3 demonstrates its robustness in the presence of short-lived explosive trends. Section 4 includes an extensive Monte Carlo study that verifies its asymptotic unbiasedness and illustrates the finite-sample performance. Section 5 presents an empirical application of the WV estimator in a volatility forecasting exercise. Section 6 concludes. Proofs and additional results can be found in the Online Appendix.

2. Volatility estimation based on candlestick wicks

2.1. Wick-based volatility estimator

For a finite time interval $[0, t]$, e.g., a trading day, we apply an equidistant partition at $0 < \Delta_n < 2\Delta_n < \dots < n\Delta_n \leq t$ 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 OHLC of a one-dimensional process $X = (X_t)_{t \geq 0}$ of the efficient logarithmic price over the i th interval can be expressed as, respectively,

$$O_i = X_{(i-1)\Delta_n}, \quad H_i = \sup_{s \in I_{n,i}} X_s, \quad L_i = \inf_{s \in I_{n,i}} X_s, \quad C_i = X_{i\Delta_n}. \quad (1)$$

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

$$w_i = H_i - L_i, \quad r_i = C_i - O_i. \quad (2)$$

The wick-based volatility (WV) estimator based on the differences between ranges and absolute returns is defined as

$$WV_{t,n} = \frac{1}{\Lambda_2} \sum_{i=1}^n (w_i - |r_i|)^2, \quad (3)$$

with

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

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 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$:

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad (5)$$

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). For the WV estimator, it holds that as $\Delta_n \rightarrow 0$,

$$WV_{t,n} \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^2 ds, \quad (6)$$

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 finite t .

Theorem 1 indicates that WV 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

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

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 $\Delta_n \rightarrow 0$, we have

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

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 $\zeta(s)$ is the Riemann zeta function.

We denote by $\xrightarrow{\mathcal{L}-s}$ the stable convergence in law, and by $\mathcal{MN}$ a mixed normal distribution, i.e., a normal distribution conditional on the realization of its F -conditional variance, which is a random variable.

Remark 2. Compared with the DV estimator of [Andersen et al. \(2023\)](#) which features a variance factor of 3, the asymptotic variance of WV is about four times smaller under infill asymptotics. This result may initially seem surprising, given that [Kolokolov et al. \(2025\)](#) demonstrate that DV attains the variance lower bound of drift-robust IV estimator based on returns from two adjacent blocks. The improvement of WV over DV stems from the additional information contained in high-frequency intraday ranges, which leads to a different limiting statistical experiment. This information also results in the reduced variance of the RRV estimator of [Christensen and Podolskij \(2007\)](#) compared to RV that can itself reach the semiparametric efficiency bound of return-based IV estimation ([Renault et al., 2017](#)). Finally, if robustness to locally explosive trends is not a priority, a variance-optimal candlestick-based IV estimator can be constructed in the spirit of [Garman and Klass \(1980\)](#) and [Li et al. \(2024\)](#), see Section 3.3 for more details.

For feasible implementation of the asymptotic distribution in [Theorem 2](#), we can estimate the integrated quarticity (IQ) with the following wick-based quarticity (WQ) estimator constructed analogously to WV:

$$WQ_{t,n} = \frac{n}{\Lambda_4} \sum_{i=1}^n (w_i - |r_i|)^4. \quad (9)$$

With techniques similar to [Theorem 1](#), we can establish the consistency of WQ, as well as a feasible CLT by the property of stable convergence:

Corollary 1 (Feasible Inference). Under the same conditions as in [Theorem 1](#), it holds that

$$WQ_{t,n} \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^4 ds, \quad (10)$$

$$\sqrt{\frac{n}{\Theta WQ_{t,n}}} \left(WV_{t,n} - \int_0^t \sigma_s^2 ds \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad (11)$$

Remark 3. The WQ estimator shares the same robustness properties as WV in the presence of both jumps ([Theorem 3](#)) and short-lived explosive trends ([Theorems 4 to 6](#)). The proofs are analogous and thus omitted here.

2.3. Jumps

We examine the behavior of WV 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:

$$X_t = X'_t + \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| \leq 1\}} (\underline{p} - \underline{q})(ds, dx) + \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| > 1\}} \underline{p}(ds, dx), \quad (12)$$

where X' follows Eq. (5), $\underline{p}(dt, dx)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with a compensator $\underline{q}(dt, dx) = dt \otimes \lambda(dx)$, 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 $(\tau_m)_{m \geq 1}$ of stopping times increasing to ∞ , and a sequence of deterministic nonnegative functions f_m on $\mathbb{R}$ for each m , which satisfies $|\delta(\omega, t, x)| \wedge 1 \leq f_m(x)$ for all (ω, t, x) with $t \leq \tau_m(\omega)$, and $\int_{\mathbb{R}} |f_m|^r \lambda(dx) < \infty$ for some $r \in [0, 1)$.

Remark 4. 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.

To demonstrate the robustness of WV in the presence of jumps, we denote by $WV_{t,n}$ and $WV'_{t,n}$ the WV estimators constructed from the discontinuous Itô semimartingale X in Eq. (12) and its continuous part X' , respectively. The difference between the observed $WV_{t,n}$ and the infeasible $WV'_{t,n}$ quantifies the additional bias induced by jumps, and the following result shows that this bias is asymptotically negligible as it does not affect the asymptotic distribution of the WV estimator.

Theorem 3 (Jump Robustness). Assume that the efficient price X follows a discontinuous Itô semimartingale in Eq. (12) with Assumption 2 satisfied. Then as $\Delta_n \rightarrow 0$,

$$WV_{t,n} - WV'_{t,n} = o_p(\sqrt{\Delta_n}). \quad (13)$$

Theorem 3 indicates the robustness of WV to both finite-activity and infinite-activity but finite-variation jumps in the efficient price. In each of the intervals containing “big” jumps, i.e., jumps with sizes of a higher asymptotic order than the continuous component, both the range and absolute return are dominated by the discontinuity under infill asymptotics.³ Consequently, the “big” jumps are canceled out in the range-return differences, or candlestick wicks, and only contribute to an $o_p(\sqrt{\Delta_n})$ bias in WV. For “small” jumps that appear everywhere and cannot be mechanically canceled in the wicks, their collective impact results in another $o_p(\sqrt{\Delta_n})$ bias. Since the biases from both “big” and “small” jumps diminish faster than the convergence rate of WV, their combined impact becomes asymptotically negligible when normalized as in the CLT in Theorem 2. Therefore, the asymptotic distribution of WV remains unaffected.

It is noteworthy that the range-return differencing approach achieves jump robustness in a straightforward and practical way without requiring additional tuning parameters, which provides a novel and easy-to-implement alternative to traditional return-based methods, such as bipower estimators (Barndorff-Nielsen and Shephard, 2004, 2006) and the return truncation technique of Mancini (2009) used in both the truncated realized volatility (TRV) and DV estimators, when both the range and return information are available.

3. Short-lived explosive trends

As noted by Andersen et al. (2023), market participants may imperfectly react to relevant new information, and sometimes induce short-lived explosive trends in asset prices. Phenomena such as “gradual jumps” and “flash crashes” will be discussed in this section as representative examples.

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

$$\int_{\tau}^{\tau+\Delta_n} \mu_s ds = O_p(\Delta_n) \quad \text{and} \quad \int_{\tau}^{\tau+\Delta_n} \sigma_s dW_s = O_p(\sqrt{\Delta_n}), \quad \text{as } \Delta_n \rightarrow 0. \quad (14)$$

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 process X defined on $(\Omega, \mathcal{F}, (F_t)_{t \geq 0}, \mathbb{P})$ is decomposed as $X = X' + H$, where X' follows Eq. (5). The drift burst component H is given by

$$H_t = \int_0^t \mu_s^b ds = \int_0^t \frac{c_1 \mathbb{1}_{\{s < \tau\}} + c_2 \mathbb{1}_{\{s > \tau\}}}{|s - \tau|^\alpha} ds, \quad (15)$$

with $\tau \geq 0$, and constants c_1 , c_2 , and α .

³ As discussed in Section 2.3 of Aït-Sahalia and Jacod (2009), each individual interval of length Δ_n mostly contain a single “big” jump under infill asymptotics; see more details in Lemma A.4 in Online Appendix A.5.

It is assumed that the explosive drift term μ_t^b has a singularity at the “burst time” τ , and spikes dramatically in the vicinity of τ . The order of magnitude of the H increment near τ is given by

$$\int_{\tau-A_n}^{\tau} \frac{c_1}{(\tau-s)^\alpha} ds \asymp \int_{\tau}^{\tau+A_n} \frac{c_2}{(s-\tau)^\alpha} ds = O_p(\Delta_n^{1-\alpha}). \quad (16)$$

We 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 explosive trend.⁴ As t moves further away from the burst time τ , the size of the H increments gradually decreases.

We allow for different constants c_1 and c_2 before and after τ .⁵ When $1/2 < \alpha < 1$ and $c_1 c_2 < 0$, the trajectory shows a “V-shape” (or “Λ-shape”) in the neighborhood of τ , due to a discontinuity in the sign of H which locally dominates log-returns (Flora and Renò, 2024). Different choices of c_1 and c_2 can be employed to mimic patterns akin to (i) a gradual jump ($c_1 = 0$ or $c_2 = 0$), or (ii) a V-shaped flash crash ($c_1 < 0$ and $c_2 > 0$). In the following theoretical derivation, we focus on these two scenarios of short-lived explosive trends.

The robustness of WV in the presence of drift burst is summarized in the following proposition. Similar to Theorem 3, we use $WV_{t,n}$ and $WV'_{t,n}$ to represent the WV estimators constructed from X and X' under Assumption 3, respectively, and the result below is the counterpart to Theorem 3 within the drift burst framework:

Theorem 4. Let Assumption 3 hold with $1/2 < \alpha < 1$. For the WV estimator, it holds that

(i) If $c_1 = 0$ or $c_2 = 0$ (Gradual Jump):

$$WV_{t,n} - WV'_{t,n} = O_p(\Delta_n^{1/2\alpha}), \quad (17)$$

(ii) If $c_1 < 0$ and $c_2 > 0$ (Flash Crash):

$$WV_{t,n} - WV'_{t,n} = O_p(\Delta_n^{1/2\alpha}) + O_p(\Delta_n^{2-2\alpha}). \quad (18)$$

Under scenario (i), similar to the DV estimator, the difference between two WV estimators arises from two sources. First, with $1/2 < \alpha < 1$, the component H dominates the price movement in the vicinity of τ , i.e., the explosive region. For each interval within this region, both the range and return are dominated by a common excessive component, i.e., the H increment of a higher asymptotic order than $\sqrt{\Delta_n}$. By construction, the self-differencing feature of the candlestick wicks completely eliminates its impact on WV. Second, the H increments outside the vicinity of τ are not offset by self-differencing, but they only have a diminished impact on WV. The biases induced by the component H in both the explosive and non-explosive regions are $O_p(\Delta_n^{1/2\alpha})$, which are asymptotically negligible under the same normalization by $\sqrt{\Delta_n}$ as in the CLT in Theorem 2, and thus have no impact on both the consistency and asymptotic distribution in Section 2.2.

The flash crash under scenario (ii) can be interpreted as a combination of two inverted gradual jumps from scenario (i). However, the V-shape around the reversal point τ introduces an additional bias in WV. The bias arises from the specific short interval that contains τ (referred to as “V-interval” hereafter). In the V-interval, the abrupt reversal in price movement results in an unbalanced range and return. Consequently, the H increment over the V-interval is not fully offset by self-differencing, which leads to an addition bias of order $O_p(\Delta_n^{2-2\alpha})$. While WV remains a consistent IV estimator for all $1/2 < \alpha < 1$, the bias from the V-interval may invalidate the CLT in Theorem 2 when the V-shape is steep. Specifically, for $1/2 < \alpha < 3/4$, the bias is of order $o_p(\sqrt{\Delta_n})$, which aligns with the result in scenario (i). However, for $3/4 < \alpha < 1$, the CLT becomes invalid since $\Delta_n^{2-2\alpha} \gg \sqrt{\Delta_n}$. Fig. 2 illustrates a finite-sample example of the candlestick in the V-interval, where the candlestick has a long lower wick, i.e., the so-called *hammer* candlestick.

