Journal of Econometrics 238 (2024) 105548

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Optimal nonparametric range-based volatility estimation ☆

Tim Bollerslev [a,b], Jia Li [c,*], Qiyuan Li [c]

[a] Department of Economics, Duke University, Durham, NC 27708, United States of America
[b] NBER, United States of America
[c] School of Economics, Singapore Management University, Singapore

ARTICLE INFO

JEL classification:
C14
C22

Keywords:
Spot volatility
Nonparametric estimation
Range-based estimation
High-frequency data
Decision theory

ABSTRACT

We present a general framework for optimal nonparametric spot volatility estimation based on intraday range data, comprised of the first, highest, lowest, and last price over a given time-interval. We rely on a decision-theoretic approach together with a coupling-type argument to directly tailor the form of the nonparametric estimator to the specific volatility measure of interest and relevant loss function. The resulting new optimal estimators offer substantial efficiency gains compared to existing commonly used range-based procedures.

1. Introduction

Most financial and macroeconomic time series exhibit time-varying volatility. Accurate assessments of said volatilities are important for financial decision making and the evaluation of economic policies alike. Accordingly, a large econometrics literature has emerged over the past several decades dedicated to the development of ever more reliable volatility estimation procedures. We add to this burgeon literature by providing new optimal range-based volatility estimators.[1] We rely on a novel decision-theoretic approach together with a coupling-type asymptotic representation to explicitly tailor the form of the optimal estimator to the volatility measure of interest and relevant loss function. In so doing, we demonstrate nontrivial efficiency gains for the new optimal estimators compared to commonly used procedures.

Prompted by the increased availability of high-frequency intraday prices for a variety of financial assets and markets, most of the volatility estimation procedures proposed in the more recent literature have been nonparametric, built on the notion of ever finer sampled returns and corresponding infill asymptotic arguments (see, e.g., the introductory discussion in Andersen and Bollerslev (2018)). In a stylized theoretical setting, the use of finely sampled intraday returns naturally affords more accurate volatility estimates than the use of coarser, say daily, returns. Empirically, however, the presence of market microstructure "noise" presents formidable challenges to the direct use of ultra high-frequency returns, necessitating more advanced robust inference procedures and/or the use of "not-too-finely" sampled intraday returns (see, e.g., the discussion in Jacod et al. (2017) and Li and Linton (2022), along with the many additional references therein).

☆ We would like to thank Xiaohong Chen, Frank Diebold (discussant), Rob Engle, Peter Phillips, Oliver Linton, Jun Yu, and participants at the SoFiE conference, SH3 conference, and various seminars for their inspiring comments. J. Li's research was partially supported by Ministry of Education in Singapore tier-1 grant 22-SOE-SMU-016.

* Corresponding author.
E-mail addresses: boller@duke.edu (T. Bollerslev), jiali@smu.edu.sg (J. Li), qyli.2019@phdecons.smu.edu.sg (Q. Li).

[1] Following the existing literature, we will refer to any estimator that exploits not only the information in the high and low prices over a given time interval, but also the first and last prices over the interval, as a "range-based" estimator. When there is no ambiguity, we will also frequently use the word "volatility" as a catchall for any scale measure, the variance included.

https://doi.org/10.1016/j.jeconom.2023.105548
Received 13 September 2023; Accepted 8 October 2023
Available online 27 October 2023

Meanwhile, pioneering work by Parkinson (1980) and Garman and Klass (1980), dating back almost half-a-century, first demonstrated the increased accuracy for daily variance estimation afforded by harnessing the richer information embedded in the daily high-low range and so-called "candlestick charts," comprised of the open, high, low, and close prices over the day.[2] This type of daily data has long been freely available for a vast array of financial assets. It is now also readily available on an intraday basis.[3] Importantly, and in parallel to the common use of "not-too-finely" sampled high-frequency intraday returns, intraday candlesticks sampled at "not-too-fine" a frequency offer a similar built-in robustness to market microstructure noise, and as such holds the promise of easy-to-implement improved volatility estimation.[4] Yet, it remains an open question how to optimally exploit the full information inherent in such candlesticks for said estimation.

We rely on ideas from decision theory to provide a definitive answer to this question. Classical decision theory generally invokes specific parametric distributional assumptions to determine the optimal estimator that minimizes the specific risk. By contrast, our high-frequency framework adopts a nonparametric approach. We leverage the infill asymptotic "coupling" method recently developed by Bollerslev et al. (2021) to bridge the gap between our setting and the classical decision-theoretic approach. This enables us to derive unique optimal high-frequency range-based spot volatility estimators corresponding to particular loss functions (e.g., Quadratic or Stein) and volatility measures (e.g., σ_t, σ_t^2, σ_t^4, or σ_t^{-1}). For spot estimation based on a single candlestick, we derive closed-form analytical expressions for the optimal estimators. These estimators are non-standard, but straightforward to implement in practice. In cases involving multiple candlesticks, we provide semi-closed form solutions for the optimal estimators and illustrate how to employ machine learning tools to numerically compute the optimal estimation functions.

Our results are most closely related to the recent work of Li et al. (2022). Extending the original analysis in Garman and Klass (1980) based on the assumption of a continuous-time price process with constant volatility to a high-frequency nonparametric infill asymptotic setting, Li et al. (2022) propose a range-based estimator for the spot volatility that achieves the minimum asymptotic variance within the class of unbiased linear estimators. Their proposal may be regarded as the best linear unbiased estimator (BLUE) for spot volatility. While that analysis is informative, it is also incomplete, and by design much simpler than the present analysis. In particular, a priori restricting the functional form of the estimator to be linear simplifies the search for the "optimal" estimator to a search for the optimal set of weights, as opposed to a search for the risk-minimizing estimator in an infinite-dimensional functional space. Importantly, restricting the functional form also does not guarantee that the resulting "shape-constrained" optimal estimator is actually *the* optimal estimator.[5] Indeed, as we demonstrate below, the "unconstrained" optimal nonparametric range-based spot volatility estimators derived here often provide nontrivial efficiency gains compared to existing procedures hitherto derived in the literature under various simplifying assumptions, the classical Garman–Klass estimator and the BLUE estimator of Li et al. (2022) included.

Further relating our work to the existing high-frequency literature on nonparametric volatility estimation, most of the prior theoretical work on optimal estimation of spot volatility (see, e.g., Foster and Nelson (1996), Comte and Renault (1998), Kristensen (2010), and Chapter 13 in Jacod and Protter (2012)) has primarily been concerned with rate optimality. However, that optimality criterion sheds little light on the estimators' actual finite-sample performance.[6] Another strand of the literature has instead been concerned with the semiparametric efficient estimation of integrated volatility functionals (see, e.g., Mykland and Zhang (2009), Jacod and Rosenbaum (2013), Renault et al. (2017), and Li and Liu (2021)). The optimality concept typically adopted in that literature has been built on the convolution theorem and the related local asymptotic minimaxity results for locally asymptotically mixed normal (LAMN) models (see, e.g., Le Cam (1960), Hájek (1972), Jeganathan (1982, 1983)). By contrast, our coupling theory directly links the nonparametric range-based spot volatility estimation/decision problem with a *non*-Gaussian limit experiment. As a result, the functional form of our new optimal estimators generally depend on the loss function and are quite nonstandard, although straightforward to implement in practice.

The remainder of this paper is organized as follows. In Section 2, we start by outlining our nonparametric high-frequency setting and basic assumptions, followed by a discussion of our key coupling arguments. We then introduce the new optimal range-based spot volatility estimators and provide a characterization of their asymptotic properties. Section 3 illustrates the practical applicability of the new estimators, and shows the intraday candlestick-based spot volatility estimates for a market portfolio for each of the eight 2022 prescheduled Federal Open Market Committee (FOMC) announcement days. We conclude with a few suggestions for future research. All proofs are included in the Appendix, while additional theoretical and numerical results can be found in the online Supplemental Appendix.

[2] Candlestick charts are also routinely used by finance practitioners in the formulation of technical trading strategies. The first such documented use of candlestick charting dates back to the 18th century and the Japanese rice trader Munehisa Homma; see, for example, Nison (2001) for an introduction to the main ideas.

[3] High-frequency candlestick data is provided by various online trading platforms (e.g. E-Trade, Robinhood), publicly available databases (e.g., Yahoo Finance), and commercial databases (e.g., Bloomberg, Tick Data, TAQ).

[4] Extending our ideas to range-based estimation with even finer sampled intraday candlesticks for which the noise cannot be ignored would be an interesting direction for future research. However, as discussed further below, the requisite task of pinning down the fine structure of the noise and the underlying economic mechanisms presents formidable challenges beyond our main research question.

[5] Of course, seemingly ad hoc functional-form restrictions do not necessarily result in efficiency loss. For example, in Gaussian linear regression models, the ordinary least-squares estimator is also the uniformly minimum-variance unbiased estimator by the Lehmann–Scheffé theorem (see, e.g., Shaffer (1991)). That is, the BLUE estimator is also the best unbiased estimator (BUE); see also the related recent discussion pertaining to possibly non-Gaussian linear regression models in Hansen (2022) and Pötscher and Preinerstorfer (2022).

[6] Kristensen (2010) does seek to characterize the optimal choice of the smoothing kernel. However, the underlying assumption that the volatility process has differentiable sample paths rules out all Brownian stochastic volatility models, as well as any model featuring volatility jumps.

2. Nonparametric range-based volatility estimation

2.1. Theoretical setting and decision-theoretic framework

The (log) price process P is assumed to follow an Itô semimartingale defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ of the form

$$P_t = P_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t, \quad 0 \leq t \leq T, \tag{2.1}$$

where the drift process b and the volatility process σ are both c àdlàg adapted, W is a standard Brownian motion, and J is a pure-jump process driven by a Poisson random measure. We are interested in the optimal nonparametric estimation of the pth power of the spot volatility, σ_t^p, at some fixed time t under a standard infill asymptotic setting with the sampling interval $\Delta_n \to 0$. We will focus our discussion in the main text on cases with $p = 2$ (variance) and $p = 1$ (volatility). The same ideas may similarly be applied in the construction of optimal estimators for other powers p.[7]

The baseline Itô semimartingale in (2.1) is directly motivated by no-arbitrage arguments. However, it is well-known that the process is misspecified empirically at ultra high, or tick level, frequencies. In addition to a host of market microstructure frictions that "contaminate" the actually observed prices (see, e.g., Diebold and Strasser (2013) for a discussion of the underlying economic mechanisms), prices are also not truly recorded on a continuous-time scale. The most commonly used approach to circumvent these difficulties for the purpose of volatility estimation is to "down-sample" the available data, and rely on returns at "not-too-high" a frequency $1/\Delta_n$. The practical choice of Δ_n has typically been guided by the so-called volatility signature plot introduced by Andersen et al. (2000) (see also the discussion in Hansen and Lunde (2006), and the recent formalization of that approach in Aït-Sahalia and Xiu (2019)). The new estimation method proposed here is similarly intended to be used with "not-too-finely" sampled data. Put differently, acknowledging that the workhorse Itô semimartingale model is only meant as a plausible approximation over "coarser" time scales, effectively allows us to follow the common approach in the literature and remain agnostic about the fine structure of the market microstructure noise.[8]

The existing high-frequency econometrics literature on nonparametric volatility estimation has primarily been focused on estimators formed using high-frequency returns; i.e., $P_{i\Delta_n} - P_{(i-1)\Delta_n}$. We augment the information in the high-frequency return by "looking inside" the Δ_n time-interval through the lens of high-frequency candlesticks. More specifically, denote the ith sampling interval by $\mathcal{T}_i = [(i-1)\Delta_n, i\Delta_n]$. The corresponding candlestick then provides information on the open, high, low, and close prices, formally defined by $P_{(i-1)\Delta_n}$, $\sup_{t\in\mathcal{T}_i} P_t$, $\inf_{t\in\mathcal{T}_i} P_t$, and $P_{i\Delta_n}$, respectively. This information may be summarized in the form of the three (normalized) returns

$$r_i \equiv \frac{P_{i\Delta_n} - P_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}, \quad u_i \equiv \frac{\sup_{t\in\mathcal{T}_i} P_t - P_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}, \quad l_i \equiv \frac{\inf_{t\in\mathcal{T}_i} P_t - P_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}, \tag{2.2}$$

where r_i denotes the usual open-close return traditionally used for high-frequency-based volatility estimation, and u_i (resp. l_i) refers to the high (resp. low) return brought by the candlestick (to help fix ideas, see Fig. 1). All range-based estimators may be expressed as functions of (r_i, u_i, l_i). To facilitate our representation and subsequent discussion of the optimal estimators, it is convenient to also define the scaled range $w_i \equiv u_i - l_i$ (as also indicated in Fig. 1), and a measure of asymmetry $a_i \equiv |u_i + l_i - r_i|$. The asymmetry measure quantifies the absolute difference between the lengths of the "wicks" above and below the rectangular box of the candlestick. The candlestick is symmetric if and only if $a_i = 0$.