We note that the specific V-shape bias can be effectively eliminated with a truncation threshold for the wick lengths. Specifically, we define the wick-truncated WV estimator as

$$\overline{WV}_{t,n} = \frac{1}{\Lambda_2} \sum_{i=1}^n (w_i - |r_i|)^2 \mathbb{1}_{\{|w_i - |r_i|| \leq \zeta \Delta_n^\varpi\}}, \quad (19)$$

for some $\zeta > 0$ and $\varpi \in (1/4, 1/2)$, where the truncation threshold follows the form of Mancini (2009). Theorem 5 demonstrates that the wick truncation effectively eliminates the impact of the V-shape bias on the CLT.

Theorem 5. Let Assumption 3 hold with $1/2 < \alpha < 1$. It holds that

$$\overline{WV}_{t,n} - WV'_{t,n} = O_p(\Delta_n^{1/2\alpha}) = o_p(\sqrt{\Delta_n}). \quad (20)$$

⁴ We follow Andersen et al. (2023) to consider such episodic Itô semimartingale violations with only an exploding drift. It does not necessarily permit local arbitrage opportunities in the specification of Christensen et al. (2022), which accommodates simultaneous drift and volatility bursts with different rates of explosion.

⁵ The constants can be readily generalized to some continuous deterministic functions, see Flora and Renò (2024).

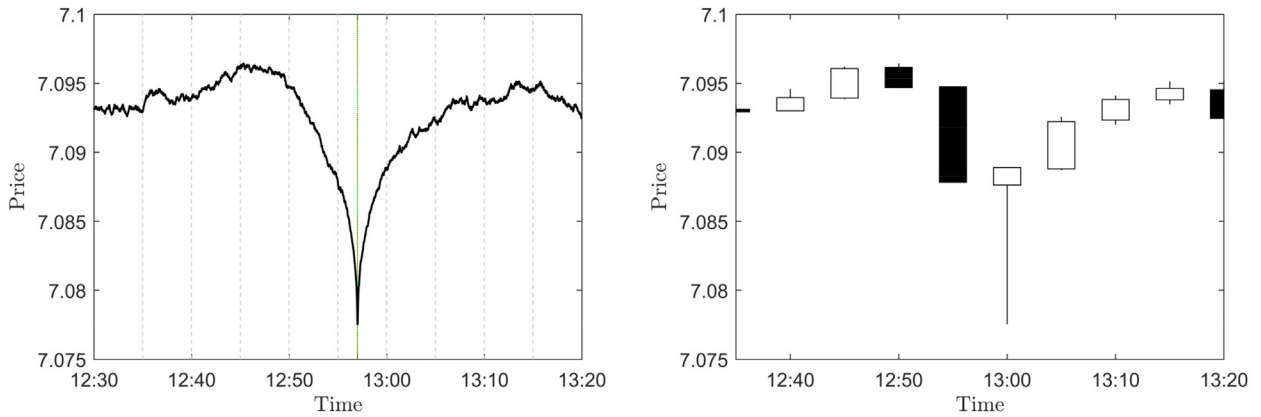


Fig. 2. 5-min candlesticks around a simulated V-shaped flash crash.

Unlike the return truncation applied in both TRV and DV, the WV estimator does not rely on the wick truncation to eliminate the impact of both jumps and monotone explosive trends. In the limit, the truncation occurs only on the individual hammer candlestick of the V-interval, such that it excludes the potential impact of the V-shape bias of order $O_p(\Delta_n^{2-2\alpha})$ on the CLT. Therefore, for all $1/2 < \alpha < 1$, the bias due to drift burst diminishes faster than the convergence rate $\sqrt{\Delta_n}$, and the overall estimation bias of wick-truncated WV under the drift burst model is identical to that under the continuous Itô semimartingale assumption:

$$\overline{WV}_{t,n} - \int_0^t \sigma_s^2 ds = O_p(\sqrt{\Delta_n}). \quad (21)$$

The above representation is similar to Theorem 2 in Andersen et al. (2023). We note that the order of the estimation bias in Eq. (21) is not affected by the drift explosion rate α , whereas in Andersen et al. (2023), the same result holds only when an additional condition on the truncation threshold is satisfied. In other words, the robustness of our wick-based estimator holds under slightly weaker condition than those required for DV. Furthermore, the result can be extended to the case with stochastically distributed explosion times over $[0, t]$, as illustrated next in Section 3.2.

As demonstrated in Lemma 3.1 and Theorem 3.1 of Laurent et al. (2024), in the absence of jumps, the standard RV estimator remains consistent in the presence of drift burst under Assumption 3, and the CLT also remains intact when the drift explodes mildly with $1/2 < \alpha < 3/4$. When $3/4 < \alpha < 1$, the standard CLT of RV fails, similar to the original WV in the steep V-shape case: The drift-induced bias is of order $O_p(\Delta_n^{2-2\alpha})$, which diminishes at a slower rate than $\sqrt{\Delta_n}$. However, for $3/4 < \alpha < 1$, the CLT distortion for WV arises only from the V-interval (scenario (ii) in Theorem 4), whereas for RV, all intervals in the vicinity of τ contribute to its CLT distortion. Despite the consistency of RV in both scenarios in Theorem 4, it is subject to severe finite-sample drift-induced bias, as evidenced by daily RV examples in Table 1 of Andersen et al. (2023).

In addition, similar to the WV, while WQ remains consistent under both scenarios in Theorem 4 for any $1/2 < \alpha < 1$, the same wick truncation can be employed to mitigate the finite-sample impact from the individual hammer candlestick:

$$\overline{WQ}_{t,n} = \frac{n}{\Lambda_4} \sum_{i=1}^n (w_i - |r_i|)^4 \mathbb{1}_{\{|w_i - |r_i|| \leq \zeta \Delta_n^{\alpha}\}}. \quad (22)$$

For both WV and WQ, we note that the wick truncation has asymptotically no effect on intervals that do not contain V-shapes, even in the presence of monotone explosive trends or various types of jumps.⁶ In the remainder of this paper, we will consistently use the wick-truncated versions in Eqs. (19) and (22), unless otherwise specified.

3.2. Persistent noise model

As an extension of the gradual jump model of Barndorff-Nielsen et al. (2009), Andersen et al. (2023) 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:

$$X_t = X'_t + \sum_{0 \leq s \leq t} \Delta X_s + H_t, \quad (23)$$

⁶ See Online Appendix A.7 for further discussion.

where X' follows Eq. (5). We denote by τ_i the first occurrence of the i th persistent noise episode, so that $\tau_1, \tau_2, \dots, \tau_N \in [0, t]$ form an increasing sequence of stopping times, with N finite almost surely. The persistent noise component is given by

$$H_t = \sum_{i: \tau_i \leq t} H_t^{(i)} \mathbb{1}_{\{\epsilon_t^{(i)} \geq 0\}}, \quad (24)$$

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

$$H_t^{(i)} = f^{(i)}(\Delta X_{\tau_i}, \eta_{\tau_i}) g^{(i)}(t), \quad (25)$$

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:

$$g_{gj}^{(i)}(t) = \left[1 - \left(\frac{t - \tau_i}{\bar{\tau}_i - \tau_i} \right)^\beta \right] \mathbb{1}_{\{t \in [\tau_i, \bar{\tau}_i]\}}, \quad (26)$$

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

$$g_{fc}^{(i)}(t) = c_1^{(i)} \left[1 - \left(\frac{\check{\tau}_i - t}{\check{\tau}_i - \tau_i} \right)^{\beta_1} \right] \mathbb{1}_{\{t \in [\tau_i, \check{\tau}_i]\}} + c_2^{(i)} \left[1 - \left(\frac{t - \check{\tau}_i}{\bar{\tau}_i - \check{\tau}_i} \right)^{\beta_2} \right] \mathbb{1}_{\{t \in [\check{\tau}_i, \bar{\tau}_i]\}}, \quad (27)$$

for some $\mathcal{F}_{\tau_i}$ -measurable random variables $c_1^{(i)}$ and $\check{\tau}_i > \tau_i$, some $\mathcal{F}_{\check{\tau}_i}$ -measurable random variables $c_2^{(i)}$ and $\bar{\tau}_i > \check{\tau}_i$, and constants $0 < \beta_1, \beta_2 < 1/2$. Moreover,

$$\epsilon_t^{(i)} = \sum_{s \in [\tau_i, t]} \Delta \epsilon_s \quad (28)$$

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

Each of the episodes is activated and terminated by the realizations of τ_i and $\bar{\tau}_i$, or randomly ended by $\epsilon_t^{(i)}$ before $\bar{\tau}_i$. 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 $H^{(i)}$ which correspond to two scenarios in [Theorem 4](#):

- (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. (26), 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. (27). 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 reversal point at a random time $\check{\tau}_i$.

The scenario (i) and (ii) correspond to two phenomena discussed earlier: 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 [Theorem 5](#) when the observed prices persistently deviate from the fundamental values due to short-lived market inefficiency.

Theorem 6. Assume that the market price X follows a contaminated semimartingale in Eq. (23) with $r = 0$ in [Assumption 2](#), and there exists a persistent noise episode within $[0, t]$. The function $g^{(1)}$ in the persistent noise component $H_t^{(1)}$ takes one of the two forms in Eqs. (26) and (27) with $0 < \beta < 1/2$ or $0 < \beta_1, \beta_2 < 1/2$. For the WV estimator, it holds that

$$\overline{WV}_{t,n} - WV'_{t,n} = \begin{cases} O_p\left(\Delta_n^{\frac{1}{2(1-\beta)}}\right) = o_p(\sqrt{\Delta_n}), & \text{when } g^{(1)} = g_{gj}^{(1)}, \\ O_p\left(\Delta_n^{\frac{1}{2(1-\beta_1 \wedge \beta_2)}}\right) = o_p(\sqrt{\Delta_n}), & \text{when } g^{(1)} = g_{fc}^{(1)}. \end{cases} \quad (29)$$

In [Theorem 6](#), our discussion is confined to a simplified scenario featuring a single persistent noise episode. Similar to the result in [Theorem 4](#) with a drift burst, the additional bias diminishes faster than the convergence rate of WV, such that it does not affect the asymptotic distribution in [Theorem 2](#), and the order of overall estimation bias of WV keeps identical to that under the continuous Itô semimartingale assumption. The results obtained in the simplified scenario can be straightforwardly extended to more general cases involving a finite number of non-overlapping episodes. Simulation results in [Section 4](#) demonstrate that the WV estimator remains unbiased in the presence of a gradual jump accompanied by an intermittent small flash crash.

It is important to note that the persistent noise discussed in this section is not what is traditionally referred to as “market microstructure noise” in the high-frequency financial econometrics literature. Market microstructure noise captures a variety of idiosyncrasies inherent in the trading process, such as bid–ask bounces and the discreteness of price changes, among other factors that contaminate observed transaction prices at high frequencies. As highlighted by [Bollerslev et al. \(2024\)](#) and discussed in the

Introduction, the wide availability of intraday candlesticks is a crucial motivation for our “inexpensive” volatility estimator. The WV estimator enables practitioners to utilize summary information for each “not-too-finely” sampled interval without the requirement of more granular transaction records, which parallels the common use of sparsely sampled data in both literature and practice to mitigate the impact of market microstructure noise (Aït-Sahalia et al., 2005; Liu et al., 2015; Aït-Sahalia and Xiu, 2019). However, a comprehensive investigation into the asymptotic properties of WV constructed from finely sampled, noise-contaminated candlesticks requires more explicit assumptions about the fine structure of market microstructure noise. This, in turn, necessitates a better understanding of various market frictions, particularly over non-trivial episodes of extreme price movements, which are difficult to rationalize within the standard framework and remain contentious in recent literature (Bollerslev et al., 2024), and is thus beyond the scope of this paper.

3.3. A Hausman test for extreme price movements

In this section, we introduce a test designed to identify the presence of extreme price movements, which include both jumps (as defined in Section 2.3) and short-lived explosive trends (as discussed in Sections 3.1 and 3.2). The test contrasts these phenomena against the null hypothesis ($\mathbb{H}_0$) of a continuous Itô semimartingale. In this section, we do not attempt to differentiate between these two types of extreme price movements, as both are responses to relevant new information, albeit at different speeds due to market efficiency (Aït-Sahalia et al., 2025).

The test statistic we construct builds on the insights of Hausman (1978) and its recent high-frequency application in Aït-Sahalia and Xiu (2019): It compares an asymptotically variance-optimal IV estimator, which is consistent under $\mathbb{H}_0$ but not robust under $\mathbb{H}_1$, against a suboptimal robust estimator. The wick-truncated WV estimator clearly satisfies the criteria of the suboptimal estimator. For the variance-optimal estimator under $\mathbb{H}_0$, we propose to use the variance-optimal one within the class of estimators constructed from a linear combination of second-order candlestick features $\theta_i = (w_i^2, w_i|r_i|, r_i^2)'$.

We start with some additional notations: Let $\omega = (\omega_1, \omega_2, \omega_3)'$ denote a 3-by-1 weight vector, where the weights satisfy $\sum_{i=1}^3 \omega_i = 1$. The set of all such weight vectors is denoted by $\mathcal{W}$. Let $\Sigma = \text{diag}(4 \ln 2, 3/2, 1)$ denote a 3-by-3 diagonal matrix. Each vector ω identifies a corresponding candlestick-based volatility (KV) estimator based on θ_i for all $i \in \{1, \dots, n\}$:

$$KV_{t,n}(\omega) = \sum_{i=1}^n \omega' \Sigma^{-1} \theta_i, \quad (30)$$

and the collection of all such estimators is denoted by $\mathcal{V}_{\mathcal{W}}$. We have the following result:

Theorem 7. Assume that the efficient price X follows a continuous Itô semimartingale in Eq. (5) with Assumption 1 satisfied under $\mathbb{H}_0$. For all $\omega \in \mathcal{W}$, it holds that

$$\frac{1}{\sqrt{\Delta_n}} \left(KV_{t,n}(\omega) - \int_0^t \sigma_s^2 ds \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \Theta(\omega) \int_0^t \sigma_s^4 ds \right), \quad (31)$$

where

$$\Theta(\omega) = \omega' \Sigma^{-1} \Omega \Sigma^{-1} \omega - 1, \quad (32)$$

with the 3-by-3 covariance matrix Ω specified in Eq. (A.85) in Online Appendix A.10. The variance factor $\Theta(\omega)$ is minimized by $\omega^* \approx (1.7103, -0.7647, 0.0544)'$, and the corresponding variance-optimal estimator in $\mathcal{V}_{\mathcal{W}}$ is defined as

$$OKV_{t,n} = KV_{t,n}(\omega^*) \approx \sum_{i=1}^n 0.6169 w_i^2 - 0.5098 w_i |r_i| + 0.0544 r_i^2, \quad (33)$$

with the variance factor $\Theta^* = \Theta(\omega^*) \approx 0.2594$.

We note that the class $\mathcal{V}_{\mathcal{W}}$ nests the standard RV estimator with $\omega = (0, 0, 1)'$, the RRV estimator with $\omega = (1, 0, 0)'$, and WV with $\omega = \Lambda_2^{-1}(4 \ln 2, -3, 1)'$. The variance-optimal OKV can be conveniently expressed as a linear combination of these three estimators:

$$OKV_{t,n} \approx 0.2 WV_{t,n} + RRV_{t,n} - 0.2 RV_{t,n}, \quad (34)$$

which can also be interpreted as a realized version of the optimal candlestick-based spot volatility estimator proposed by Li et al. (2024).⁷ This form highlights that, under $\mathbb{H}_1$, if either RV or RRV is not robust to jumps or explosive trends, neither will OKV. Consequently, a Hausman test can be constructed by comparing OKV with the wick-truncated WV, which retains robustness under $\mathbb{H}_1$.

The Hausman test statistic is defined as follows:

$$T_{t,n} = \frac{n(OKV_{t,n} - \overline{WV}_{t,n})^2}{\Xi \overline{WQ}_{t,n}}, \quad (35)$$

⁷ Following Meilijson (2011) and Bollerslev et al. (2024), the variance factor Θ^* can be further reduced to approximately 0.2587 by including quadratic terms involving a candlestick asymmetry measure. However, as this substantially complicates the form of the estimator without an economically significant improvement in precision, we will not explore it further in this discussion.

where $\Xi = \Theta - \Theta^* \approx 0.4651$. For a given significance level α , we reject $\mathbb{H}_0$ if $T_{t,n}$ exceeds $c_{1-\alpha}$, i.e., the $(1 - \alpha)\%$ quantile of the χ^2 -distribution with the degree of freedom being equal to 1. The following result demonstrates that the Hausman test is asymptotically valid under both $\mathbb{H}_0$ and $\mathbb{H}_1$:

Corollary 2. *The Hausman test statistic is correctly sized under $\mathbb{H}_0$:*

$$\mathbb{P}(T_{t,n} > c_{1-\alpha}) \rightarrow \alpha. \quad (36)$$

Under $\mathbb{H}_1$ where X follows (i) the discontinuous Itô semimartingale in Eq. (12), (ii) the drift burst model under [Assumption 3](#) for $\alpha \in (3/4, 1)$, or (iii) the persistent noise model under [Assumption 4](#) for β or $\beta_1 \vee \beta_2 \in (0, 1/4)$, the test is consistent with

$$\mathbb{P}(T_{t,n} > c_{1-\alpha}) \rightarrow 1. \quad (37)$$

The above result shows that the Hausman test is consistent against the alternative of finite-activity and infinite-activity but finite-variation jumps. However, it is not consistent in the presence of drift burst or persistent noise with relatively mild explosion rates, i.e., when $1/2 < \alpha < 3/4$ or $1/4 < \beta < 1/2$. This is due to the fact that, under these parameter choices, the asymptotic distributions of both RV and RRV remain unaffected, which can be verified following Theorem 3.1 of [Laurent et al. \(2024\)](#). Consequently, the test asymptotically lacks power against such mildly explosive trends. However, our simulation results in [Table 3](#) reveal that the test exhibits non-trivial power in finite samples. This finite-sample power originates from the substantial drift-induced bias in RV and RRV (and hence in OKV) under empirically realistic sampling frequencies. The finite-sample bias inflates the test statistic and results in strong finite-sample power. This phenomenon thus echoes the findings of [Christensen et al. \(2022\)](#), which shows that, even when the theory predicts no asymptotic power, finite-sample power of drift burst tests can still be achieved.

Unlike the DV-based test proposed by [Andersen et al. \(2023\)](#), which is specifically designed to detect short-lived explosive trends, our test also demonstrates power against the jump alternative. It is possible to generalize our result by augmenting OKV with the standard return truncation technique of [Mancini \(2009\)](#), thereby rendering the Hausman test robust to the presence of jumps. However, as our goal is to develop an omnibus test to identify both types of extreme price movements, we do not formalize this extension here and leave it as a direction for future research.

4. Monte Carlo simulations

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

4.1. Simulation design

We generate synthetic intraday candlesticks with a Heston model for the latent efficient price process X ([Heston, 1993](#)):

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

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

$$J_t = \sum_{i=1}^{N_t} Z_i, \quad (39)$$

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

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

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. We follow the persistent noise model of [Andersen et al. \(2023\)](#) to simulate three different patterns of short-lived explosive trends⁸:

- **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. (25), 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 $g_{g_j}^{(1)}$ in Eq. (26), and $(\tau, \bar{\tau}) = (0.5, 0.59)$.

⁸ As shown in Eqs. (56) and (57) in [Andersen et al. \(2023\)](#), there exists an asymptotic correspondence between the two models of explosive trends in [Section 3](#), and they are equivalent with identical asymptotic analyses when $\beta = 1 - \alpha$. The simulation results with the drift burst model in [Assumption 3](#) indicate the same qualitative conclusions.

Table 1
Monte Carlo bias results (%).

Panel A: WV without wick truncation											
Interval	$H = 0$		Gradual jump			Flash crash			Gradual jump with an intermittent flash crash		
	No jump	Jumps	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
5 s	-0.31	-0.36	-0.81	-0.83	-0.73	-0.92	-0.78	-0.42	-0.92	-0.86	-0.73
10 s	-0.29	-0.21	-1.13	-0.96	-0.93	-0.95	-0.12	1.20	-1.06	-0.94	-0.56
30 s	-0.17	-0.10	-1.95	-1.68	-1.40	-1.69	0.08	2.64	-2.00	-1.51	-0.86
1 min	-0.13	-0.09	-2.84	-2.52	-1.89	-3.08	-1.13	1.83	-3.06	-2.51	-1.59
2 min	-0.12	-0.02	-3.92	-3.28	-2.43	10.46	22.12	35.33	-4.65	-3.70	-2.46
3 min	-0.13	-0.05	-4.74	-4.04	-3.12	10.72	23.05	36.79	-0.63	2.25	4.59
5 min	-0.18	0.50	-5.66	-5.03	-3.90	13.32	28.34	42.96	-7.13	-5.66	-4.44
Panel B: WV with wick truncation ($C_c = 3$)											
Interval	$H = 0$		Gradual jump			Flash crash			Gradual jump with an intermittent flash crash		
	No jump	Jumps	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
5 s	-0.31	-0.36	-0.81	-0.83	-0.73	-0.94	-0.99	-0.99	-0.92	-0.86	-0.75
10 s	-0.29	-0.21	-1.13	-0.96	-0.93	-1.40	-1.44	-1.32	-1.06	-0.97	-0.82
30 s	-0.17	-0.10	-1.95	-1.68	-1.40	-1.97	-2.35	-2.08	-2.00	-1.52	-0.93
1 min	-0.13	-0.10	-2.85	-2.52	-1.94	-3.09	-1.22	-1.63	-3.07	-2.51	-1.61
2 min	-0.12	-0.03	-3.96	-3.30	-2.60	-6.18	-5.56	-4.34	-4.67	-3.70	-2.53
3 min	-0.14	-0.07	-4.77	-4.14	-3.29	-4.16	-6.47	-5.22	-0.70	1.76	2.77
5 min	-0.18	0.47	-5.80	-5.12	-4.13	7.55	-2.53	-6.48	-7.16	-5.70	-4.47
Panel C: WV with wick truncation ($C_c = 2$)											
Interval	$H = 0$		Gradual jump			Flash crash			Gradual jump with an intermittent flash crash		
	No jump	Jumps	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25	$\beta = 0.45$	0.35	0.25
5 s	-0.60	-0.65	-1.02	-1.01	-0.89	-1.19	-1.22	-1.12	-1.13	-1.07	-0.99
10 s	-0.61	-0.54	-1.30	-1.12	-1.08	-1.69	-1.57	-1.44	-1.32	-1.37	-1.27
30 s	-0.58	-0.52	-2.07	-1.79	-1.53	-2.82	-2.55	-2.14	-2.16	-1.87	-1.79
1 min	-0.65	-0.60	-2.93	-2.62	-2.06	-3.21	-3.04	-2.99	-3.14	-2.60	-1.82
2 min	-0.74	-0.70	-4.00	-3.33	-2.64	-6.37	-5.59	-4.42	-4.71	-3.73	-2.59
3 min	-0.97	-0.95	-4.82	-4.22	-3.42	-7.58	-6.61	-5.29	-3.21	-3.73	-3.84
5 min	-1.44	-0.66	-5.89	-5.22	-4.34	-6.62	-7.68	-6.67	-7.23	-5.78	-4.54

Relative biases (%) of WV constructed from 5, 10, 30, 60, 120, 180, and 300-second candlesticks for 3000 days. The OHLC for each candlestick interval are obtained from the Euler discretization with a mesh size of 10^{-5} seconds. Panel B reports the relative biases of the wick-truncated WV with $C_c = (3, 2)$. The DGP follows the Heston model in Eq. (38), and we follow the persistent noise model of Andersen et al. (2023) to simulate the three different patterns of short-lived explosive trends.