To more clearly highlight the key novelty of our approach, we first focus on estimators based on a single high-frequency candlestick "neighboring" t in the sense that $|i\Delta_n - t| = o(1)$.[9] Optimal estimation with multiple adjacent candlesticks is discussed in Section 2.4. Accordingly, we will express our estimators for σ_t^p generically as

$$S = f(r_i, u_i, l_i), \tag{2.3}$$

for some function $f(\cdot)$. Since spot volatility is fundamentally a "scale parameter," we will restrict our attention to *scale-equivariant* estimators, requiring the function $f(\cdot)$ to be homogeneous of degree p, that is, $f(\lambda x) = \lambda^p f(x)$ for any $\lambda > 0$. We will further refer to the

[7] Analogous results for $p = 4$ (quarticity) and $p = -1$ (precision) are presented in the online Supplemental Appendix.

[8] Alternatively, one could impose more explicit assumptions about the form of the noise, and the way in which the prices observed at ultra high frequencies differ from the efficient prices. However, it is far from obvious how the noise component should be modeled, plus the "right" choice is invariably asset and/or market specific. For instance, are the conditional moments of the noise constant or time-varying; does the noise exhibit conditional and/or unconditional serial dependence; should the noise be treated as "small" (i.e., local-to-zero) or "large;" is the noise correlated with the latent efficient price (see, e.g., Kalnina and Linton (2008), Zhang et al. (2005), Jacod et al. (2017), and Li and Linton (2022)). Further complicating matters, the broader econometrics literature on nonclassical measurement errors (see, e.g., Schennach (2020)) also calls into question the "classical" additive separability and mean independence assumptions routinely invoked in the high-frequency econometrics literature, and instead suggests that the noise may be better accounted for using nonclassical models (as in, e.g., Berkson (1950) and Hyslop and Imbens (2001)). Hence, while it is conceivable that the new approach developed here could be extended to allow for the use of ultra high-frequency data by explicating the "fine structure" of the noise, any associated theoretical efficiency claims would come with the perhaps even more challenging task of justifying the *additional* requisite assumptions.

[9] Note, the index i generally also depends on n. We purposely suppress this dependence in our notation so as to avoid nested subscripts.

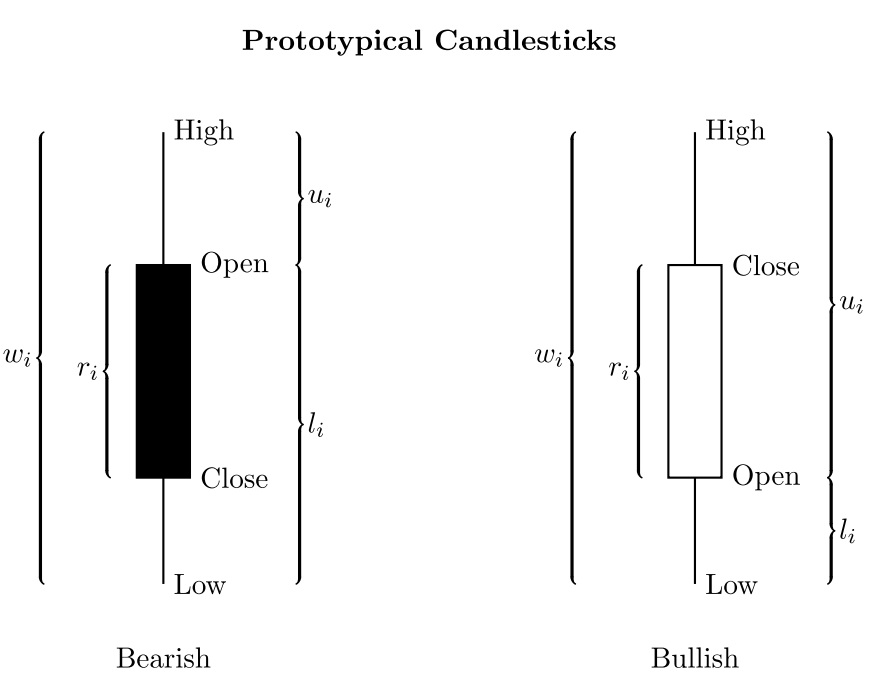

Fig. 1. The figure shows two prototypical candlesticks, bearish and bullish, comprised of the open, high, low, and close prices. The corresponding return r_i, range w_i, high return u_i, and low return l_i, as defined in Eq. (2.2), are explicitly highlighted.

estimator as *regular* if $f(\cdot)$ is continuous (Lebesgue) almost everywhere. This regularity requirement seems rather innocuous. However, it ensures that any candidate estimator has a limit distribution that is also scale-equivariant.[10] As shown in Theorem 1 below, it also proves sufficient to "couple" the original nonparametric estimation problem with a much simpler limit decision problem.

In the analysis of scale estimation problems, it is also standard to gauge the estimator's performance by a scale-invariant loss function. For any non-negative loss function $L(\cdot)$ this is readily achieved by considering the scaled loss $L(S/\sigma_t^p)$. Correspondingly, the *risk* of the estimator S may be succinctly expressed as

$$R(S; L) \equiv \mathbb{E}[L(S/\sigma_t^p)]. \tag{2.4}$$

It is impossible to obtain an optimal estimator that minimizes $R(S; L)$ under the general nonparametric model in (2.1). Intuitively, as the joint distribution of the data vector (r_i, u_i, l_i) is determined by the unknown joint law of the (b, σ, W, J) process this would essentially amount to an optimization problem with an infinite-dimensional nuisance parameter. Importantly, however, under mild regularity conditions on the price process, the multiplicative estimation error S/σ_t^p may be shown to be *asymptotically pivotal* for any regular scale-equivariant estimator. Consequently, the asymptotic loss and risk are both nuisance-free, simplifying the optimality analysis.

The following regularity condition, which is standard in the literature on nonparametric volatility estimation (see, e.g., Jacod and Protter (2012), Jacod et al. (2021), Bollerslev et al. (2021), and Li et al. (2022)) suffices for this pivotalization scheme to obtain.

Assumption 1. Suppose that the price process P has the form in (2.1) and that there exists a sequence $(T_m)_{m\geq 1}$ of stopping times increasing to infinity and a sequence $(K_m)_{m\geq 1}$ of finite constants such that the following conditions hold for each $m \geq 1$: (i) for all $t \in [0, T_m]$, $|b_t| + |\sigma_t| + |\sigma_t|^{-1} + F_t(\mathbb{R} \setminus \{0\}) \leq K_m$, where F_t denotes the spot Lévy measure of J; (ii) for some constant $\kappa > 0$, $\mathbb{E}[|\sigma_{t\wedge T_m} - \sigma_{s\wedge T_m}|^2] \leq K_m|t - s|^{2\kappa}$ for all $s, t \in [0, T]$.

Assumption 1 necessitates that various processes are bounded by a finite constant K_m up to a stopping time T_m, without requiring the bound to hold over the entire sample span. This setup is commonly employed when applying localization, a standard technique in stochastic calculus used for extending limit theorems under weaker conditions. For a comprehensive discussion on its application in the analysis of high-frequency data, see, e.g., Section 4.4.1 in Jacod and Protter (2012). The parameter κ, defined as the Hölder continuity index for the volatility process σ under the L_2 norm, pertains to the smoothness of σ. If the volatility is driven by a Brownian motion, κ is at most $1/2$, and the volatility path is non-differentiable everywhere. This setting differs from typical nonparametric problems, where unknown functions are often assumed to be differentiable of higher order. Values of $\kappa < 1/2$ also permits the volatility to exhibit "rough" paths, as emphasized by Gatheral et al. (2018) among others.

[10] This mirrors the notion of regularity in Gaussian shift limit experiments that requires the estimator to be asymptotically location-equivariant (see, e.g., Van der Vaart (1998)).

The following theorem stipulates a general asymptotic representation for any regular scale-equivariant estimator $S = f(r_i, u_i, l_i)$ allowed under these mild conditions. By linking the nonparametric estimation problem with that in a limit *non*-Gaussian experiment, the result differs notably from the Gaussian shift experiment commonly used in the analysis of semiparametric efficiency, the estimation of integrated volatility functionals included.

Theorem 1. *Under Assumption 1, any regular scale-equivariant estimator $S = f(r_i, u_i, l_i)$ with $|i\Delta_n - t| \to 0$ may be expressed as*

$$\frac{S}{\sigma_t^p} = f(\zeta_i) + o_p(1), \tag{2.5}$$

where $\zeta_i \equiv (\zeta_{i,r}, \zeta_{i,u}, \zeta_{i,l})$ and

$$\zeta_{i,r} \equiv \frac{W_{i\Delta_n} - W_{(i-1)\Delta_n}}{\sqrt{\Delta_n}}, \quad \zeta_{i,u} \equiv \frac{\sup_{s\in\mathcal{T}_i}(W_s - W_{(i-1)\Delta_n})}{\sqrt{\Delta_n}}, \quad \zeta_{i,l} \equiv \frac{\inf_{s\in\mathcal{T}_i}(W_s - W_{(i-1)\Delta_n})}{\sqrt{\Delta_n}}.$$

The theorem shows that the multiplicative estimation error in S may be decomposed into a nondegenerate leading term $f(\zeta_i)$ and an asymptotically negligible $o_p(1)$ term. The $o_p(1)$ term absorbs various nonparametric biases stemming from the drift, time-variation of volatility, and jumps. If the price was simply a scaled Brownian motion, this term would be identically equal to zero. Importantly, the distribution of the ζ_i random variable that determines the leading $f(\zeta_i)$ term is known in finite samples. To appreciate this point, let B denote a generic copy of the standard Brownian motion on the unit interval $[0,1]$ with $B_0 = 0$. It then follows that

$$\zeta_i \overset{d}{=} \tilde{\zeta} \equiv \Big(B_1, \sup_{t\in[0,1]} B_t, \inf_{t\in[0,1]} B_t \Big). \tag{2.6}$$

Since this distribution, and by implication the $f(\zeta_i)$ term in (2.5), are both nuisance-free, the multiplicative estimation error S/σ_t^p is therefore also asymptotically pivotal.[11]

If the loss function $L(\cdot)$ is continuous, Theorem 1 further implies an analogous coupling result for the estimation loss

$$L(S/\sigma_t^p) = L\big(f(\zeta_i)\big) + o_p(1). \tag{2.7}$$

Following the literature (e.g., Le Cam (1986) and Van der Vaart (1998)), this naturally suggests defining the asymptotic risk of any regular scale-equivariant estimator as the expected value of the limit loss $L\big(f(\zeta_i)\big)$. By (2.6) the asymptotic risk may thus be expressed as

$$\widetilde{R}(S; L) \equiv \mathbb{E}\big[L\big(f(\tilde{\zeta})\big)\big]. \tag{2.8}$$

The distribution of $\tilde{\zeta}$ is known in finite sample and so $\widetilde{R}(S; L)$ can be readily evaluated for any loss function $L(\cdot)$ and estimator $f(\cdot)$. We will refer to a regular scale-equivariant estimator S as *optimal*, or more precisely as an Asymptotic Minimum-Risk scale-Equivariant (AMRE) estimator, if it minimizes $\widetilde{R}(S; L)$. Since this asymptotic risk does not depend on any unknown quantities in the nonparametric model in (2.1), this optimality concept is valid in a *uniform* sense. As such, it also implies asymptotic admissibility and minimaxity (within the class of regular scale-equivariant estimators). Consequently, any suboptimal estimator is necessarily asymptotically inadmissible.