- **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 $g_{fc}^{(1)}$ in Eq. (27), $(\tau, \bar{\tau}, \bar{\tau}) = (0.41, 0.49, 0.57)$, and $c_{1,2}^{(1)} = 1$.⁹
- **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 $g_{gj}^{(1)}$ in Eq. (26), and $(\tau_1, \bar{\tau}_1) = (0.5, 0.65)$. $\eta_{\tau_2} = 0.75\%$, $g^{(2)}$ takes the form $g_{fc}^{(2)}$ in Eq. (27), $(\tau_2, \bar{\tau}_2, \bar{\tau}_2) = (0.55, 0.59, 0.63)$, and $c_{1,2}^{(2)} = 1$.

For each scenario, we consider three different choices of parameter $\beta = \beta_{1,2} \in \{0.45, 0.35, 0.25\}$, which controls the steepness of short-lived explosive trends. 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 price, and corresponds to a less sticky expectation of market participants. Fig. 3 shows examples of simulated (close) prices e^X , efficient (in blue) and observed (in black), for all three scenarios with $\beta = \beta_{1,2} = 0.45$.

In this section, we first follow Li et al. (2024) and Bollerslev et al. (2024) to examine the unbiasedness of WV in “continuous time”: We simulate the sample paths based on a Euler scheme with mesh size being $10 \mu s$ (10^{-5} s) for 3000 days. Intraday candlesticks are constructed at intervals of 5, 10, and 30 s, as well as 1, 2, 3, and 5 min. Next, we evaluate the finite-sample performance of WV: Over a total of 10 000 days, we obtain OHLC for each interval of 1, 2, 3, 5, and 10 min with an Euler mesh size of 0.01 s. Finally, we examine the finite-sample performance of the Hausman test proposed in Section 3.3.

⁹ For flash crashes simulated with $g^{(i)} = g_{fc}^{(i)}$, we stale the observation on $\bar{\tau}_i$ to avoid an unnecessary large “jump” on $\bar{\tau}_i$. For example, when $(\tau, \bar{\tau}, \bar{\tau}) = (0.41, 0.49, 0.57) = (9594, 11466, 13338)$ s, we truncate the all H increments between 11 465 and 11 467 s.

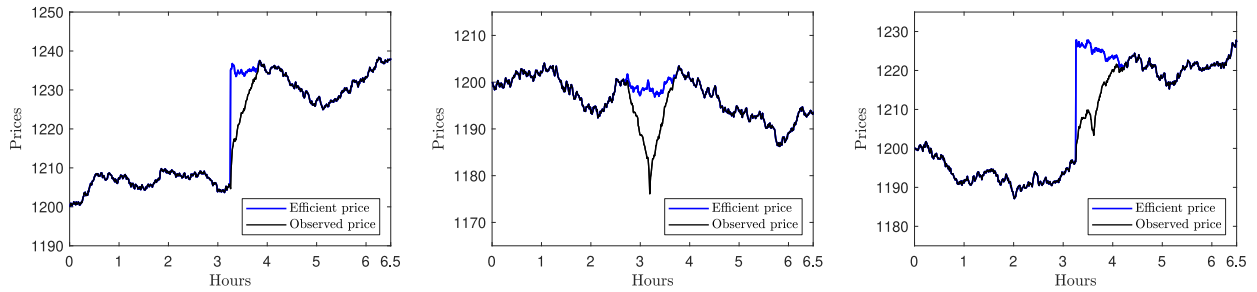


Fig. 3. Simulated close prices e^X , efficient (in blue) and observed (in black), with (i) a gradual jump (left), (ii) a flash crash (middle), and (iii) a gradual jump with an intermittent small flash crash (right), respectively. The efficient price X is simulated with the Heston model in Eq. (38). The deviation between efficient and observed prices is simulated with the persistent noise model in Assumption 4, with all parameters listed above ($\beta = \beta_{1,2} = 0.45$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2. Asymptotic unbiasedness

Table 1 reports the relative biases (%) of WV in “continuous time”. In Panel A, we find that the biases are fairly small across all observation schemes in the absence of short-lived explosive trends, which confirms the consistency of our estimator (Theorem 1) and its robustness to discontinuities (Theorem 3). The existence of gradual jumps leads to only small biases of WV constructed from candlesticks with all selected interval lengths, and the biases shrink rapidly 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 WV in the presence of short-lived explosive trends, which corroborates Theorem 4.¹⁰

We observe that flash crashes introduce a positive bias in the WV estimator without wick truncation when the intervals are relatively large, and the bias becomes more pronounced when the V-shape is steeper, i.e., for smaller values of β , as discussed in Section 3.1. For example, the relative bias of 5-min WV is 42.96% in the presence of a V-shaped flash crash with $\beta = 0.25$. Panel B and C of Table 1 report the relative biases of the wick-truncated WV in Eq. (19), with the truncation parameters $\varpi = 0.49$ and ζ determined with a data-adaptive method of Andersen et al. (2023). Specifically, we let

$$\zeta = C_\zeta \sqrt{\text{MedRV}_{t,n}}, \quad (41)$$

where MedRV is the median RV estimator of Andersen et al. (2012):

$$\text{MedRV}_{t,n} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{n}{n-2} \right) \sum_{i=2}^{n-1} \text{median}(|r_{i-1}|, |r_i|, |r_{i+1}|)^2, \quad (42)$$

and we consider both $C_\zeta = 3$ and 2. Andersen et al. (2023) employ $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}) = (4, 4\sqrt{2})$ and $(3, 3\sqrt{2})$ for TRV and WV, respectively, based on the ratio $\sqrt{2}$ between standard deviations of (absolute) differenced and undifferenced i.i.d. Brownian returns. Analogous to their more aggressive choice of $(C_\zeta^{\text{TRV}}, C_\zeta^{\text{DV}}) = (3, 3\sqrt{2})$, our choice $C_\zeta = 2$ for WV is approximately the same quantile (99.7%) of range-return differences from a standard Brownian motion.

As demonstrated in Panels B and C of Table 1, the wick truncation has nearly no impact in the absence of V-shapes. However, in the presence of V-shaped flash crashes, the wick truncation substantially mitigates the pronounced bias in finite samples, which is consistent with the discussion in Section 3.1.

4.3. Finite-sample performance

To evaluate the finite-sample performance of WV, we utilize the simulated candlesticks that mimic real-world observations accessible to retail investors. We consider two competing range-based estimators: the RRV of Christensen and Podolskij (2007), i.e.,

$$\text{RRV}_{t,n} = \frac{1}{4 \ln 2} \sum_{i=1}^n w_i^2, \quad (43)$$

and the OKV in Eq. (33). While both RRV and OKV have smaller asymptotic variance than WV, they are not robust to jumps and short-lived explosive trends. We also consider some return-based estimators with different degrees of robustness as additional

¹⁰ The bias results for several other IV estimators are presented in Online Appendix C.1, which demonstrate that WV exhibit drift-robustness comparable to DV but better than TRV in the presence of short-lived explosive trends.

Table 2
Monte Carlo RMSE results.

Panel A: $H = 0$																					
Interval	No jump							With jumps													
	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV							
1 min	0.48	0.48	0.33	0.31	0.65	0.65	0.79	0.47	0.48	2.53	1.21	6.71	0.65	0.80							
2 min	0.58	0.57	0.43	0.36	0.91	0.91	1.11	0.58	0.57	2.62	1.29	6.80	0.92	1.13							
3 min	0.69	0.67	0.51	0.42	1.11	1.12	1.37	0.69	0.67	2.68	1.37	6.78	1.14	1.40							
5 min	0.86	0.83	0.66	0.53	1.45	1.45	1.76	0.87	0.82	2.79	1.48	6.88	1.49	1.85							
10 min	1.23	1.15	0.92	0.74	2.03	2.04	2.47	1.23	1.14	2.99	1.69	7.10	2.23	2.85							
Panel B: Gradual jump																					
Interval	$\beta = 0.45$							$\beta = 0.35$							$\beta = 0.25$						
	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV
1 min	0.63	0.63	3.16	1.41	8.61	1.22	0.79	0.61	0.61	3.74	1.64	10.32	1.04	0.79	0.57	0.57	5.26	2.28	14.63	0.88	0.78
2 min	0.75	0.75	3.73	1.74	9.75	2.08	1.17	0.72	0.72	4.38	2.05	11.46	1.71	1.20	0.68	0.68	5.95	2.78	15.66	1.39	1.12
3 min	0.87	0.87	4.94	2.24	13.26	2.63	1.49	0.83	0.83	6.09	2.74	16.50	2.16	1.39	0.78	0.78	8.25	3.70	22.54	1.72	1.37
5 min	1.06	1.06	6.61	3.00	17.79	3.78	2.26	1.02	1.02	7.98	3.62	21.60	3.01	1.95	0.98	0.98	10.32	4.68	28.08	2.30	1.82
10 min	1.38	1.36	8.59	4.26	21.62	9.61	6.30	1.34	1.33	9.78	4.87	24.54	6.47	5.52	1.30	1.29	11.94	5.96	30.02	4.47	4.44
Panel C: Flash crash																					
Interval	$\beta = 0.45$							$\beta = 0.35$							$\beta = 0.25$						
	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV
1 min	0.65	0.65	3.17	1.47	8.30	1.90	0.80	0.53	0.65	3.37	1.63	8.58	1.64	0.83	0.46	0.65	3.59	1.85	8.74	1.33	0.84
2 min	0.87	0.91	3.66	2.07	8.73	3.18	1.31	0.86	0.86	3.76	2.43	8.43	2.53	1.32	0.77	0.77	3.81	2.81	7.97	1.93	1.27
3 min	0.90	1.05	4.51	2.68	10.06	4.63	1.58	0.98	0.99	4.61	3.13	9.47	3.72	1.51	0.89	0.89	4.62	3.57	8.69	2.80	1.48
5 min	1.36	1.17	6.09	3.81	12.75	7.39	2.26	1.54	1.19	6.15	4.36	11.55	5.88	2.09	1.15	1.09	5.93	4.76	10.02	4.33	2.03
10 min	1.66	1.59	10.08	6.20	20.81	15.72	4.90	2.65	2.12	9.80	6.72	18.08	12.58	4.34	3.64	1.69	8.85	6.81	14.38	9.12	4.06
Panel D: Gradual jump with an intermittent flash crash																					
Interval	$\beta = 0.45$							$\beta = 0.35$							$\beta = 0.25$						
	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV	WV(3)	WV(2)	RRV	OKV	RV	TRV	DV
1 min	0.65	0.66	3.06	1.38	8.23	1.54	0.82	0.61	0.61	3.60	1.62	9.69	1.40	0.79	0.56	0.57	4.89	2.20	13.26	1.18	0.79
2 min	0.79	0.79	3.69	1.74	9.50	2.84	1.35	0.74	0.74	4.34	2.08	11.09	2.50	1.33	0.67	0.67	5.71	2.76	14.60	2.06	1.20
3 min	0.73	0.79	4.42	2.16	11.38	2.99	1.71	0.70	0.82	5.43	2.68	14.02	2.57	1.49	0.71	0.82	7.34	3.59	19.15	2.02	1.39
5 min	1.12	1.12	6.24	2.91	16.33	4.93	3.36	1.05	1.05	7.57	3.55	19.81	4.32	2.83	0.99	0.99	9.64	4.53	25.35	3.51	2.23
10 min	1.40	1.39	8.61	4.34	21.19	11.61	9.94	1.37	1.35	9.82	5.03	23.87	8.89	10.28	1.34	1.32	11.65	6.01	28.29	6.57	8.58

RMSEs (multiplied by 10^5) of different IV estimators for 10000 days. We follow the persistent noise model of Andersen et al. (2023) to simulate the three different patterns of short-lived explosive trends. All range- and return-based estimators are constructed from the simulated 1, 2, 3, 5, and 10-min candlesticks. The choice of truncation parameters for TRV, DV and WV follows the instructions in Section 4.2, with $C_{\zeta}^{WV} = 3$ and 2 (represented by WV(3) and WV(2), respectively), $C_{\zeta}^{TRV} = 3$, and $C_{\zeta}^{DV} = 3\sqrt{2}$.

benchmarks, such as the standard RV estimator, the TRV estimator of Mancini (2009):

$$RV_{t,n} = \sum_{i=1}^n r_i^2 \quad \text{and} \quad TRV_{t,n} = \sum_{i=1}^n r_i^2 \mathbb{1}_{\{|r_i| \leq \zeta \Delta_n^m\}}, \quad (44)$$

and the DV estimator of Andersen et al. (2023):

$$DV_{t,n} = \frac{1}{2} \sum_{i=2}^n (r_i - r_{i-1})^2 \mathbb{1}_{\{|r_i - r_{i-1}| \leq \zeta \Delta_n^m\}}. \quad (45)$$

The choice of truncation parameters for TRV, DV and WV follows the instructions in Section 4.2, with $C_{\zeta}^{WV} = 3$ and 2, $C_{\zeta}^{TRV} = 3$, and $C_{\zeta}^{DV} = 3\sqrt{2}$.¹¹ For a generic IV estimator $\hat{V}_{t,n}$, its finite-sample performance is evaluated with the root-mean-square error (RMSE):

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\hat{V}_{t,n} - \int_0^t \sigma_t^2 dt \right)^2}, \quad \text{with } M = 10000. \quad (46)$$

In Panel A of Table 2, we present the RMSE results for all selected IV estimators in the absence of short-lived explosive trends. When candlesticks are simulated from a continuous Itô semimartingale, the range-based estimators demonstrate lower finite-sample RMSEs compared to their return-based counterparts, which is consistent with the discussion on efficiency in Remark 2. Among the

¹¹ We also consider alternative parameter choices $(C_{\zeta}^{TRV}, C_{\zeta}^{DV}) = (4, 4\sqrt{2})$ used for comparison in Andersen et al. (2023). 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.3 in Online Appendix C.2, which is consistent with the Monte Carlo results in Andersen et al. (2023).

range-based estimators, the OKV proposed in Section 3.3 achieves the smallest RMSEs across all observation schemes, while the WV trades off some efficiency for increased robustness. In the presence of jumps, the WV estimator, along with the return-based TRV and DV, demonstrates excellent robustness and obtains RMSEs comparable to the no-jump scenario. In contrast, other estimators display significantly inflated RMSEs. The OKV, however, is less affected by jumps than the RRV, as the weighting scheme in Eq. (34) partially offsets the impact of jumps.

Panels B, C, and D in Table 2 report the RMSE results in the presence of gradual jumps and flash crashes. The WV estimators consistently achieve the smallest RMSEs among all selected range- and return-based measures, highlighting its robustness against different extreme market scenarios and across various observation frequencies. For instance, the WV estimators maintain small RMSEs (less than 4×10^5) even when constructed from 10-min candlesticks, a frequency at which other estimators become much less reliable, which underscores its effectiveness as a practical and robust estimation technique. In contrast, we observe that some return-based estimators originally designed to be robust to jumps, such as TRV (as well as BV and MedRV in Table C.3 in Online Appendix C.2), exhibit larger RMSEs than DV. This difference becomes more pronounced as the explosive trends occur more gradually over time rather than abruptly, i.e., with larger values of β , which causes these movements to deviate further from the sharp, sudden price changes (become less jump-like) that these estimators are specifically designed to handle.¹²

We proceed to evaluate the finite-sample performance of the Hausman test in Eq. (35). Following the simulation setup outlined in Section 4.1 for the continuous Itô semimartingale under the null, we consider three alternative scenarios: (i) jumps of moderate or large sizes, (ii) a single gradual jump, and (iii) a single V-shaped flash crash. Specifically, jumps are simulated with the compound Poisson process in Eq. (39) with $\lambda = 1$ and two different standard deviations for the random jump size Z_i : $\varsigma = 0.9\%$ for moderate jumps and $\varsigma = 1.8\%$ for larger jumps. For the gradual jump and flash crash, their magnitudes remain deterministic with the parameter choices in Section 4.1, but their locations are uniformly randomized within the interval $[0, 1]$.

Table 3 reports the finite-sample size and power (at 5% nominal level) of the Hausman test constructed from intraday candlesticks of 1, 2, 3, 5, and 10 min. In Panel A, where the log-price follows a continuous Itô semimartingale, the empirical rejection rates of the Hausman test are close to the nominal 5% level, though it tends to be slightly oversized. In Panels B, C, and D, which include alternative scenarios involving jumps and short-lived explosive trends, the Hausman test demonstrates strong finite-sample power, particularly in the presence of gradual jumps and flash crashes. We note that the Hausman test achieves the highest rejection rates in the presence of some sharp explosive trends simulated with $\beta = 0.25$, which corroborates the theoretical results in Corollary 2. As discussed following that corollary in Section 3.3, while the Hausman test is not consistent in the presence of explosive trends with relatively mild explosion rates, our simulation results with $\beta = 0.45$ and $\beta = 0.35$ reveal that the test still exhibits non-trivial power in finite samples.

In Table 3, we also present the empirical rejection rates of three additional nonparametric tests, each with a distinct null hypothesis. The first is the high-frequency jump test (referred to as BNS in Table 3) by Barndorff-Nielsen and Shephard (2006), which relies on the relative magnitudes of the return-based BV and RV estimators. The second is the differenced return-based test (ALTZ) by Andersen et al. (2023), which is designed specifically to test for the presence of short-lived explosive trends.¹³ The third is a nonparametric test for non-negligible drift (KRZ) proposed by Kolokolov et al. (2025), which compares the standard RV and the average of DV estimators constructed from m -lag differenced returns, with the critical regions obtained by wild bootstrap.¹⁴ Our results indicate that the BNS test suffers from different levels of size distortion when gradual jumps or flash crashes are present. Similarly, the ALTZ test exhibits size distortion in the presence of relatively large jumps. These limitations highlight the challenge of separately testing for “discontinuous” and “continuous” extreme price movements. Moreover, both the ALTZ and KRZ tests exhibit weaker power than the Hausman test in identifying short-lived explosive trends, particularly in flash crash scenarios with smaller β , which suggests the difficulty of identifying extreme price movements that both occur and recover rapidly when relying solely on return information over “not-too-finely” sampled intervals. The new Hausman test, based on readily available summarized information, offers an easy-to-implement and reliable alternative for identifying intraday extreme price movements.

5. Empirical analysis

In this section, we use the WV 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.

¹² When ultra-high-frequency data is available, one can construct some noise-corrected return-based estimators, e.g., the pre-averaged RV and BV of Jacod et al. (2009) and Podolskij and Vetter (2009), from all tick-level data. Online Appendix C.3 presents a comparison between these pre-averaged estimators and WV in the presence of market microstructure noise, which shows that the WV constructed from summary information over sparsely sampled intervals can achieve comparable performance to the pre-averaged estimators based on more granular observations.

¹³ We construct the test statistic of Andersen et al. (2023) based on second-order return autocovariances, i.e., $T_1^n(2)$, in line with their empirical applications. The bandwidth for the local DV estimators is selected as $k_n = \lceil \sqrt{n/2} \rceil$.

¹⁴ Kolokolov et al. (2025) show that the average of DV estimators is BUMVU for every fixed m . We follow their recommendation in Section 7 to use the maximum of multiple test statistics across all $m \in \{1, 2, \dots, 21\}$, where the data-driven critical regions obtained by wild bootstrap ensures robustness against multiple testing issue. The MATLAB code was kindly provided by Aleksey Kolokolov.

Table 3

Monte Carlo size and power results (%).

Monte Carlo size and power results (%).

Panel A: No jump & No explosive trends

Interval	Tests			
Interval	Haus.	BNS	ALTZ	KRZ
1 min	6.53	5.21	5.09	4.79
2 min	6.09	5.18	5.08	5.17
3 min	6.02	4.99	5.42	5.25
5 min	5.88	5.55	5.71	4.86
10 min	6.14	5.71	6.18	5.07

Panel B: Jumps

Interval	Moderate jumps				Large jumps			
	Haus.	BNS	ALTZ	KRZ	Haus.	BNS	ALTZ	KRZ
1 min	74.22	82.32	4.42	5.46	87.05	91.74	9.97	5.20
2 min	68.08	77.30	5.16	5.03	84.44	88.93	10.36	5.28
3 min	65.45	73.10	6.18	4.67	82.42	87.26	11.62	4.94
5 min	60.72	67.11	5.77	4.48	80.67	84.03	12.84	4.79
10 min	53.18	54.91	7.41	4.91	76.55	77.13	13.31	5.03

Panel C: Gradual jump

Interval	$\beta = 0.45$				$\beta = 0.35$				$\beta = 0.25$			
	Haus.	BNS	ALTZ	KRZ	Haus.	BNS	ALTZ	KRZ	Haus.	BNS	ALTZ	KRZ
1 min	79.20	2.47	94.31	91.03	93.81	17.59	91.61	79.62	99.74	56.01	83.06	60.52
2 min	85.82	2.52	94.05	87.35	94.96	16.84	87.35	73.36	99.47	53.68	73.26	53.07
3 min	88.98	2.72	92.53	82.19	95.65	16.88	83.49	67.15	99.43	51.22	67.46	48.83
5 min	92.12	3.28	88.23	70.50	96.91	15.94	77.58	56.28	99.41	48.72	61.27	41.16
10 min	93.93	3.54	70.25	38.81	96.72	15.69	59.64	33.25	98.71	44.14	48.84	26.92

Panel D: Flash crash

Interval	$\beta = 0.45$				$\beta = 0.35$				$\beta = 0.25$			
	Haus.	BNS	ALTZ	KRZ	Haus.	BNS	ALTZ	KRZ	Haus.	BNS	ALTZ	KRZ
1 min	90.08	0.14	81.28	75.76	97.82	0.29	60.44	54.44	99.59	0.57	30.83	33.89
2 min	89.64	0.24	72.31	65.75	96.96	0.38	47.04	46.74	99.01	1.09	21.97	27.03
3 min	83.61	0.16	60.79	56.06	93.37	0.58	35.06	38.54	96.77	1.24	16.47	22.23
5 min	72.71	0.26	35.74	39.33	84.61	0.57	19.27	25.88	91.25	1.48	9.96	15.87
10 min	55.76	0.78	1.15	9.67	69.37	1.28	1.17	7.39	79.72	2.29	1.45	5.27

Finite-sample size and power (%) of different nonparametric tests for jumps and/or short-lived explosive trends at 5% nominal level: the Hausman test in Eq. (35), BNS (Barndorff-Nielsen and Shephard, 2006), ALTZ (Andersen et al., 2023), and KRZ (Kolokolov et al., 2025). We follow the persistent noise model of Andersen et al. (2023) to simulate the three different patterns of short-lived explosive trends. All range- and return-based test statistics are constructed from the simulated 1, 2, 3, 5, and 10-min candlesticks. The wick truncation follows the instructions in Section 4.2 with $C_c = 3$.

5.1. Data

To replicate the OHLCs readily available to general investors, we construct 5-min candlesticks with high-frequency transaction records of SPY obtained from the daily Trade and Quote (TAQ) dataset for the period ranging from January 2, 2014 to December 29, 2023. The tick-level transactions are timestamped in milliseconds until mid-2015 and in microseconds since then.¹⁵ As is standard in empirical research with TAQ data, we use the standard data filters as in Barndorff-Nielsen et al. (2009) to eliminate data errors, remove all transactions in the original record that are later corrected, canceled or otherwise invalidated. In addition, we remove all trading days with an early market closure, and restrict our sample to transactions between 9:30:00 – 16:00:00 Eastern Time (ET). The final sample comprises of $T = 2493$ days.