Theorem 1 is based on the Itô semimartingale model (2.1), which as previously noted does not explicitly incorporate microstructure noise. However, the same result remains valid if the observed prices are affected by "small" noise. Specifically, if the magnitude of the noise is of order $o_p(\Delta_n^{1/2})$, the "noisy" observation of (r_i, u_i, l_i) deviates from their true value by $o_p(1)$, which, according to the continuous mapping theorem, implies that (2.5) also holds for the "noisy" estimator. Intuitively, the $o_p(\Delta_n^{1/2})$ rate requirement for the noise is more plausible when Δ_n is not "too-small", consistent with the idea of not using "too-finely" sampled data, or coarse sampling.[12]

In order to construct an AMRE estimator, it is helpful to recognize that the asymptotic risk of a regular estimation function $f(\cdot)$ can be equivalently represented in terms of its finite-sample risk in a limit parametric model, where the (log) price process P is a simple scaled Brownian motion (i.e., $P_t = \sigma W_t$). This, in turn, facilitates the use of classical finite-sample theory for optimal equivariant estimation in determining the optimal estimation function and the AMRE estimator.[13] The AMRE estimators presented in Sections 2.2 and 2.4 below, as well as the additional estimators discussed in the online Supplemental Appendix, are developed using this approach.

[11] This nuisance-free limit distribution also permits the construction of asymptotically valid confidence intervals for σ_t^p. For any $\alpha \in (0,1)$, let L and U be constants such that $\mathbb{P}\big(L \le 1/f(\tilde{\zeta}) \le U\big) = 1 - \alpha$. Then $[LS, US]$ is a confidence interval for σ_t^p at asymptotic level $1-\alpha$. The length of the interval is minimized by taking $[L, U]$ as the $1-\alpha$ level highest density interval of the distribution of $1/f(\tilde{\zeta})$.

[12] In line with the existing empirical literature on high-frequency-based volatility estimation, we recommend adopting a $\Delta_n = 5$-min sampling scheme as the default choice in practice. Simulation results in the online Supplemental Appendix also demonstrate that the noise, when calibrated to empirically realistic levels, has a negligible effect on the resulting 5-min estimators. As the noise level becomes higher, the noise leads to larger positive bias in the volatility estimates, and the shrinkage estimator derived from minimizing the quadratic loss tends to outperform other estimators.

[13] According to Corollary 3.3.4 in Lehmann and Casella (1998), the solution to the functional minimization problem $\min_f \mathbb{E}\big[L\big(f(\tilde{\zeta})\big)\big]$ exists and is unique, provided that an equivariant estimator with finite risk exists and the function $x \mapsto L(e^x)$ is strictly convex and not monotone.

2.2. Optimal estimation for spot variance and volatility

To streamline the presentation and more clearly highlight our main theoretical contributions, we will focus our discussion on the optimal estimation of the spot variance σ_t^2 and the spot volatility σ_t.[14] We will restrict our attention to optimal estimators based on Stein's loss and the standardized quadratic loss,

$$L_{\text{Stein}}(x) \equiv x - \log(x) - 1, \qquad L_{\text{Quad}}(x) \equiv |x-1|^2. \tag{2.9}$$

These specific loss functions arguably constitute the two most commonly used losses in practice. AMRE estimators for other, possibly case-specific, loss functions could be derived similarly.

To facilitate the representation of the optimal estimators, it is helpful to define the functions $G_q(\cdot)$ and $H_q(\cdot)$ for any integer $q \geq 0$ as,

$$\begin{aligned} G_q(x) \equiv\ & \psi_q\Big(\frac{1-x}{2}\Big) + \psi_q\Big(\frac{1+x}{2}\Big) - \frac{x}{q+1}\left(\psi_{q+1}\Big(\frac{1-x}{2}\Big) - \psi_{q+1}\Big(\frac{1+x}{2}\Big)\right) \\ & - \frac{1-x^2}{4(q+1)(q+2)}\left(\psi_{q+2}\Big(\frac{1-x}{2}\Big) + \psi_{q+2}\Big(\frac{1+x}{2}\Big)\right), \end{aligned} \tag{2.10}$$

$$\begin{aligned} H_q(x) \equiv\ & \psi_q\Big(1-\frac{x}{2}\Big) + \psi_q\Big(\frac{x}{2}\Big) - \frac{x}{q+1}\left(\psi_{q+1}\Big(1-\frac{x}{2}\Big) - \psi_{q+1}\Big(\frac{x}{2}\Big)\right) \\ & + \frac{x^2}{4(q+1)(q+2)}\left(\psi_{q+2}\Big(1-\frac{x}{2}\Big) + \psi_{q+2}\Big(\frac{x}{2}\Big)\right), \end{aligned} \tag{2.11}$$

where $\psi_q(x)$ denotes the polygamma function of order q, that is, the $(q+1)$th-order derivative of the logarithm of the Gamma function. The $G_q(\cdot)$ and $H_q(\cdot)$ functions are both continuous almost everywhere, making them suitable for constructing regular estimators.[15] Using these definitions, the subsequent theorem offers explicit analytical expressions for the AMRE estimators of the spot variance and spot volatility under each of the two loss functions.

Theorem 2. *Under the same setting as Theorem 1, we have*

(a) the AMRE range-based estimator for σ_t^2 under Stein's loss is asymptotically unbiased and given by

$$\hat{\sigma}^2_{\text{Stein}} \equiv \frac{4w_i^2}{3} \cdot \frac{G_0(a_i/w_i) - H_0(|r_i|/w_i)}{G_2(a_i/w_i) - H_2(|r_i|/w_i)},$$

while the AMRE range-based estimator for σ_t^2 under standardized quadratic loss equals

$$\hat{\sigma}^2_{\text{Quad}} \equiv \frac{32w_i^2}{5} \cdot \frac{G_2(a_i/w_i) - H_2(|r_i|/w_i)}{G_4(a_i/w_i) - H_4(|r_i|/w_i)};$$

(b) the AMRE range-based estimator for σ_t under Stein's loss is asymptotically unbiased and given by

$$\hat{\sigma}_{\text{Stein}} \equiv \frac{\sqrt{2\pi}}{3} w_i \cdot \frac{G_0(a_i/w_i) - H_0(|r_i|/w_i)}{H_1(|r_i|/w_i) - G_1(a_i/w_i)},$$

while the AMRE range-based estimator for σ_t under standardized quadratic loss equals

$$\hat{\sigma}_{\text{Quad}} \equiv 2\sqrt{\frac{2}{\pi}} w_i \cdot \frac{H_1(|r_i|/w_i) - G_1(a_i/w_i)}{G_2(a_i/w_i) - H_2(|r_i|/w_i)}.$$

Comment 1. The asymptotic unbiasedness of the $\hat{\sigma}^2_{\text{Stein}}$ and $\hat{\sigma}_{\text{Stein}}$ estimators is reminiscent of the classical finite-sample result that minimum-risk estimators under Stein's loss are guaranteed to be unbiased. As demonstrated by Brown (1968), Stein's loss is also the unique loss function (up to affine transformations) that satisfies this property. Consequently, AMRE estimators under other loss functions are necessarily asymptotically biased.

Comment 2. The AMRE estimators depend solely on the shape of the candlestick, as summarized by the scaled range $w_i \equiv u_i - l_i$, the scaled absolute return $|r_i|$, and the a_i asymmetry measure. These shape-related features remain unaffected by a "color change" or an "upside-down flip" of the candlestick.[16] Consequently, the optimal volatility estimators are also invariant to these "directional" transformations. This feature reduction is due to a sufficiency argument, as formalized by Lemma A.1 in the Appendix, which shows that the shape features $(w_i, |r_i|, a_i)$ are indeed sufficient statistics for σ in the limit model $P_t = \sigma W_t$. Recall that according to the Rao–Blackwell theorem, optimal estimators depend on data only through sufficient statistics.

[14] Analogous derivations for the spot quarticity σ_t^4 and spot precision σ_t^{-1} are provided in the online Supplemental Appendix.

[15] The almost everywhere continuity of the $G_q(\cdot)$ and $H_q(\cdot)$ functions follows from the fact that polygamma functions are formally meromorphic, meaning that they are analytic except for a discrete set of points.

[16] More precisely, the color change corresponds to changing the sign of r_i and the upside-down flip amounts to swapping the upper and lower shadows of the candlestick.

The AMRE estimation functions defined in Theorem 2 are relatively complex. Clearly, it would have been challenging to accurately "intuit" these specific functional forms when searching for optimal estimators within a restricted class of functions. Since the AMRE estimators are uniquely determined (Corollary 3.3.4 in Lehmann and Casella (1998)), any ad hoc restrictions on the functional form used in the derivation of "shape-constrained" optimal estimators would therefore also generally result in suboptimal and, as previously mentioned, asymptotically inadmissible estimators.

For a more direct comparison, recall that Garman and Klass's (1980) minimum-variance unbiased quadratic estimator for spot variance is given by[17]

$$\begin{aligned}\hat{\sigma}^2_{\text{GK}} &\equiv 0.511(u_i - l_i)^2 - 0.019\big(r_i(u_i + l_i) - 2u_i l_i\big) - 0.383 r_i^2 \\ &= 0.5015 w_i^2 + 0.0095 a_i^2 - 0.3925 r_i^2,\end{aligned}$$

while the BLUE for spot volatility proposed by Li et al. (2022) is given by

$$\hat{\sigma}_{\text{BLUE}} \equiv 0.811 w_i - 0.369 |r_i|.$$

Meanwhile, approximating the functional forms of the AMRE estimators described in Theorem 2 by cubic polynomials of the (maximal invariant) ratio statistics, $|r_i|/w_i$ and a_i/w_i, the spot variance estimator may be expressed as[18]

$$\begin{aligned}\hat{\sigma}^2_{\text{Stein}} &\approx 0.5921 w_i^2 - 0.2066 |r_i| w_i - 0.1289 a_i^2 - 0.5874 r_i^2 - 0.0001 \frac{a_i^3}{w_i} \\ &\quad + 0.0382 \frac{|r_i| a_i^2}{w_i} - 0.0001 \frac{r_i^2 a_i}{w_i} + 0.3872 \frac{|r_i|^3}{w_i}, \\ \hat{\sigma}^2_{\text{Quad}} &\approx 0.4936 w_i^2 - 0.0002 a_i w_i - 0.2436 |r_i| w_i - 0.1003 a_i^2 + 0.0001 |r_i| a_i \\ &\quad - 0.4316 r_i^2 - 0.0006 \frac{a_i^3}{w_i} + 0.0883 \frac{|r_i| a_i^2}{w_i} - 0.0005 \frac{r_i^2 a_i}{w_i} + 0.3282 \frac{|r_i|^3}{w_i},\end{aligned}$$

while the analogous approximations for the AMRE spot volatility estimators take the form

$$\begin{aligned}\hat{\sigma}_{\text{Stein}} &\approx 0.7859 w_i - 0.1010 |r_i| - 0.0888 \frac{a_i^2}{w_i} - 0.4798 \frac{r_i^2}{w_i} - 0.0178 \frac{a_i^2 |r_i|}{w_i^2} + 0.2341 \frac{|r_i|^3}{w_i^2}, \\ \hat{\sigma}_{\text{Quad}} &\approx 0.7526 w_i - 0.1366 |r_i| - 0.0846 \frac{a_i^2}{w_i} - 0.0001 \frac{a_i |r_i|}{w_i} - 0.4345 \frac{r_i^2}{w_i} + 0.0181 \frac{a_i^2 |r_i|}{w_i^2} \\ &\quad - 0.0001 \frac{a_i r_i^2}{w_i^2} + 0.2284 \frac{|r_i|^3}{w_i^2}.\end{aligned}$$

While not exact, these cubic expansions formally highlight the differences between the AMRE estimators and the shape-constrained estimators, by explicating the former's dependence on additional higher-order nonlinear features.

To help more clearly visualize these differences, Fig. 2 present the estimation functions for the three spot variance estimators: $\hat{\sigma}^2_{\text{Stein}}$, $\hat{\sigma}^2_{\text{Quad}}$, and $\hat{\sigma}^2_{\text{GK}}$. As the estimators are all scale-equivariant, we compare them without loss of generality under the scale normalization $w_i = 1$ (i.e., $|r_i|$ and a_i are interpreted in a relative sense). In the left panel of Fig. 2, we further fix the asymmetry factor at $a_i = 0$, and plot the spot variance estimators as functions of the absolute return $|r_i|$. Looking at the two asymptotically unbiased estimators, $\hat{\sigma}^2_{\text{Stein}}$ and $\hat{\sigma}^2_{\text{GK}}$, the former is higher when $|r_i|$ is close to 0 or 1, and lower when $|r_i|$ takes medium values.[19] Meanwhile, the estimation function associated with $\hat{\sigma}^2_{\text{Quad}}$ is systematically below the estimation functions for the two unbiased estimators, indicating that the AMRE estimator under quadratic loss exhibits a certain "shrinkage" and therefore also is downward biased.