5.2. Heterogeneous autoregressive (HAR) model

The HAR model of Corsi (2009) is designed to parsimoniously capture the dependence structure of return volatility across different horizons, which to a great extent approximates its well-documented long memory properties. Renowned for its consistent and remarkable predictive performance, the HAR model serves as the predominant benchmark in volatility forecasting research. In this section, $\hat{V}_t^*$ denotes the benchmark IV estimator at day t as the forecasting target, and $\hat{V}_t$ represents a generic IV estimator. We

¹⁵ 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.

define the following moving averages of $\hat{V}_t$ as:

$$\hat{V}_{w,t} = \frac{1}{5} \sum_{i=1}^5 \hat{V}_{t-i+1} \quad \text{and} \quad \hat{V}_{m,t} = \frac{1}{22} \sum_{i=1}^{22} \hat{V}_{t-i+1}, \quad (47)$$

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

$$\hat{V}_t^* = \omega + \beta_d \hat{V}_{t-1} + \beta_w \hat{V}_{w,t-1} + \beta_m \hat{V}_{m,t-1} + \varepsilon_t, \quad (48)$$

which can be easily estimated via ordinary least squares (OLS). As evidenced by numerous empirical studies, incorporating a volatility measure of the continuous component on the right-hand side (RHS) can effectively improve the predictive power of the HAR model for the left-hand side (LHS) target variable $\hat{V}_t^*$ (Andersen et al., 2007).

The inherent trade-off between robustness and efficiency, as discussed in Section 3.3, motivates the use of hybrid estimators (also known as pre-test estimators) that combine two realized measures. For example, we consider a mixture of the OKV estimator in Eq. (33) for days without extreme price movements, and the WV estimator for days when such phenomena are present:

$$HWV_t = OKV_t \mathbb{1}_{\{T_t \leq c_{1-a}\}} + \overline{WV}_t \mathbb{1}_{\{T_t > c_{1-a}\}}, \quad (49)$$

where the hybrid WV (HWV) estimator combines the efficiency of OKV and the robustness of WV based on the results of the Hausman test in Eq. (35). Similarly, we consider two other hybrid combinations of IV estimators: HBV, which combines RV and BV with the BNS test of Barndorff-Nielsen and Shephard (2006), and HDV, a mixture of TRV and DV based on the ALTZ test of Andersen et al. (2023).

In addition to the implementation of better volatility measures in the standard HAR model, we also consider two important extensions of the original HAR-RV model for comparative study. The first extension is the quarticity expanded HAR (HARQ) model proposed by Bollerslev et al. (2016). Motivated by the fact that the persistence of RV is influenced by temporal variations in its measurement errors, the HARQ-RV model allows for a time-varying coefficient for the previous day's RV on the RHS. The coefficient depends on the realized quarticity (RQ) that captures the heteroskedasticity in the error¹⁶:

$$\hat{V}_t^* = \omega + \left(\beta_d + \beta_q \sqrt{RQ_{t-1}} \right) RV_{t-1} + \beta_w RV_{w,t-1} + \beta_m RV_{m,t-1} + \varepsilon_t, \quad (50)$$

where $RQ_t = 3^{-1} n \sum_{i=1}^n r_i^4$. A similar extension can be applied to the HAR-WV model by adopting the HARQ structure, replacing RV and RQ with our WV and WQ estimators, respectively:

$$\hat{V}_t^* = \omega + \left(\beta_d + \beta_q \sqrt{WQ_{t-1}} \right) \overline{WV}_{t-1} + \beta_w \overline{WV}_{w,t-1} + \beta_m \overline{WV}_{m,t-1} + \varepsilon_t. \quad (51)$$

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

$$\hat{V}_t^* = \omega + \beta_d^- RV_{t-1}^- + \beta_d^+ RV_{t-1}^+ + \beta_w RV_{w,t-1} + \beta_m RV_{m,t-1} + \varepsilon_t, \quad (52)$$

where the semivariances are given by

$$RV_t^- = \sum_{i=1}^n r_i^2 \mathbb{1}_{\{r_i < 0\}} \quad \text{and} \quad RV_t^+ = \sum_{i=1}^n r_i^2 \mathbb{1}_{\{r_i > 0\}}. \quad (53)$$

The intuition that “good” and “bad” volatilities are not created equal motivates the decomposition of the original RV into separate up and downside semivariances. The empirical results in Patton and Sheppard (2015) indicate that this decomposition leads to more accurate volatility forecasts, with the “bad” volatility predominantly driving the short-run changes in the future. Similarly, we consider the SHAR-WV model with the same structure:

$$\hat{V}_t^* = \omega + \beta_d^- \overline{WV}_{t-1}^- + \beta_d^+ \overline{WV}_{t-1}^+ + \beta_w \overline{WV}_{w,t-1} + \beta_m \overline{WV}_{m,t-1} + \varepsilon_t, \quad (54)$$

with the wick-based semivariances:

$$\begin{aligned} \overline{WV}_t^- &= \frac{1}{A_2} \sum_{i=1}^n (w_i - |r_i|)^2 \mathbb{1}_{\{|w_i - |r_i|| \leq \zeta A_n^w\}} \mathbb{1}_{\{r_i < 0\}}, \\ \overline{WV}_t^+ &= \frac{1}{A_2} \sum_{i=1}^n (w_i - |r_i|)^2 \mathbb{1}_{\{|w_i - |r_i|| \leq \zeta A_n^w\}} \mathbb{1}_{\{r_i \geq 0\}}. \end{aligned} \quad (55)$$

¹⁶ 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.

5.3. Empirical results

In this section, we estimate the standard HAR model in Eq. (48) with various return- and range-based realized volatilities as predictors on the RHS. Specifically, we include

- (i) Return-based estimators: RV, BV, MedRV, TRV, and DV;
- (ii) Range-based estimators: RRV, OKV, and WV.
- (iii) Hybrid estimators in the form of Eq. (49): HBV, HDV, and HWV.

All estimators are constructed from intraday OHLCs contained in 5-min candlesticks. In addition to the standard HAR model, we estimate the HARQ and SHAR models with both RV and WV, as specified in Eqs. (50), (51), (52), and (54), respectively. With an initial in-sample period of the first 252 trading days, we generate one-day-ahead out-of-sample forecasts for three realized volatilities with different levels of robustness: RV, TRV, and WV. The in-sample estimation and out-of-sample forecasting procedures are repeated with both rolling windows (RW) and expanding windows (EW).

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:

$$\text{MSE} = \frac{1}{M} \sum_{t=1}^M (\hat{V}_t^* - \tilde{V}_t^*)^2 \quad \text{and} \quad \text{QLIKE} = \frac{1}{M} \sum_{t=1}^M \left(\frac{\hat{V}_t^*}{\tilde{V}_t^*} - \ln \left(\frac{\hat{V}_t^*}{\tilde{V}_t^*} \right) - 1 \right), \quad (56)$$

where $\hat{V}_t^*$ and $\tilde{V}_t^*$ denote the ex-post estimate and the forecast of the target volatility measure on day t , respectively, and M represents the total number of out-of-sample days.

Table 4 reports the MSE and QLIKE results for one-day-ahead out-of-sample forecasts of RV, TRV, and WV, respectively. Compared with the simple HAR models augmented with return-based volatility estimators, range-based estimators generally demonstrate superior predictive power, which is consistent with the established view since Parkinson (1980) and Garman and Klass (1980) that ranges embed richer information. It is notable that the HAR-WV model achieves the smallest values of both loss functions in all cases. Given that the symmetric MSE function penalizes outliers heavily, it is particularly sensitive to excessively misinformative forecasts. The MSE results in Table 4 indicate that the HAR-WV model can effectively reduce the occurrence of extremely inaccurate forecasts in both the left and right tails.

Among the standard HAR models augmented with hybrid estimators, HAR-HBV and HAR-HWV achieve slightly lower MSE and QLIKE results than the better individual performers in their respective combinations when re-estimated in rolling windows. However, this improvement becomes less pronounced when the models are estimated with expanding in-sample windows.

For the modified HAR models, both the HARQ-RV and SHAR-RV models can achieve smaller MSE and QLIKE results than the original HAR-RV. This demonstrates that the consideration of either the measurement errors in RV or the volatility asymmetry helps to exploit concealed information due to aggregation, and leads to more accurate forecasts. Furthermore, the HARQ-WV (resp. SHAR-WV) extension outperforms HARQ-RV (resp. SHAR-RV) with smaller MSE and QLIKE values across all cases.¹⁷

Furthermore, to assess whether the models are statistically significantly different in terms of forecasting performance, we employ the model confidence set (MCS) approach of Hansen et al. (2011), which involves a sequence of tests for equal predictive ability and retains the best model in the MCS at a given confidence level. We consider both MSE and QLIKE results in Table 4, and estimate the bootstrap distributions of the range test statistics in Section 3.1.2 of Hansen et al. (2011), each with 5000 replications and a block size of $\lfloor \sqrt[3]{M} \rfloor = 13$ days.¹⁸

In Table 4, we further report the MCS results (with asterisks) for our selection of volatility models at two confidence levels, 75% and 90%, following Hansen et al. (2003, 2011).¹⁹ The MSE-based MCS tends to be more conservative, which is consistent with the empirical results of Hansen et al. (2003), and eliminates no models at either confidence level in our application. By contrast, the QLIKE-based MCS is more selective: It excludes both HAR-RV and HAR-BV under the rolling-window scheme, and excludes nearly all competitors of HARQ-WV when evaluating the expanding-window forecasts, highlighting its significantly superior forecasting performance among all selected models. Additional comparisons within selected subsets of models (in Online Appendix C.5) lead to similar findings: a simple HAR model that uses in-sample WV or HWV estimates significantly outperforms the same model based on other return- or range-based estimators, which provides further evidence that the summary information embedded in candlesticks can enhance volatility forecasts.

¹⁷ We also consider noise-corrected IV estimators constructed from all tick-level transaction data, which are accessible to practitioners with superior data availability. A comparison of volatility forecasting based on range-based estimators (with 5-min candlesticks) and noise-corrected RV estimators (with tick-level data) demonstrates comparable performance, which is presented in Online Appendix C.4.

¹⁸ The block size aligns with the optimal choice suggested by Hall et al. (1995) for bootstrapping dependent data. Alternatively, it can be set based on the maximum lag (selected via information criteria) of AR models fitted to the loss series for all selected models. The MCS results with such data-driven bootstrap block sizes lead to similar conclusions to those from Table 4, and are thus omitted here.

¹⁹ The MCS p -values are reported in Table C.7 in Online Appendix C.5.

Table 4
Daily out-of-sample 5-min HAR volatility forecasts.