In the right panel of Fig. 2, we fix $|r_i| = 0.3$ and plot the estimators as functions of the asymmetry factor a_i.[20] While $\hat{\sigma}^2_{\text{GK}}$ displays a slightly positive dependence on the asymmetry factor, the two AMRE estimators evidence a more pronounced negative dependence. Comparing the left and right panels further reveals that the absolute return has a greater impact on variance estimation than the asymmetry factor.

A similar comparison for the three spot volatility estimators, $\hat{\sigma}_{\text{Stein}}$, $\hat{\sigma}_{\text{Quad}}$, and $\hat{\sigma}_{\text{BLUE}}$, is provided in Fig. 3. The overall patterns generally mirror those of Fig. 2. Underscoring the difference between the two AMRE and the BLUE estimator of Li et al. (2022), which does not depend on a_i, the right panel clearly shows that the two optimal estimators both depend negatively, and nontrivially, on the asymmetry factor.

We turn next to a more direct assessment of how these differences in the functional forms of the estimators translate into asymptotic biases, variances, and differences in Stein and quadratic risks.

[17] A simplified "practical" version of the Garman–Klass estimator, defined by $0.5 w_i^2 - \big(2\log(2) - 1\big) r_i^2 \approx 0.5 w_i^2 - 0.386 r_i^2$, has also sometimes been used in empirical applications.

[18] The approximation for $\hat{\sigma}^2_{\text{Stein}}$ is constructed by projecting $\big(G_0(a_i/w_i) - H_0(|r_i|/w_i)\big)/\big(G_2(a_i/w_i) - H_2(|r_i|/w_i)\big)$ onto a cubic polynomial of the maximal invariant $(|r_i|/w_i, a_i/w_i)$ under the L_2 distance. The approximations for the other AMRE estimators are obtained similarly.

[19] As a point of reference, in the Brownian limit model, the interquartile range of $|r_i|/w_i$ spans 0.243 to 0.676, while the interdecile range covers 0.099 to 0.817.

[20] Since a_i measures the absolute difference between the lengths of the upper and lower shadows of the candlestick, it takes values in $[0, w_i - |r_i|]$.

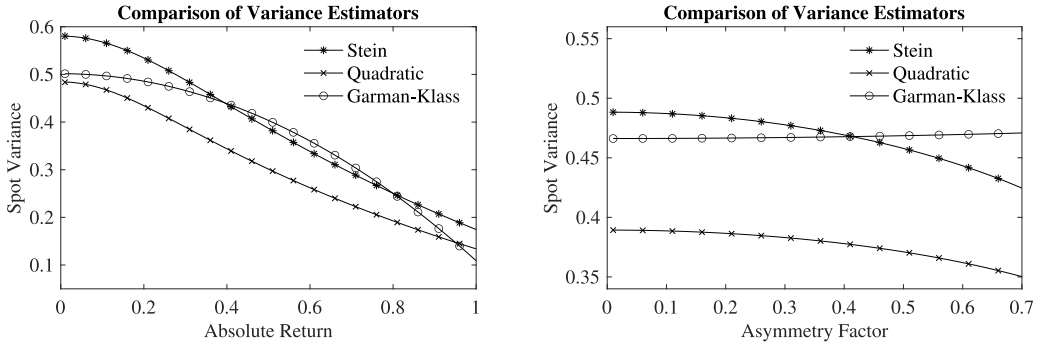

Fig. 2. Alternative range-based variance estimators. The figure plots the AMRE estimators for the variance under Stein's loss (Stein) and quadratic loss (Quadratic), together with the Garman–Klass estimator. The range w_i is normalized to unity. The left panel plots the spot variance estimator as a function of the absolute return $|r_i|$, with the asymmetry factor fixed at $a_i = 0$. The right panel plots the spot variance estimator as a function of the asymmetry factor a_i, with the absolute return fixed at $|r_i| = 0.3$.

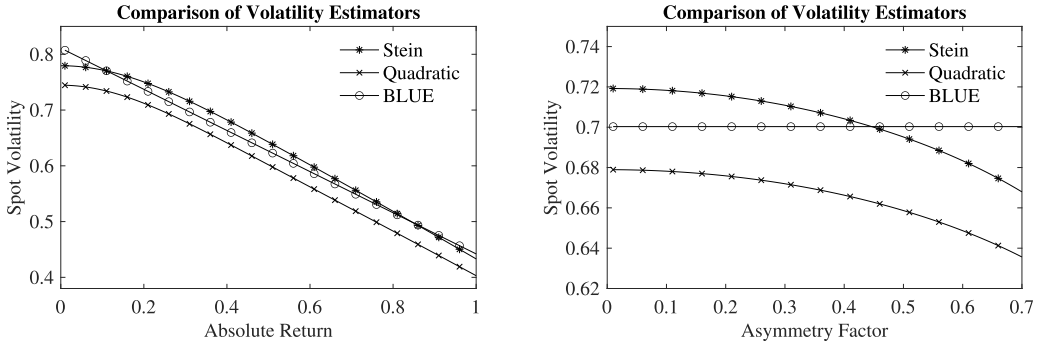

Fig. 3. Alternative range-based volatility estimators. The figure plots the AMRE estimators for the volatility under Stein's loss (Stein) and quadratic loss (Quadratic), together with the BLUE estimator proposed by Li et al. (2022). The range w_i is normalized to unity. The left panel plots the spot volatility estimator as a function of the absolute return $|r_i|$, with the asymmetry factor fixed at $a_i = 0$. The right panel plots the spot volatility estimator as a function of the asymmetry factor a_i, with the absolute return fixed at $|r_i| = 0.3$.

Table 1
Asymptotic properties of spot variance estimators.

Estimator	Bias	Variance	Relative efficiency	
			Stein	Quadratic
$\hat{\sigma}^2_{\text{Stein}}$	0.000	0.259	1.000	0.803
$\hat{\sigma}^2_{\text{Quad}}$	−0.205	0.165	0.813	1.000
$\hat{\sigma}^2_{\text{GK}}$	0.000	0.270	0.968	0.770
$(\hat{\sigma}_{\text{BLUE}})^2$	0.062	0.295	0.976	0.698

Note: The table reports the asymptotic biases, variances, and relative efficiency under Stein's and quadratic risks for each of the spot variance estimators indicated in the first column.

2.3. Risk comparisons

We will focus our comparisons of the risks of the different estimators by considering the relative efficiency, defined as the ratio between the risk of the relevant AMRE estimator and the estimator under consideration. Table 1 reports the results for the three spot variance estimators: $\hat{\sigma}^2_{\text{Stein}}$, $\hat{\sigma}^2_{\text{Quad}}$, and $\hat{\sigma}^2_{\text{GK}}$. Since $\hat{\sigma}_{\text{BLUE}}$ is the BLUE for spot volatility, we also include $(\hat{\sigma}_{\text{BLUE}})^2$ as a fourth contender for estimating the spot variance. Table 2 presents the analogous results for the three spot volatility estimators: $\hat{\sigma}_{\text{Stein}}$, $\hat{\sigma}_{\text{Quad}}$, and $\hat{\sigma}_{\text{BLUE}}$. For comparison, we also include the transformed $(\hat{\sigma}^2_{\text{GK}})^{1/2}$ spot volatility estimator.[21]

Looking first at Table 1, the optimal $\hat{\sigma}^2_{\text{Quad}}$ estimator exhibits substantial downward asymptotic bias. This "shrinkage" feature is attributable to the fact that the quadratic loss assigns a heavier penalty to overestimation than underestimation, and as such the corresponding optimal estimator naturally sacrifices some downward bias in order to further reduce the variance. Indeed, the

[21] All of the numbers are computed numerically based on ten million Monte Carlo draws of a standard Brownian motion $(B_t)_{t\in[0,1]}$ (recall (2.6)) as the simulated sample averages corresponding to $\mathbb{E}[f(\tilde{\zeta})] - 1$, $Var(f(\tilde{\zeta}))$, and $\mathbb{E}[L(f(\tilde{\zeta}))]$, respectively.

Table 2
Asymptotic properties of spot volatility estimators.

Estimator	Bias	Variance	Relative efficiency	
			Stein	Quadratic
$\hat{\sigma}_{\text{Stein}}$	0.000	0.061	1.000	0.967
$\hat{\sigma}_{\text{Quad}}$	−0.058	0.055	0.909	1.000
$(\hat{\sigma}^2_{\text{GK}})^{1/2}$	−0.030	0.060	0.938	0.952
$\hat{\sigma}_{\text{BLUE}}$	0.000	0.063	0.968	0.937

Note: The table reports the asymptotic biases, variances, and relative efficiency under Stein's and quadratic risks for each of the spot volatility estimators indicated in the first column.

asymptotic variance of the $\hat{\sigma}^2_{\text{Quad}}$ estimator is notably lower than the corresponding numbers for all of the other estimators. Compared to the optimal estimator $\hat{\sigma}^2_{\text{Quad}}$, the relative efficiencies of the shape-constrained $\hat{\sigma}^2_{\text{GK}}$ and $(\hat{\sigma}_{\text{BLUE}})^2$ estimators equal 77.0% and 69.8% respectively.[22]

Although the shape-constrained estimators clearly demonstrate suboptimal performance under quadratic loss, they exhibit "near-optimal" behavior under Stein's loss. Specifically, the relative efficiencies of $\hat{\sigma}^2_{\text{GK}}$ and $(\hat{\sigma}_{\text{BLUE}})^2$ are 96.8% and 97.6%, respectively, when compared to the AMRE $\hat{\sigma}^2_{\text{Stein}}$. In other words, in scenarios where an economic agent's loss function closely resembles Stein's loss, the classical Garman–Klass and the BLUE estimators are both reasonable practical choices. Of course, since the AMRE estimator can also easily be calculated in practice using our explicit closed form solution, there is really no need to suffer any efficiency loss, however small it might be.

Turning to Table 2 and spot volatility estimation, the results again evidence notable bias for the $\hat{\sigma}_{\text{Quad}}$ estimator. In general, the efficiency gaps between the shape-constrained volatility estimators and the optimal estimator are smaller than for variance estimation. Intuitively, the optimal estimation of σ_t is "easier" than the optimal estimation of its nonlinear transform σ_t^2.[23] As such, the Garman–Klass estimator and the simple linear estimator proposed by Li et al. (2022) turn out to perform quite well for spot volatility estimation under both quadratic and Stein's loss functions, although both estimators, strictly speaking, are inadmissible.

More broadly, these numerical comparisons also demonstrate that the relative asymptotic risks of alternative estimators, and in turn the design of optimal estimators, can depend quite strongly on the underlying loss function. This reflects the finite-sample nature of our coupling-based asymptotic analysis in a non-Gaussian limit experiment. By contrast, in the conventional "large sample" asymptotic setting with Gaussian shift limit experiments, different loss functions (as long as they are bowl-shaped) result in the same optimal estimators (see, e.g., Chapter 8 in Van der Vaart (1998)).

Acknowledging the practical challenge of precising the loss function in some applications, we observe an intriguing pattern for the risk comparisons in Tables 1 and 2. In both tables, the AMRE estimators derived under Stein's loss exhibit lower risks than the shape-constrained estimators, not only under Stein's loss (which holds by construction), but also under quadratic loss.[24] Hence, for users who are uncertain about their specific loss function, we recommend employing $\hat{\sigma}^2_{\text{Stein}}$ and $\hat{\sigma}_{\text{Stein}}$ as "general purpose" estimators for spot variance and spot volatility estimation, respectively.

2.4. Optimal estimators with multiple candlesticks

The estimators discussed above rely on a single candlestick for optimally estimating the spot volatility, or the spot variance. In this section we describe how to combine multiple adjacent candlesticks (over asymptotically shrinking time intervals) for optimally estimating σ_t^p.

To set out the notation, given a fixed integer $k \geq 1$, let $\boldsymbol{c}_i = (r_{i+j-1}, u_{i+j-1}, l_{i+j-1})_{1\leq j\leq k}$ collect the observed features of k successive candlesticks starting at the ith observation. Denote the corresponding estimator for σ_t^p that utilizes the k candlesticks by $f(\boldsymbol{c}_i)$. A direct extension of Theorem 1 produces the following analogous coupling result for the k-candlestick setting.