	RV		TRV		WV	
	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE
Panel A: RW forecasts						
HAR-RV	3.03**	0.36*	1.82**	0.37*	1.11**	0.26
HAR-BV	3.39**	0.37*	1.98**	0.38*	1.18**	0.26
HAR-MedRV	2.99**	0.36*	1.74**	0.37**	1.03**	0.26*
HAR-TRV	2.52**	0.35**	1.55**	0.36**	0.95**	0.25**
HAR-DV	2.55**	0.35**	1.66**	0.37**	1.03**	0.25**
HAR-RRV	2.38**	0.35**	1.47**	0.36**	0.90**	0.25**
HAR-OKV	2.19**	0.35**	1.35**	0.36**	0.83**	0.25**
HAR-WV	2.14**	0.34**	1.32**	0.36**	0.83**	0.24**
HAR-HBV	3.01**	0.36**	1.81**	0.37**	1.11**	0.26*
HAR-HDV	2.52**	0.35**	1.55**	0.36**	0.95**	0.25**
HAR-HWV	2.12**	0.34**	1.30**	0.36**	0.82**	0.24**
SHAR-RV	2.89**	0.36*	1.81**	0.37**	1.11**	0.26**
SHAR-WV	2.37**	0.36*	1.41**	0.37**	0.82**	0.25**
HARQ-RV	2.11**	0.33**	1.37**	0.35**	0.98**	0.25**
HARQ-WV	1.94**	0.33**	1.35**	0.35**	0.96**	0.25**
Panel B: EW forecasts						
HAR-RV	2.19**	0.37	1.33**	0.39	0.82**	0.27
HAR-BV	2.46**	0.38	1.45**	0.41	0.87**	0.28
HAR-MedRV	2.16**	0.36	1.28**	0.39	0.76**	0.27
HAR-TRV	1.75**	0.31	1.06**	0.33	0.65**	0.23
HAR-DV	1.87**	0.32	1.22**	0.33	0.76**	0.23
HAR-RRV	1.76**	0.33	1.08**	0.35	0.66**	0.24
HAR-OKV	1.64**	0.32	1.00**	0.33	0.61**	0.23
HAR-WV	1.56**	0.30	0.95**	0.31	0.58**	0.21
HAR-HBV	2.18**	0.37	1.32**	0.39	0.81**	0.27
HAR-HDV	1.75**	0.31	1.06**	0.33	0.65**	0.22
HAR-HWV	1.60**	0.30	0.97**	0.31	0.60**	0.21
SHAR-RV	1.98**	0.35	1.26**	0.38	0.78**	0.26
SHAR-WV	1.55**	0.32	0.95**	0.32	0.57**	0.21
HARQ-RV	1.79**	0.28**	1.21**	0.29**	0.78**	0.21
HARQ-WV	1.53**	0.28**	0.98**	0.29**	0.59**	0.19**

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 with both rolling windows and expanding windows, respectively. All return- and range-based estimators, along with the associated test statistics, are constructed from intraday OHLCs contained in 5-min candlesticks. The choice of truncation parameters for TRV, DV and WV follows the instructions in Section 4.2 with $(C_{\zeta}^{WV}, C_{\zeta}^{TRV}, C_{\zeta}^{DV}) = (3, 3, 3\sqrt{2})$.

* Model confidence sets (MCS) at confidence levels 75%.

** Model confidence sets (MCS) at confidence levels 90%.

5.4. Impact of extreme price movements

To further investigate the sources of forecast error reduction in the HAR-WV model, we split the out-of-sample period into two complementary subsets of days: (i) days that follow extreme price movements identified on preceding days (“EPM + 1” days), and (ii) all remaining days. Extreme price movements are characterized by either jumps or episodes of explosive trends. When such events occur on day t , they tend to inflate estimates of non-robust volatility measures such as RV on day t . Consequently, the HAR model with non-robust volatility estimators on the RHS is likely to produce overestimated volatility forecasts for day $t + 1$.

Table 5 presents the MSE and QLIKE results for all standard and modified HAR models across the out-of-sample subsets classified based on this criterion, where the existence of extreme price movements is identified with the Hausman test in Eq. (35). Firstly, we observe a substantial reduction in MSEs across all models on days without price jumps and explosive trends on preceding days. The standard HAR models (with various RHS variables) exhibit relatively comparable MSE and QLIKE results, which indicates nearly uniform performance across all HAR models when applied to periods of moderate market behavior. However, the existence of extreme price movements tends to distort the estimation of dependence structure in volatility, and therefore leads to uniformly deteriorated forecasts across these models. Secondly, among the three volatility estimators used as one-day-ahead forecasting targets on the LHS, we find that the forecasts for WV exhibit superior accuracy in the presence of extreme price movements with all chosen RHS variables. In contrast, RV predictions are significantly more susceptible to the influence of extreme price movements, which leads to substantially larger forecast errors. Thirdly, for each target variable on the LHS, the HAR-WV model demonstrates the least vulnerability to extreme events among the standard HAR models and consistently delivers the most accurate one-day-ahead forecasts. The enhanced predictability from robustness is further evidenced by the dramatically inflated MSE values observed in standard HAR models that rely on non-robust RV, which highlight how extreme price movements can severely bias their forecasts.

Table 5

Daily out-of-sample 5-min HAR volatility forecasts on different subsets of days.

Trading days	RV				TRV				WV			
	EPM + 1		Other		EPM + 1		Other		EPM + 1		Other	
	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE	MSE	QLIKE
Panel A: RW forecasts												
HAR-RV	8.59	0.42	2.11	0.35	4.82	0.45	1.33	0.36	2.33	0.30	0.90	0.26
HAR-BV	11.20	0.41	2.10	0.36	5.82	0.44	1.34	0.37	2.78	0.29	0.91	0.26
HAR-MedRV	9.04	0.39	2.00	0.36	4.75	0.42	1.25	0.37	2.19	0.28	0.84	0.25
HAR-TRV	5.29	0.36	2.06	0.35	3.05	0.39	1.30	0.36	1.35	0.26	0.88	0.24
HAR-DV	4.51	0.37	2.23	0.35	2.72	0.39	1.48	0.36	1.19	0.25	1.00	0.25
HAR-RRV	4.59	0.39	2.01	0.35	2.77	0.43	1.25	0.35	1.29	0.29	0.84	0.24
HAR-OKV	3.48	0.38	1.97	0.34	2.17	0.42	1.22	0.35	0.98	0.28	0.81	0.24
HAR-WV	2.53	0.33	2.08	0.35	1.63	0.35	1.31	0.36	0.71	0.23	0.85	0.25
HAR-HBV	8.70	0.40	2.07	0.35	4.81	0.43	1.31	0.36	2.35	0.29	0.89	0.25
HAR-HDV	5.28	0.36	2.06	0.35	3.05	0.39	1.30	0.36	1.35	0.26	0.88	0.24
HAR-HWV	2.46	0.30	2.06	0.36	1.58	0.32	1.27	0.37	0.69	0.21	0.84	0.25
SHAR-RV	6.74	0.44	2.26	0.35	4.40	0.46	1.38	0.36	2.23	0.32	0.93	0.25
SHAR-WV	3.13	0.37	2.36	0.37	2.23	0.40	1.51	0.37	0.90	0.25	0.94	0.25
HARQ-RV	1.17	0.40	2.26	0.33	0.64	0.44	1.47	0.34	0.37	0.30	1.08	0.25
HARQ-WV	1.28	0.31	2.05	0.33	0.76	0.37	1.45	0.34	0.48	0.25	1.04	0.25
Panel B: EW forecasts												
HAR-RV	5.65	0.44	1.62	0.36	3.04	0.48	1.05	0.38	1.51	0.32	0.70	0.27
HAR-BV	7.50	0.45	1.63	0.37	3.87	0.48	1.06	0.40	1.89	0.33	0.71	0.28
HAR-MedRV	5.97	0.42	1.53	0.35	3.09	0.45	0.98	0.38	1.47	0.30	0.64	0.26
HAR-TRV	3.10	0.35	1.52	0.30	1.72	0.38	0.95	0.32	0.78	0.25	0.63	0.22
HAR-DV	2.57	0.36	1.75	0.31	1.51	0.38	1.17	0.33	0.70	0.25	0.77	0.22
HAR-RRV	3.04	0.39	1.55	0.32	1.74	0.42	0.97	0.34	0.81	0.28	0.64	0.23
HAR-OKV	2.36	0.38	1.52	0.31	1.39	0.41	0.94	0.32	0.62	0.27	0.61	0.22
HAR-WV	1.89	0.31	1.50	0.30	1.11	0.29	0.92	0.31	0.48	0.21	0.60	0.22
HAR-HBV	5.65	0.43	1.61	0.36	3.02	0.47	1.04	0.38	1.50	0.31	0.70	0.27
HAR-HDV	3.12	0.35	1.52	0.30	1.74	0.37	0.95	0.32	0.79	0.25	0.63	0.22
HAR-HWV	1.84	0.27	1.56	0.30	1.11	0.29	0.95	0.32	0.46	0.19	0.62	0.21
SHAR-RV	3.79	0.49	1.68	0.33	2.27	0.52	1.09	0.35	1.17	0.36	0.71	0.25
SHAR-WV	1.73	0.34	1.57	0.31	1.07	0.36	0.98	0.32	0.46	0.23	0.61	0.21
HARQ-RV	1.34	0.38	1.87	0.26	0.74	0.41	1.29	0.27	0.34	0.28	0.85	0.19
HARQ-WV	1.55	0.30	1.53	0.26	0.92	0.31	0.99	0.27	0.37	0.21	0.63	0.19

MSE ($\times 10^8$) and QLIKE of daily out-of-sample volatility forecasts for the SPDR S&P 500 ETF Trust (SPY) on (i) days that follow extreme price movements identified on preceding days (“EPM+1” days), and (ii) all remaining days. The existence of jumps or short-lived explosive trends is identified with the Hausman test in Eq. (35). The HAR model is re-estimated via OLS with both rolling windows and expanding windows, respectively. All return- and range-based estimators are constructed from intraday OHLs contained in 5-min candlesticks. The choice of truncation parameters for TRV, DV and WV follows the instructions in Section 4.2 with $(C_{\epsilon}^{WV}, C_{\epsilon}^{TRV}, C_{\epsilon}^{DV}) = (3, 3, 3\sqrt{2})$.

Additionally, we find that both HARQ models exhibit relatively more robust performance on “EPM+1” days. This improvement, however, is largely attributable to the application of the “insanity filter” (Swanson and White, 1997), which replaces severely biased forecasts with their corresponding in-sample means. By mitigating extreme over-predictions, the filter reverses the relative performance of HARQ-RV and HARQ-WV in terms of MSE, with the HARQ-RV model obtaining apparently lower MSEs on “EPM+1” days. This observation also indirectly highlights the utility of the filter in standard HAR models, particularly when the RHS volatility measures are not robust to extreme price movements. However, the QLIKE function, which penalizes under-predictions more heavily than MSE and is less affected by the filter, continues to highlight the superior performance of HARQ-WV on days following extreme price movements.

To gain deeper insights into how extreme price movements impact the accuracy of volatility forecasts, we focus on the days surrounding these events. Specifically, we evaluate the performance of one-day-ahead RV and WV forecasts with three standard HAR models over a period spanning five days before and five days after the identified extreme price movements. To isolate the effect of individual events, we exclude any days that experienced other extreme price movements within the preceding five days.

Fig. 4 illustrates the MSE results for the one-day-ahead rolling-window RV (left) and WV (right) forecasts across the three models. On days with identified extreme price movements (“EPM Days” with $t = 0$), we observe a clear spike in MSEs for all models. This increase is due to biases in the IV proxies as forecasting targets. Since WV exhibits much lower bias than RV in the presence of extreme price movements, the spike in MSEs on $t = 0$ is much smaller in the right panel. More importantly, a substantial spike in MSEs is observed on “EPM+1” days for both HAR-RV and HAR-TRV models in both panels. As previously discussed, the HAR-WV model outperforms both HAR-RV and HAR-TRV on “EPM+1” days due to its robustness to extreme price movements. Moreover, the MSEs of all models decay as we move further away from the identified extreme events, and return to baseline levels when $t = 5$. This analysis provides further insights into the impact of extreme price movements on volatility forecasting and highlights the main sources of the forecasting accuracy gain of the HAR-WV model.

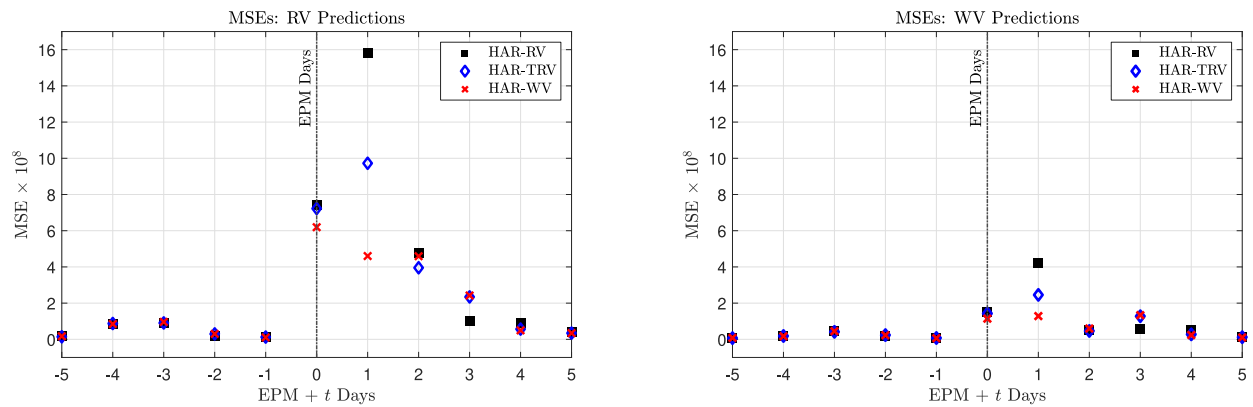


Fig. 4. MSE results for the one-day-ahead RV (left) and WV (right) forecasts across three standard HAR models. The existence of jumps or short-lived explosive trends is identified with the Hausman test in Eq. (35). The HAR model is re-estimated via OLS with rolling windows. All return- and range-based estimators are constructed from intraday OHLCs contained in 5-min candlesticks.