Corollary 1. *Under Assumption 1, given any fixed integer $k \geq 1$, any regular scale-equivariant estimator $S = f(\boldsymbol{c}_i)$ with $|i\Delta_n - t| \to 0$ may be expressed as*

$$\frac{S}{\sigma_t^p} = f(\zeta_i, \zeta_{i+1}, \dots, \zeta_{i+k-1}) + o_p(1),$$

where the variables ζ_{i+j}, $j = 1, \dots, k$, are defined as in Theorem 1.

[22] Interestingly, $\hat{\sigma}^2_{\text{Stein}}$ exhibits lower quadratic risk than $\hat{\sigma}^2_{\text{GK}}$. Since both of these two estimators are asymptotically unbiased, this suggests that under quadratic loss the Garman–Klass estimator is asymptotically inadmissible, not only within the class of regular scale-equivariant estimators, but also within the subclass of asymptotically unbiased estimators.

[23] Consistent with this intuition, Tables S1 in the online Supplemental Appendix shows that the efficiency gap between the AMRE estimators and the shape-constrained estimators is also larger for the quarticity σ_t^4, an even "more nonlinear" transform of the volatility than the variance. For example, under quadratic loss, the relative efficiencies of $(\hat{\sigma}^2_{\text{GK}})^2$ and $(\hat{\sigma}_{\text{BLUE}})^4$ are only 31.2% and 25.5%, respectively, in comparison with the AMRE estimator for σ_t^4.

[24] Additional results reported in the online Supplemental Appendix show that this phenomenon remains true for estimating the spot quarticity, σ_t^4, and the spot precision, σ_t^{-1}.

Table 3
Asymptotic properties of spot variance estimators based on two candlesticks.

Estimator	Bias	Variance	Relative efficiency	
			Stein	Quadratic
$\hat{\sigma}^2_{\text{Stein}}(2)$	0.000	0.128	1.000	0.891
$\hat{\sigma}^2_{\text{Quad}}(2)$	−0.103	0.103	0.833	1.000
$\hat{\sigma}^2_{\text{GK}}$	0.000	0.135	0.923	0.844
$(\hat{\sigma}_{\text{BLUE}})^2$	0.062	0.147	0.923	0.755

Note: The table reports the asymptotic biases, variances, and relative efficiency under Stein's and quadratic risks for each of the spot variance estimators indicated in the first column.

Building on the same reasoning outlined in Section 2.2, we may therefore couple the original estimation problem with that in the Brownian limit experiment. Moreover, by a direct extension of the proof of Theorem 2, we can also derive semi-closed form expressions for the AMRE estimators that utilize k successive candlesticks. Concretely, the AMRE estimators under Stein's loss and standardized quadratic loss may be expressed as

$$\begin{aligned}\hat{\sigma}^p_{\text{Stein}}(k) &= w_i^p \cdot \frac{1}{F_{k,p}\left(\frac{|r_i|}{w_i}, \frac{a_i}{w_i}, \frac{w_{i+1}}{w_i}, \frac{|r_{i+1}|}{w_i}, \frac{a_{i+1}}{w_i}, \dots, \frac{w_{i+k-1}}{w_i}, \frac{|r_{i+k-1}|}{w_i}, \frac{a_{i+k-1}}{w_i}\right)}, \\ \hat{\sigma}^p_{\text{Quad}}(k) &= w_i^p \cdot \frac{F_{k,p}\left(\frac{|r_i|}{w_i}, \frac{a_i}{w_i}, \frac{w_{i+1}}{w_i}, \frac{|r_{i+1}|}{w_i}, \frac{a_{i+1}}{w_i}, \dots, \frac{w_{i+k-1}}{w_i}, \frac{|r_{i+k-1}|}{w_i}, \frac{a_{i+k-1}}{w_i}\right)}{F_{k,2p}\left(\frac{|r_i|}{w_i}, \frac{a_i}{w_i}, \frac{w_{i+1}}{w_i}, \frac{|r_{i+1}|}{w_i}, \frac{a_{i+1}}{w_i}, \dots, \frac{w_{i+k-1}}{w_i}, \frac{|r_{i+k-1}|}{w_i}, \frac{a_{i+k-1}}{w_i}\right)}.\end{aligned} \tag{2.12}$$

The function $F_{k,q} : \mathbb{R}^{3k-1} \to \mathbb{R}$ that enters these expressions for $q \in \{p, 2p\}$ is formally defined as a conditional expectation function:

$$\begin{aligned}&F_{k,q}\left(\frac{|r_i|}{w_i}, \frac{a_i}{w_i}, \frac{w_{i+1}}{w_i}, \frac{|r_{i+1}|}{w_i}, \frac{a_{i+1}}{w_i}, \dots, \frac{w_{i+k-1}}{w_i}, \frac{|r_{i+k-1}|}{w_i}, \frac{a_{i+k-1}}{w_i}\right) \\ &\quad \equiv \mathbb{E}\left[\xi^q_{w,1} \middle| \frac{\xi_{w,j}}{\xi_{w,1}} = \frac{w_{i+j-1}}{w_i}, \frac{\xi_{r,j}}{\xi_{w,1}} = \frac{|r_{i+j-1}|}{w_i}, \frac{\xi_{a,j}}{\xi_{w,1}} = \frac{a_{i+j-1}}{w_i} \text{ for all } 1 \le j \le k\right],\end{aligned} \tag{2.13}$$

where $(\xi_{w,j}, \xi_{r,j}, \xi_{a,j})$, $j = 1, 2, \dots, k$, are independent copies of

$$\left(\sup_{t\in[0,1]} B_t - \inf_{t\in[0,1]} B_t, |B_1|, \left|\sup_{t\in[0,1]} B_t + \inf_{t\in[0,1]} B_t - B_1\right|\right). \tag{2.14}$$

In parallel to the optimal estimators that rely on a single candlestick, the two AMRE estimators defined in (2.12) are also structurally similar. The w_i^p component, in particular, acts as a generic scale-equivariant estimator for σ_i^p, while the $F_{k,q}(\cdot)$ function depends on candlestick observations solely through the maximal invariant statistics, defined as the shape features $(w_{i+j-1}, |r_{i+j-1}|, a_{i+j-1})_{1\le j\le k}$ normalized by w_i. The earlier closed form solutions for the single-candlestick case, or $k = 1$, detailed in Theorem 2 were obtained by explicitly deriving the functional form of $F_{1,q}(\cdot)$. Regrettably, analytical solutions for $F_{k,q}(\cdot)$ for $k \ge 2$ are currently unattainable.

Nonetheless, the semi-closed form solutions in (2.12) still suggest a strategy for numerically computing the optimal estimation functions. In particular, since $F_{k,q}(\cdot)$ is defined as the conditional expectation of w_i^q given the maximal invariant statistics under the limit experiment, one may simulate the $(\xi_{w,j}, \xi_{r,j}, \xi_{a,j})_{1\le j\le k}$ variables as i.i.d. copies of the Brownian functionals defined in (2.14) and then calculate the requisite conditional expectation function in (2.13) via Monte Carlo integration. This calculation, which formally entails the formation of a predictor that minimizes the mean squared error, may be conveniently implemented using popular machine learning tools such as neural networks, or random forests. We stress that for a given k and q, the function $F_{k,q}(\cdot)$ only needs to be computed once.

To illustrate the idea, consider the case with two candlesticks, or $k = 2$. Employing a neural network to compute the conditional expectation functions $F_{2,p}(\cdot)$ and $F_{2,2p}(\cdot)$ numerically, Tables 3 and 4 report the resulting asymptotic bias, variance, and relative efficiency for the AMRE estimators for estimating the spot variance and volatility, respectively, obtained by using these numerical approximations in place of the true unknown functions in (2.12).[25] The $k = 2$ versions of the shape-constrained $\hat{\sigma}^2_{\text{GK}}$ and $\hat{\sigma}_{\text{BLUE}}$ estimators, also included in the tables, are constructed as simple averages of their respective single-candlestick estimates, following the suggestion of Li et al. (2022).

The general results are qualitatively very similar to the ones for the single-candlestick estimators previously reported in Tables 1 and 2. The optimal estimators are notably more accurate, especially for estimating the spot variance under quadratic loss. At the same time, the "near optimality" property of the shape-constrained estimators under Stein's loss observed for the single-candlestick

[25] More specifically, we rely on a logistic sigmoid activation function, and an architecture comprised of an input layer with 32 neurons, followed by two hidden layers with 16 and 8 neurons, respectively. We train the model based on five million random draws of $(\xi_{w,j}, \xi_{r,j}, \xi_{a,j})_{j=1,2}$, where the Brownian motion $(B_t)_{t\in[0,1]}$ is generated using an Euler scheme with a mesh size of 10^{-7}. Underscoring the accuracy of the approach, using the same procedures to calculate the functions for $k = 1$ results in numerical solutions that are practically indistinguishable from the closed form solutions detailed in Theorem 2.

Table 4
Asymptotic properties of spot volatility estimators based on two candlesticks.

Estimator	Bias	Variance	Relative efficiency	
			Stein	Quadratic
$\hat{\sigma}_{\text{Stein}}(2)$	0.000	0.030	1.000	0.966
$\hat{\sigma}_{\text{Quad}}(2)$	−0.025	0.029	0.939	1.000
$(\hat{\sigma}^2_{\text{GK}})^{1/2}$	−0.030	0.030	0.940	0.935
$\hat{\sigma}_{\text{BLUE}}$	0.000	0.031	0.942	0.935

Note: The table reports the asymptotic biases, variances, and relative efficiency under Stein's and quadratic risks for each of the spot volatility estimators indicated in the first column.

case does not appear to hold as well. For instance, the relative efficiency of the Garman–Klass variance estimator drops from 96.8% in the $k = 1$ case to 92.3% in the $k = 2$ case, while the relative efficiency of the BLUE volatility estimator drops from 96.8% to 94.2%. These findings further motivate the use of the AMRE estimators in practice.

Putting the results in Tables 3 and 4 further into perspective, it is, of course, not surprising that the spot estimators constructed by combining two candlesticks exhibit smaller theoretical asymptotic variances than their single-candlestick counterparts. At the same time, the temporal aggregation of multiple candlesticks can easily be harmful in practice, especially when the volatility fluctuates rapidly. In such situations, the limit experiment with constant volatility that formally underlies the theoretical asymptotic comparisons will likely not provide a good finite-sample guide either. Of course, this type of empirical scenario is precisely when spot estimation can be most useful and informative. The empirical application discussed in the next section further highlights these issues.

3. An empirical illustration

To demonstrate the practical applicability and insights afforded by the new optimal estimators, we present spot volatility estimates for a market portfolio on the eight 2022 prescheduled FOMC announcement days. Putting the results into perspective, at the start of the year U.S. inflation had already soared to its highest level since the 1980s. In response to this, the Federal Reserve indicated at its January 2022 meeting that it would soon be appropriate to raise the target range for the federal funds rate. Subsequently, the target rate was indeed increased by 25 basis points (bps) in March, followed by a more substantial 50 bps hike in May. The pace of rate increases further accelerated to 75 bps for the next four meetings, before moderating to a 50 bps rise at the final 2022 meeting in December. Each of these rate increases were detailed in a short formal release by the FOMC at exactly 14:00 EST, followed by additional comments and a press conference led by Federal Reserve Chairman Jerome Powell starting half-an-hour later.

It is well established that financial markets often reacts quite strongly to the initial 14:00 FOMC announcement.[26] It is much less clear, however, what happens to market volatility at the exact time of the FOMC announcement, let alone in its immediate aftermath and during the subsequent press conference. We shed new light on this issue by utilizing intraday candlestick data for the S&P 500 index, in the form of the VOO exchange-traded fund (ETF) managed by the Vanguard Group, to estimate high-frequency spot volatilities. Guided by the simulation results discussed above, to mitigate the impact of microstructure noise, we employ 5-min VOO candlesticks, sourced directly from Bloomberg. We focus our analyses on the 5-min $\hat{\sigma}_{\text{Stein}}$ AMRE spot volatility estimator; comparisons with the other estimators discussed above are presented in the online Supplemental Appendix.[27]

Fig. 4 displays the resulting estimates, together with 90% level pointwise confidence intervals. To facilitate comparison across the different days, all of the plots are presented on a uniform percentage daily scale. As the figure shows, the market volatility generally spikes immediately following the initial FOMC announcement at 14:00.[28] The volatility then generally reverts towards a more "normal" level in the half-hour window between the initial release and the start of the press conference. By comparison, the volatility patterns observed during the press conference appear less systematic. In addition to reiterating key policy decisions, also summarized in the initial release, the press conference and the subsequent interaction with the media often provide additional forward guidance about future Fed policies, interspersed with comments about the general economic outlook as perceived by the Fed. The staggered information flow delivery throughout this process naturally manifest in event specific volatility spikes linked to the exact timing of the new information.

[26] Andersen et al. (2007), Lee and Mykland (2008), Lee (2012), and Bollerslev et al. (2018), among others, have previously associated high-frequency jumps in asset prices with FOMC announcements. FOMC announcements have also been used as a powerful tool for the high-frequency identification of monetary policy shocks, as exemplified by Cochrane and Piazzesi (2002), Rigobon and Sack (2004), Bernanke and Kuttner (2005), and Nakamura and Steinsson (2018), while Savor and Wilson (2014), Lucca and Moench (2015), Cieslak et al. (2019), and Ai and Bansal (2018) have emphasized the significance of an FOMC announcement risk premium and pre-announcement drifts.

[27] The online Supplemental Appendix also reports analogous results for the Dollar/Yen exchange rate.

[28] Interestingly, for some of the days, most notably March 16, May 4, and June 15, the volatility actually increased slightly in advance of the official 14:00 announcement. Whether these "abnormal" pre-announcement increases can be attributed to information leaks during the Fed's official blackout period may warrant further scrutiny.

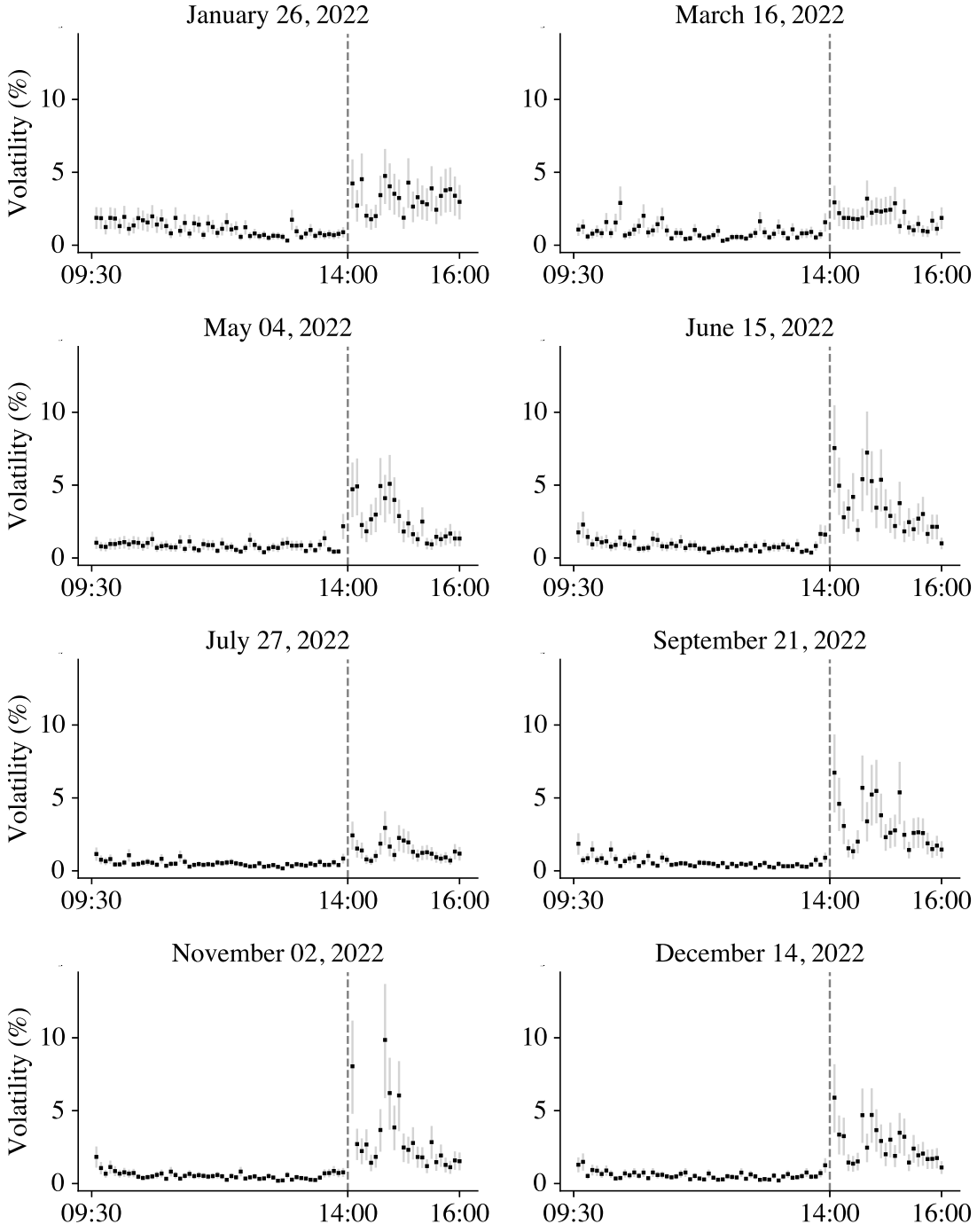

Fig. 4. Spot volatility estimates for the VOO ETF on FOMC announcement days. The figure plots the $\hat{\sigma}_{\text{Stein}}$ spot volatility estimates based on 5-min VOO ETF candlesticks, expressed in daily percentage units. Pointwise confidence intervals at the 90% level is calculated as detailed in footnote 11. The vertical lines included in each of the panels indicate the official 14:00 FOMC announcement times.

The November 2 announcement provides an interesting case in point. In line with the general pattern noted above, the spot volatility shows an initial burst at 14:00, followed by a gradual decline to a lower, albeit still elevated, level at 14:30. Then, concurrent with the start of Chairman Powell's speech, the 14:30–14:35 volatility estimate increased moderately, reflecting the limited new information presented in the opening, relatively structured, portion of the speech. This modest uptick is then followed by a dramatic volatility spike, of even greater magnitude than the initial surge that accompanied the 14:00 announcement. This volatility spike coincided with the time at which Powell concluded his opening remarks and began the press conference by

mentioning that the ultimate level of interest rates would be "*higher than previously expected*".[29] Powell's brief interaction with the media shortly thereafter further underscored the looming uncertainty surrounding the central bank's monetary tightening agenda.[30] These comments on the likely trajectory of monetary policy offered crucial forward guidance, and their unexpected nature in effect amounted to a "policy shock" resulting in a sharp increase in market volatility at the time.

In sum, asset price volatilities often experience large changes over short time windows in response to the release of important new economic information. The new optimal high-frequency candlestick estimators developed here allows for meaningful estimation of such changes, which would otherwise be obscured by the use of longer estimation windows, in turn affording a more nuanced depiction and better understanding of the economic mechanisms at work.

4. Concluding remarks

We develop a new class of optimal range-based nonparametric volatility estimators. The new optimal estimators are explicitly geared to the volatility object of interest and relevant loss function. They involve complex, yet closed-form and easy-to-evaluate, nonlinear functions of the range, the absolute return, and a measure of asymmetry. The efficiency gains provided by the new estimators compared to currently used suboptimal range-based estimators rooted in ad hoc functional-form assumptions can be substantial.

Looking ahead, the same infill asymptotic decision-theoretic framework developed here, based on coupling the nonparametric volatility estimation problem with a finite-sample optimal estimation problem, could possibly be adapted to study other outstanding optimal nonparametric inference problems. High-frequency range-based estimators have also previously been used for the estimation of integrated volatility over non-trivial time intervals (e.g., Christensen and Podolskij (2007)). The new optimal estimators developed here could similarly be employed for that purpose, as well as the estimation of other volatility functionals. The integrated quarticity, in particular, has proven notoriously difficult to accurately estimate in practice, yet it plays a crucial role in assessing the estimation error of traditional realized volatility type estimators (e.g., Barndorff-Nielsen and Shephard (2002) and Bollerslev et al. (2016)). Prior empirical uses of range-based volatility estimators for modeling and forecasting time-varying volatility abounds (early contributions include Gallant et al. (1999) and Alizadeh et al. (2002)). The range-based estimators developed here may naturally be used in that context as well for obtaining more accurate inference. They could also help sharpen the inference in the recent and growing literature that relies on high-frequency identification through heteroskedasticity (following Rigobon (2003)), and volatilities being higher over short "treatment" windows following news events (e.g., Nakamura and Steinsson (2018) and Bollerslev et al. (2018)). We leave further work on all of these theoretical and more empirically oriented issues for future research.

Appendix A. Proofs

Proof of Theorem 1. In this proof, we focus on a specific time point, denoted as t, and examine i_n such that $i_n\Delta_n = t + o(1)$. To simplify our notation, we write i instead of i_n. We employ a generic constant $K > 0$, which may vary throughout the proof. Relying on a standard localization technique, we can strengthen Assumption 1 by assuming that the boundedness conditions hold uniformly over the whole sample. For more details on the localization method, refer to Section 4.4.1 in Jacod and Protter (2012).

Under Assumption 1(i), the probability of the interval $\mathcal{I}_i$ containing at least one price jump is $O(\Delta_n)$. Consequently, price jumps occur in $\mathcal{I}_i$ with asymptotically negligible probability. As our analysis focuses on this particular interval, we can assume without loss of generality that there are no jumps.

Denote $r_i' \equiv \sigma_{(i-1)\Delta_n}\zeta_{i,r}$, $u_i' \equiv \sigma_{(i-1)\Delta_n}\zeta_{i,u}$, and $l_i' \equiv \sigma_{(i-1)\Delta_n}\zeta_{i,l}$. Since there is no jump within the $\mathcal{I}_i$ interval, we can rewrite (r_i, u_i, l_i) as

$$
\begin{aligned}
r_i &= \Delta_n^{-1/2}\left(\int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s\right),\\
u_i &= \Delta_n^{-1/2}\sup_{s\in\mathcal{I}_i}\left(\int_{(i-1)\Delta_n}^{s} b_u du + \int_{(i-1)\Delta_n}^{s} \sigma_u dW_u\right),\\
l_i &= \Delta_n^{-1/2}\inf_{s\in\mathcal{I}_i}\left(\int_{(i-1)\Delta_n}^{s} b_u du + \int_{(i-1)\Delta_n}^{s} \sigma_u dW_u\right).
\end{aligned}
$$

Under Assumption 1(i), it is easy to see that

$$
\left|\int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds\right| \le \int_{(i-1)\Delta_n}^{i\Delta_n} |b_s| ds = O_p(\Delta_n). \tag{A.1}
$$

Moreover, by the Burkholder–David–Gundy inequality and Assumption 1(ii), we have

$$
\begin{aligned}
\mathbb{E}\left[\sup_{s\in\mathcal{I}_i}\left|\int_{(i-1)\Delta_n}^{s}(\sigma_u - \sigma_{(i-1)\Delta_n})dW_u\right|^2\right] &\le K\mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n}|\sigma_u - \sigma_{(i-1)\Delta_n}|^2 du\right]\\
&\le K\Delta_n^{1+2\kappa},
\end{aligned}
$$

[29] A complete transcript of Powell's statement is available at: https://www.federalreserve.gov/monetarypolicy/fomcpresconf20221102.htm.