6. Conclusions

Motivated by the rich information embedded in intraday candlesticks and their wide availability to retail investors, we propose a novel nonparametric candlestick-based estimator for integrated variance (IV), namely the wick-based volatility (WV) estimator. The WV estimator provides robust inference of IV in the presence of short-lived extreme price movements, including jump discontinuities, gradual jumps, and flash crashes. We demonstrate that WV consistently estimates IV with an asymptotic variance approximately four times smaller than that of the differenced-return volatility (DV) estimator introduced by Andersen et al. (2023). Simulation results highlight the reliable finite-sample performance of WV across various practically relevant scenarios. Furthermore, an empirical application to volatility forecasting shows that the HAR-WV model can effectively reduce the occurrence of extremely misleading forecasts due to extreme price movements on preceding days, which leads to superior forecasting accuracy compared to a broad class of estimators in the literature.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the Co-Editor Michael Jansson, the Associate Editor, and two anonymous referees for their constructive comments and suggestions. We also thank Torben Andersen, Tim Bollerslev, Carsten Chong, Kim Christensen, Francis Diebold, Dick van Dijk, Dobrislav Dobrev, Peter Hansen, Aleksey Kolokolov, Manh Pham, Eric Renault, Roberto Renò, Neil Shephard, Ruixun Zhang (discussant), as well as participants at the Financial Econometrics Conference in Honour of Stephen Taylor, QFFE 2023, IAAE 2023, Asian Meeting of the Econometric Society 2023, EcoSta 2023, 3rd International Econometrics PhD Conference, CFE-CMStatistics 2023, Virtual Workshop for Junior Researchers in Time Series 2024, SoFiE 2024, EEA-ESEM 2024, NBER-NSF Time Series Conference 2024, 5th International Symposium on Interval Data Modelling, and seminars at various institutions, for helpful comments and suggestions. We further acknowledge the insightful comments from Man AHL (London) on the empirical applications. This paper was previously circulated under the title “Nonparametric Range-Based Estimation of Integrated Variance with Episodic Extreme Return Persistence.” Shifan Yu acknowledges financial support from the Economic and Social Research Council (ESRC), United Kingdom North West Social Science Doctoral Training Partnership (ES/P000665/1). The usual disclaimer applies.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jeconom.2025.106014>.

References

- Aït-Sahalia, Y., Jacod, J., 2009. Testing for jumps in a discretely observed process. *Ann. Stat.* 37 (1), 184–222.
- Aït-Sahalia, Y., Jacod, J., 2014. *High-Frequency Financial Econometrics*. Princeton University Press.
- Aït-Sahalia, Y., Li, C.X., Li, C., 2025. So many jumps, so little news. Working Paper.
- Aït-Sahalia, Y., Mykland, P.A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. *Rev. Financ. Stud.* 18 (2), 351–416.

- Aït-Sahalia, Y., Xiu, D., 2019. A Hausman test for the presence of market microstructure noise in high frequency data. *J. Econom.* 211 (1), 176–205.
- Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002. Range-based estimation of stochastic volatility models. *J. Financ.* 57 (3), 1047–1091.
- Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Internat. Econ. Rev.* 39 (4), 885–905.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Rev. Econ. Stat.* 89 (4), 701–720.
- Andersen, T.G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. *J. Econom.* 169 (1), 75–93.
- Andersen, T.G., Li, Y., Todorov, V., Zhou, B., 2023. Volatility measurement with pockets of extreme return persistence. *J. Econom.* 237 (2), 105439.
- Andersen, T.G., Todorov, V., Zhou, B., 2025. Real-time detection of local no-arbitrage violations. *Quant. Econ.* (forthcoming).
- Bajgrowicz, P., Scaillet, O., Treccani, A., 2016. Jumps in high-frequency data: Spurious detections, dynamics, and news. *Manag. Sci.* 62 (8), 2198–2217.
- Ball, C.A., Torous, W.N., 1984. The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. *J. Bus.* 57 (1), 97–112.
- Bandi, F.M., Renò, R., 2016. Price and volatility co-jumps. *J. Financ. Econ.* 119 (1), 107–146.
- Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: Trades and quotes. *Econom. J.* 119 (3), 1–32.
- Barndorff-Nielsen, O.E., Kinnebrook, S., Shephard, N., 2010. Measuring downside risk – realised semivariance. In: Bollerslev, T., Russell, J., Watson, M. (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Oxford University Press, New York, p. 2010.
- Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. *J. Financ. Econom.* 2 (1), 1–37.
- Barndorff-Nielsen, O.E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. *J. Financ. Econom.* 4 (1), 1–30.
- Beckers, S., 1983. Variances of security price returns based on high, low, and closing prices. *J. Bus.* 56 (1), 97–112.
- Bellia, M., Christensen, K., Kolokolov, A., Pelizzon, L., Renò, R., 2024. Do Designated Market Makers Provide Liquidity During Extreme Price Movements? Working Paper.
- Bollerslev, T., Li, J., Li, Q., 2024. Optimal nonparametric range-based volatility estimation. *J. Econom.* 238 (1), 105548.
- Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2016. Exploiting the errors: A simple approach for improved volatility forecasting. *J. Econom.* 192 (1), 1–18.
- Brandt, M.W., Diebold, F.X., 2006. A no-arbitrage approach to range-based estimation of return covariances and correlations. *J. Bus.* 79 (1), 61–74.
- CFTC, 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., Kolokolov, A., 2024. An unbounded intensity model for point process. *J. Econom.* 244 (1), 105840.
- Christensen, K., Oomen, R., Podolskij, M., 2014. Fact or friction: Jumps at ultra high frequency. *J. Financ. Econ.* 114 (3), 576–599.
- Christensen, K., Oomen, R., Renò, R., 2022. The drift burst hypothesis. *J. Econom.* 227 (2), 461–497.
- Christensen, K., Podolskij, M., 2007. Realized range-based estimation of integrated variance. *J. Econom.* 141 (2), 323–349.
- Christensen, K., Podolskij, M., 2012. Asymptotic theory of range-based multipower variation. *J. Financ. Econom.* 10 (3), 417–456.
- Christensen, K., Podolskij, M., Vetter, M., 2009. Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance Stoch.* 13 (2), 239–268.
- Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *J. Financ. Econom.* 7 (2), 174–196.
- Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. *J. Econom.* 159 (2), 276–288.
- Duembgen, M., Podolskij, M., 2015. High-frequency asymptotics for path-dependent functionals of Itô semimartingales. *Stochastic Process. Appl.* 125 (4), 1195–1217.
- Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. *J. Financ.* 58 (3), 1269–1300.
- Flora, M., Renò, R., 2024. V-shapes. Working Paper.
- Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. *J. Bus.* 53 (1), 67–78.
- Grigelionis, B., 1980. A martingale approach to the statistical problems of point processes. *Scand. J. Stat.* 7 (4), 190–196.
- Hall, P., Horowitz, J.L., Jing, B.-Y., 1995. On blocking rules for the bootstrap with dependent data. *Biometrika* 82 (3), 561–574.
- Hansen, P.R., Lunde, A., Nason, J.M., 2003. Choosing the best volatility models: The model confidence set approach. *Oxf. Bull. Econ. Stat.* 65, 839–861.
- Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. *Econometrica* 79 (2), 453–497.
- Hausman, J.A., 1978. Specification tests in econometrics. *Econometrica* 46 (6), 1251–1271.
- Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.* 6 (2), 327–343.
- Holden, C.W., Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *J. Financ.* 69 (4), 1747–1785.
- Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. *J. Financ. Econom.* 3 (4), 456–499.
- Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.* 119 (7), 2249–2276.
- Jacod, J., Li, Y., Zheng, X., 2019. Estimating the integrated volatility with tick observations. *J. Econom.* 208 (1), 80–100.
- Jacod, J., Todorov, V., 2010. Do price and volatility jump together? *Ann. Appl. Probab.* 20 (4), 1425–1469.
- Jacod, J., Todorov, V., 2014. Efficient estimation of integrated volatility in presence of infinite variation jumps. *Ann. Stat.* 42 (3), 1029–1069.
- Jacod, J., Todorov, V., 2018. Limit theorems for integrated local empirical characteristic exponents from noisy high-frequency data with application to volatility and jump activity estimation. *Ann. Appl. Probab.* 28 (1), 511–576.
- Kirilenko, A., Kyle, A.S., Samadi, M., Tuzun, T., 2017. The flash crash: High-frequency trading in an electronic market. *J. Financ.* 72 (3), 967–998.
- Kolokolov, A., 2023. Cryptocrashes. Working Paper.
- Kolokolov, A., Renò, R., Zoi, P., 2025. BUMVU estimators. *J. Econom.* (forthcoming).
- Kunitomo, N., 1992. Improving the Parkinson method of estimating security price volatilities. *J. Bus.* 65 (2), 295–302.
- Laly, F., Petitjean, M., 2020. Mini flash crashes: Review, taxonomy and policy responses. *Bull. Econ. Res.* 72 (3), 251–271.
- Laurent, S., Renò, R., Shi, S., 2024. Realized drift. *J. Econom.* (forthcoming).
- Laurent, S., Shi, S., 2020. Volatility estimation and jump detection for drift-diffusion processes. *J. Econom.* 217 (2), 259–290.
- Li, J., Wang, D., Zhang, Q., 2024. Reading the candlesticks: An OK estimator for volatility. *Rev. Econ. Stat.* 106 (4), 1114–1128.
- Liu, L.Y., Patton, A.J., Sheppard, K., 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *J. Econom.* 187 (1), 293–311.
- Mancini, C., 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* 36 (2), 270–296.
- Mancini, C., 2023. Drift burst test statistic in the presence of infinite variation jumps. *Stochastic Process. Appl.* 163, 535–591.
- Martens, M., van Dijk, D., 2007. Measuring volatility with the realized range. *J. Econom.* 138 (1), 181–207.
- Meilijson, I., 2011. The Garman-Klass volatility estimator revisited. *Revstat* 9 (3), 199–212.
- Menkveld, A.J., Yueshen, B.Z., 2019. The flash crash: A cautionary tale about highly fragmented markets. *Manag. Sci.* 65 (10), 4470–4488.
- Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. *J. Bus.* 53 (1), 61–65.
- Patton, A.J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Rev. Econ. Stat.* 97 (3), 683–697.
- Podolskij, M., Vetter, M., 2009. Bipower-type estimation in a noisy diffusion setting. *Stochastic Process. Appl.* 119 (9), 2803–2831.
- Renault, E., Sarisoy, C., Werker, B.J., 2017. Efficient estimation of integrated volatility and related processes. *Econometric Theory* 33 (2), 439–478.
- Rogers, L.C.G., Satchell, S.E., 1991. Estimating variance from high, low and closing prices. *Ann. Appl. Probab.* 1 (4), 504–512.
- Swanson, N.R., White, H., 1997. Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. *Int. J. Forecast.* 13 (4), 439–461.
- Todorov, V., Tauchen, G., 2011. Volatility jumps. *J. Bus. Econ. Statist.* 29 (3), 356–371.
- Vetter, M., 2010. Limit theorems for bipower variation of semimartingales. *Stochastic Process. Appl.* 120 (1), 22–38.
- Yang, D., Zhang, Q., 2000. Drift-independent volatility estimation based on high, low, open, and close prices. *J. Bus.* 73 (3), 477–492.