[30] The first two questions from the media, posed by Colby Smith of the *Financial Times* and Howard Schneider of Reuters, respectively, also both concerned the potential slowdown of future rate increases.

and hence,

$$\sup_{s\in\mathcal{T}_i}\left|\int_{(i-1)\Delta_n}^{s}(\sigma_u-\sigma_{(i-1)\Delta_n})dW_u\right| = O_p(\Delta_n^{1/2+\kappa}). \tag{A.2}$$

By the triangle inequality, (A.1), and (A.2),

$$|r_i - r_i'| \le \Delta_n^{-1/2}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds\right| + \Delta_n^{-1/2}\left|\int_{(i-1)\Delta_n}^{i\Delta_n}(\sigma_s-\sigma_{(i-1)\Delta_n})dW_s\right| = O_p(\Delta_n^{(1/2)\wedge\kappa}). \tag{A.3}$$

In addition, we note that

$$\begin{aligned}|u_i - u_i'| &= \Delta_n^{-1/2}\left|\sup_{s\in\mathcal{T}_i}\left(\int_{(i-1)\Delta_n}^{s} b_u du + \int_{(i-1)\Delta_n}^{s}\sigma_u dW_u\right)\right.\\ &\qquad\left. - \sigma_{(i-1)\Delta_n}\sup_{s\in\mathcal{T}_i}(W_s - W_{(i-1)\Delta_n})\right|\\ &\le \Delta_n^{-1/2}\left(\int_{(i-1)\Delta_n}^{i\Delta_n}|b_u|du + \sup_{s\in\mathcal{T}_i}\left|\int_{(i-1)\Delta_n}^{s}(\sigma_u-\sigma_{(i-1)\Delta_n})dW_u\right|\right)\\ &= O_p(\Delta_n^{(1/2)\wedge\kappa}),\end{aligned} \tag{A.4}$$

where the last line follows from (A.1) and (A.2). Similarly, we can derive

$$|l_i - l_i'| = O_p(\Delta_n^{(1/2)\wedge\kappa}). \tag{A.5}$$

Since $f(\cdot)$ is continuous a.e., the estimates from (A.3)–(A.5) imply that $f(r_i,u_i,l_i) = f(r_i',u_i',l_i') + o_p(1)$. Since σ is bounded away from zero under Assumption 1(i), we further have

$$\frac{f(r_i,u_i,l_i)}{\sigma^p_{(i-1)\Delta_n}} = \frac{f(r_i',u_i',l_i')}{\sigma^p_{(i-1)\Delta_n}} + o_p(1) = f(\zeta_i) + o_p(1), \tag{A.6}$$

where the second equality follows from the homogeneity of $f(\cdot)$ and the definition of ζ_i. By Assumption 1(ii), $|\sigma_t - \sigma_{(i-1)\Delta_n}| = O_p(|t-i\Delta_n|^\kappa) = o_p(1)$ as $|i\Delta_n - t| \to 0$, which together with (A.6) implies the assertion of Theorem 1. □

To prove Theorem 2, we first prove two lemmas. Lemma A.1 shows the sufficiency of the shape features for volatility estimation under the limit model. Lemma A.2 derives a closed-form expression for the conditional expectation of certain Brownian functionals.

Lemma A.1. *Under the limit model $P_t = \sigma W_t$, the collection of shape features $(|r_i|, w_i, a_i)$ is a sufficient statistic for σ given the observation (r_i,u_i,l_i).*

Proof of Lemma A.1. Recall that $(B_t)_{t\in[0,1]}$ is a standard Brownian motion with $B_0 = 0$. Let $g(r,u,l)$ denote the probability density function of $(B_1, \sup_{0\le t\le 1} B_t, \inf_{0\le t\le 1} B_t)$. The density of (r_i,u_i,l_i) is then given by the function

$$(r,u,l) \mapsto \frac{1}{\sigma^3} g\Big(\frac{r}{\sigma},\frac{u}{\sigma},\frac{l}{\sigma}\Big). \tag{A.7}$$

By equation (1.15.8) in Borodin and Salminen (2002),

$$\begin{aligned}&\mathbb{P}\Big(l < \inf_{0\le t\le 1} B_t,\ \sup_{0\le t\le 1} B_t < u,\ B_1 \in dr\Big)\\ &\quad= \frac{1}{\sqrt{2\pi}}\sum_{k=-\infty}^{\infty}\left\{\exp\Big(-\frac{(2k(u-l)+r)^2}{2}\Big) - \exp\Big(-\frac{(2k(u-l)+r-2l)^2}{2}\Big)\right\}dr.\end{aligned} \tag{A.8}$$

The function $g(r,u,l)$ is thus proportional to $\sum_{k=-\infty}^{\infty}\big(A_k(r,u,l) - B_k(r,u,l)\big)$, where

$$\begin{aligned}A_k(r,u,l) &\equiv k^2\{(2k(u-l)+r)^2 - 1\}\exp\Big(-\frac{(2k(u-l)+r)^2}{2}\Big),\\ B_k(r,u,l) &\equiv k(k+1)\{(2k(u-l)+r-2l)^2 - 1\}\exp\Big(-\frac{(2k(u-l)+r-2l)^2}{2}\Big).\end{aligned}$$

By a change of variable via $w = u - l$ and $d = u + l - r$, we may identify these functions with

$$\begin{aligned}\tilde{A}_k(r,w,d) &\equiv k^2\{(2kw+r)^2 - 1\}\exp\Big(-\frac{(2kw+r)^2}{2}\Big),\\ \tilde{B}_k(r,w,d) &\equiv k(k+1)\{(2kw+w-d)^2 - 1\}\exp\Big(-\frac{(2kw+w-d)^2}{2}\Big).\end{aligned}$$

Note that for each $k \ge 0$, $\tilde{A}_k(-r,w,d) = \tilde{A}_{-k}(r,w,d)$ and $\tilde{B}_k(r,w,-d) = \tilde{B}_{-k-1}(r,w,d)$. Thus, $g(r,u,l)$ depends on (r,u,l) only through $(|r|,w,|d|)$. The assertion of the lemma then follows from the Fisher–Neyman factorization theorem. □

Lemma A.2. *Let B be a standard Brownian motion on the unit interval with $B_0 = 0$ and*

$$\xi_1 \equiv \sup_t B_t - \inf_t B_t, \quad \xi_2 \equiv \frac{|\sup_t B_t + \inf_t B_t - B_1|}{\sup_t B_t - \inf_t B_t}, \quad \xi_3 \equiv \frac{|B_1|}{\sup_t B_t - \inf_t B_t},$$

where $\sup_t$ and $\inf_t$ are taken over $[0,1]$. Then for each integer $q \geq 1$ we have

$$\mathbb{E}[\xi_1^q | \xi_2, \xi_3] = (-1)^q \frac{(q+2)}{\sqrt{2^q \pi} q!} \Gamma\Big(\frac{q+3}{2}\Big) \frac{G_q(\xi_2) - H_q(\xi_3)}{G_0(\xi_2) - H_0(\xi_3)},$$

where $G_q(\cdot)$ and $H_q(\cdot)$ are defined as in (2.10) and (2.11).

Proof of Lemma A.2. Let $g_\xi(\cdot)$ denote the joint density of (ξ_1, ξ_2, ξ_3). The conditional expectation of interest can then be written as

$$\mathbb{E}[\xi_1^q | \xi_2 = y, \xi_3 = z] = \frac{\int_0^\infty x^q g_\xi(x, y, z) dx}{\int_0^\infty g_\xi(x, y, z) dx}.$$

The main task is to calculate the numerator $\int_0^\infty x^q g_\xi(x, y, z) dx$ for $q \geq 1$ and the denominator $\int_0^\infty g_\xi(x, y, z) dx$. (The calculation for the latter is not a special case of the former by simply setting $p = 0$, as it requires a slightly more refined technical argument due to the lack of convergence of certain series.)

We first calculate $\int_0^\infty x^q g_\xi(x, y, z) dx$. From (A.8) and the definition of (ξ_1, ξ_2, ξ_3), we obtain $g_\xi(x, y, z) = 4\sqrt{2/\pi} \sum_{k=-\infty}^\infty \big(C_k(x, z) - D_k(x, y)\big)$, where

$$C_k(x, z) \equiv k^2 x^2 \big((2k+z)^2 x^2 - 1\big) \exp\Big(-\frac{(2k+z)^2 x^2}{2}\Big),$$
$$D_k(x, y) \equiv k(1+k) x^2 \big((2k+1-y)^2 x^2 - 1\big) \exp\Big(-\frac{(2k+1-y)^2 x^2}{2}\Big).$$

Since $z \in [0, 1]$, for $q \geq 1$, by a direct integration, we have

$$\int_0^\infty \sum_{k=-\infty}^\infty x^q C_k(x, z) dx = M_q \cdot \sum_{k=-\infty}^\infty \frac{k^2}{|\frac{z}{2} + k|^{q+3}}, \tag{A.9}$$

where we denote $M_r \equiv 2^{-(r+5)/2}(r+2)\Gamma\big((r+3)/2\big)$ for any $r \geq 0$. (Note that the convergence of the above series requires the integer $q \geq 1$.) We now express (A.9) using polygamma functions. By (5.15.1) in Olver et al. (2010), for $r \geq 1$ we have

$$\psi_r\Big(\frac{z}{2}\Big) = \sum_{k=0}^\infty \frac{(-1)^{r+1} r!}{(\frac{z}{2} + k)^{r+1}}. \tag{A.10}$$

Note that when $k \geq 0$, the summand in (A.9) may be rewritten in the form of the summand in (A.10) as

$$\frac{k^2}{(\frac{z}{2} + k)^{q+3}} = (-1)^{q+1} \Bigg(\frac{1}{q!} \cdot \frac{(-1)^{q+1} q!}{(\frac{z}{2} + k)^{q+1}} + \frac{1}{(q+1)!} \cdot z \frac{(-1)^{q+2}(q+1)!}{(\frac{z}{2} + k)^{q+2}} + \frac{1}{4(q+2)!} \cdot z^2 \frac{(-1)^{q+3}(q+2)!}{(\frac{z}{2} + k)^{q+3}} \Bigg). \tag{A.11}$$

Combining (A.9)–(A.11) yields

$$\sum_{k=0}^\infty \int_0^\infty x^q C_k(x, z) dx = \frac{(-1)^{q+1} M_q}{q!} \Big(\psi_q\Big(\frac{z}{2}\Big) + \frac{1}{q+1} z \psi_{q+1}\Big(\frac{z}{2}\Big) + \frac{1}{4(q+1)(q+2)} z^2 \psi_{q+2}\Big(\frac{z}{2}\Big) \Big).$$

The summation in (A.9) over $k < 0$ can be rewritten, with a change of variable $m = -k - 1$, as

$$\sum_{k=-\infty}^{-1} \frac{k^2}{(-\frac{z}{2} - k)^{q+3}} = \sum_{m=0}^\infty \frac{(m+1)^2}{(1 - \frac{z}{2} + m)^{q+3}}.$$

Using an argument similar to (A.11), we also have

$$\sum_{k=-\infty}^{-1} \int_0^\infty x^q C_k(x, z) dx = \frac{(-1)^{q+1} M_q}{q!} \Big(\psi_q\Big(1 - \frac{z}{2}\Big) - \frac{1}{q+1} z \psi_{q+1}\Big(1 - \frac{z}{2}\Big) + \frac{1}{4(q+1)(q+2)} z^2 \psi_{q+2}\Big(1 - \frac{z}{2}\Big) \Big).$$

Combining the above results for $k \geq 0$ and $k < 0$ and recalling the definition of $H_q(\cdot)$, we obtain

$$\int_0^\infty \sum_{k=-\infty}^\infty x^q C_k(x, z) dx = \frac{(-1)^{q+1} M_q}{q!} H_q(z). \tag{A.12}$$

By a similar argument leading to (A.12), we can also show that

$$\int_0^\infty \sum_{k=-\infty}^{\infty} x^q D_k(x,y)dx = \frac{(-1)^{q+1}M_q}{q!}G_q(y).$$

Hence, for $q \geq 1$,

$$\int_0^\infty x^q g_\xi(x,y,z)dx = 4\sqrt{\frac{2}{\pi}}\frac{(-1)^q M_q}{q!}\big(G_q(y) - H_q(z)\big). \tag{A.13}$$

For the denominator $\int_0^\infty g_\xi(x,y,z)dx$, by a direct integration, we have

$$\int_0^\infty \sum_{k=-\infty}^{\infty} \big(C_k(x,z) - D_k(x,y)\big)dx = \frac{\sqrt{2\pi}}{8}\sum_{k=-\infty}^{\infty}\left(\frac{k^2}{|\frac{z}{2}+k|^3} - \frac{k(k+1)}{|\frac{1-y}{2}+k|^3}\right). \tag{A.14}$$

By (5.7.6) in Olver et al. (2010), we obtain

$$\psi_0\Big(\frac{1-y}{2}\Big) - \psi_0\Big(\frac{z}{2}\Big) = \sum_{k=0}^{\infty}\left(\frac{1}{\frac{z}{2}+k} - \frac{1}{\frac{1-y}{2}+k}\right). \tag{A.15}$$

Note that when $k \geq 0$, the summand in (A.14) may be rewritten in the form of the summand in (A.10) and (A.15) as

$$\begin{aligned}\frac{k^2}{(\frac{z}{2}+k)^3} - \frac{k(k+1)}{(\frac{1-y}{2}+k)^3} &= \left(\frac{1}{\frac{z}{2}+k} - \frac{1}{\frac{1-y}{2}+k}\right) - z\frac{1}{(\frac{z}{2}+k)^2} - y\frac{1}{(\frac{1-y}{2}+k)^2} \\ &\quad - \frac{1}{8}z^2\frac{-2}{(\frac{z}{2}+k)^3} - \frac{1}{8}(1-y^2)\frac{-2}{(\frac{1-y}{2}+k)^3}.\end{aligned} \tag{A.16}$$

Combining (A.10) and (A.14)–(A.16) yields

$$\begin{aligned}\sum_{k=0}^{\infty}\int_0^\infty \big(C_k(x,z) - D_k(x,z)\big)dx = \frac{\sqrt{2\pi}}{8}\Big(&\psi_0\Big(\frac{1-y}{2}\Big) - \psi_0\Big(\frac{z}{2}\Big) - z\psi_1\Big(\frac{z}{2}\Big) - y\psi_1\Big(\frac{1-y}{2}\Big) \\ &- \frac{1}{8}z^2\psi_2\Big(\frac{z}{2}\Big) - \frac{1}{8}(1-y^2)\psi_2\Big(\frac{1-y}{2}\Big)\Big).\end{aligned}$$

The summation in (A.14) over $k < 0$ can be rewritten, with a change of variable $m = -k-1$, as

$$\sum_{k=-\infty}^{-1}\left(\frac{k^2}{(-\frac{z}{2}-k)^3} - \frac{k(k+1)}{(-\frac{1-y}{2}-k)^3}\right) = \sum_{m=0}^{\infty}\left(\frac{(m+1)^2}{(1-\frac{z}{2}+m)^3} - \frac{m(m+1)}{(\frac{1+y}{2}+m)^3}\right).$$

Using an argument similar to (A.16), we also have

$$\begin{aligned}&\sum_{k=-\infty}^{-1}\int_0^\infty \big(C_k(x,z) - D_k(x,z)\big)dx \\ &\quad= \frac{\sqrt{2\pi}}{8}\Big(\psi_0\Big(\frac{1+y}{2}\Big) - \psi_0\Big(1-\frac{z}{2}\Big) + z\psi_1\Big(1-\frac{z}{2}\Big) + y\psi_1\Big(\frac{1+y}{2}\Big) \\ &\qquad - \frac{1}{8}z^2\psi_2\Big(1-\frac{z}{2}\Big) - \frac{1}{8}(1-y^2)\psi_2\Big(\frac{1+y}{2}\Big)\Big).\end{aligned}$$

Combining the above results for $k \geq 0$ and $k < 0$ and recalling the definition of $G_0(\cdot)$, $H_0(\cdot)$, and M_0, we obtain

$$\int_0^\infty g_\xi(x,y,z)dx = 4\sqrt{\frac{2}{\pi}}M_0\big(G_0(y) - H_0(z)\big). \tag{A.17}$$

The assertion of the lemma then readily follows from (A.13), (A.17), and the fact that

$$\frac{M_q}{M_0} = \frac{(q+2)}{\sqrt{2^q\pi}}\Gamma\Big(\frac{q+3}{2}\Big). \quad \square$$

Proof of Theorem 2. We first consider the case with Stein's loss. Recall that the asymptotic risk $\mathbb{E}\big[L\big(f(\tilde{\zeta})\big)\big]$ equals the finite-sample risk of the estimator $f(r_i,u_i,l_i)$ under the limit model $P_t = \sigma W_t$. Therefore, minimizing the asymptotic risk is equivalent to finding the minimum-risk scale-equivariant estimator for σ under the limit model. By Lemma A.1 and the Rao–Blackwell theorem, this optimal estimator only depends on the shape features $(|r_i|, w_i, a_i)$. Note that $(w_i, a_i/w_i, |r_i|/w_i)$ has the same distribution as (ξ_1,ξ_2,ξ_3) defined in Lemma A.2. With an appeal to Corollary 3.3.8 in Lehmann and Casella (1998), the minimum-risk scale-equivariant estimation function under the limit problem is given by $w_i^p/\mathbb{E}[\xi_1^p|\xi_2 = a_i/w_i, \xi_3 = |r_i|/w_i]$. For estimating spot variance, taking $p = 2$ and applying Lemma A.2 with $q = 2$, we may rewrite this function in closed form as

$$\frac{4w_i^2}{3}\cdot\frac{G_0(a_i/w_i) - H_0(|r_i|/w_i)}{G_2(a_i/w_i) - H_2(|r_i|/w_i)}.$$

Recalling the meromorphic property of the polygamma functions, we see that this estimation function is continuous almost everywhere. This estimator is thus regular, and so, is also the AMRE estimator under the original nonparametric model as asserted in Theorem 2.

The proof for the quadratic loss is similar, except that we now apply (3.3.18) in Lehmann and Casella (1998) and Lemma A.2 above with $p = 2$, $q = 2$ and 4 to show that the optimal estimation function of σ^2 under the limit model is

$$w_i^2 \cdot \frac{\mathbb{E}[\xi_1^2|\xi_2 = a_i/w_i, \xi_3 = |r_i|/w_i]}{\mathbb{E}[\xi_1^4|\xi_2 = a_i/w_i, \xi_3 = |r_i|/w_i]} = \frac{32w_i^2}{5} \cdot \frac{G_2(a_i/w_i) - H_2(|r_i|/w_i)}{G_4(a_i/w_i) - H_4(|r_i|/w_i)}.$$

This estimator is also regular and thus is the AMRE estimator under the original nonparametric model as asserted.

The derivation of AMRE estimators of spot volatility follows the same lines of arguments, except for taking $p = 1$, $q = 1$ and 2. □

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2023.105548.

References

Ai, H., Bansal, R., 2018. Risk preferences and the macroeconomic announcement premium. Econometrica 86 (4), 1383–1430.
Aït-Sahalia, Y., Xiu, D., 2019. A Hausman test for the presence of market microstructure noise in high frequency data. J. Econmerrics 211 (1), 176–205.
Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002. Range-based estimation of stochastic volatility models. J. Finance 57 (3), 1047–1091.
Andersen, T.G., Bollerslev, T., 2018. Volatility. Edward Elgar Publishing.
Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Disentangling continuous and jump components in measuring, modeling and forecasting asset return volatility. Rev. Econ. Stat. 89 (4), 701–720.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2000. Great realizations. Risk 13, 105–108.
Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realised volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (2), 253–280.
Berkson, J., 1950. Are there two regressions? J. Amer. Statist. Assoc. 45 (250), 164–180.
Bernanke, B.S., Kuttner, K.N., 2005. What explains the stock market's reaction to federal reserve policy? J. Finance 60 (3), 1221–1257.
Bollerslev, T., Li, J., Liao, Z., 2021. Fixed-k inference for volatility. Quant. Econ. 12 (4), 1053–1084.
Bollerslev, T., Li, J., Xue, Y., 2018. Volume, volatility, and public news announcements. Rev. Econom. Stud. 85 (4), 2005–2041.
Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2016. Exploiting the errors: A simple approach for improved volatility forecasting. J. Econometrics 192 (1), 1–18.
Borodin, A.N., Salminen, P., 2002. Handbook of Brownian Motion-Facts and Formulae. Birkhäuser, Basel.
Brown, L., 1968. Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. Ann. Math. Stat. 39 (1), 29–48.
Christensen, K., Podolskij, M., 2007. Realized range-based estimation of integrated variance. J. Econometrics 141 (2), 323–349.
Cieslak, A., Morse, A., Vissing-Jorgensen, A., 2019. Stock returns over the FOMC cycle. J. Finance 74 (5), 2201–2248.
Cochrane, J.H., Piazzesi, M., 2002. The fed and interest rates - A high-frequency identification. Amer. Econ. Rev. 92 (2), 90–95.
Comte, F., Renault, E., 1998. Long memory in continuous-time stochastic volatility models. Math. Finance 8 (4), 291–323.
Diebold, F.X., Strasser, G., 2013. On the correlation structure of microstructure noise: A financial economic approach. Rev. Econom. Stud. 80 (4), 1304–1337.
Foster, D.P., Nelson, D.B., 1996. Continuous record asymptotics for rolling sample variance estimators. Econometrica 64 (1), 139–174.
Gallant, A.R., Te Hsu, C., Tauchen, G., 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. Rev. Econ. Stat. 81 (4), 617–631.
Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 53 (1), 67–78.
Gatheral, J., Jaisson, T., Rosenbaum, M., 2018. Volatility is rough. Quant. Finance 18 (6), 933–949.
Hájek, J., 1972. Local asymptotic minimax and admissibility in estimation. In: Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1. pp. 175–194.
Hansen, B.E., 2022. A modern Gauss-Markov theorem. Econometrica 90, 1283–1294.
Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. J. Bus. Econom. Statist. 24 (2), 127–161.
Hyslop, R., Imbens, G.W., 2001. Bias from classical and other forms of measurement error. J. Bus. Econom. Statist. 19 (4), 475–481.
Jacod, J., Li, J., Liao, Z., 2021. Volatility coupling. Ann. Statist. 49 (4), 1982–1998.
Jacod, J., Li, Y., Zheng, X., 2017. Statistical properties of microstructure noise. Econometrica 85 (4), 1133–1174.
Jacod, J., Protter, P., 2012. Discretization of Processes, Vol. 67. Springer Science & Business Media.
Jacod, J., Rosenbaum, M., 2013. Quarticity and other functionals of volatility: Efficient estimation. Ann. Statist. 41 (3), 1462–1484.
Jeganathan, P., 1982. On the convergence of moments of statistical estimators. Sankhyā: A 213–232.
Jeganathan, P., 1983. Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. Sankhyā: A 66–87.
Kalnina, I., Linton, O., 2008. Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. J. Econometrics 147 (1), 47–59.
Kristensen, D., 2010. Nonparametric filtering of the realized spot volatility: A kernel-based approach. Econom. Theory 26 (1), 60–93.
Le Cam, L., 1960. Locally asymptotically normal families of distributions. Univ. Calif. Publ. Stat. 3, 37–98.
Le Cam, L., 1986. Asymptotic Theory of Statistical Inference. John Wiley & Sons.
Lee, S., 2012. Jumps and information flow in financial markets. Rev. Financ. Stud. 25 (2), 439–479.
Lee, S., Mykland, P., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Rev. Financ. Stud. 21 (6), 2535–2563.
Lehmann, E.L., Casella, G., 1998. Theory of Point Estimation. In: Springer Texts in Statistics.
Li, Z.M., Linton, O., 2022. A ReMeDI for microstructure noise. Econometrica 90 (1), 367–389.
Li, J., Liu, Y., 2021. Efficient estimation of integrated volatility functionals under general volatility dynamics. Econom. Theory 37 (4), 664–707.
Li, J., Wang, D., Zhang, Q., 2022. Reading the candlesticks: An OK estimator for volatility. Rev. Econ. Stat. (forthcoming).
Lucca, D.O., Moench, E., 2015. The Pre-FOMC announcement drift. J. Finance 70 (1), 329–371.
Mykland, P.A., Zhang, L., 2009. Inference for continuous semimartingales observed at high frequency. Econometrica 77 (5), 1403–1445.
Nakamura, E., Steinsson, J., 2018. High-frequency identification of monetary non-neutrality: The information effect. Q. J. Econ. 133 (3), 1283–1330.
Nison, S., 2001. Japanese Candlestick Charting Teechniques, second ed. Prentice Hall Press.

Olver, F.W., Lozier, D.W., Boisvert, R.F., Clark, C.W., 2010. NIST Handbook of Mathematical Functions. Cambridge University Press.
Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53, 61–65.
Pötscher, B.M., Preinerstorfer, D., 2022. A Modern Gauss-Markov Theorem? Really?. Working Paper, University of Vienna, Department of Statistics.
Renault, E., Sarisoy, C., Werker, B.J., 2017. Efficient estimation of integrated volatility and related processes. Econom. Theory 33 (2), 439–478.
Rigobon, R., 2003. Identification through heteroskedasticity. Rev. Econ. Stat. 85 (4), 777–792.
Rigobon, R., Sack, B., 2004. The impact of monetary policy on asset prices. J. Monetary Econ. 51 (8), 1553–1575.
Savor, P., Wilson, M., 2014. Asset pricing: A tale of two days. J. Financ. Econ. 113 (2), 171–201.
Schennach, S.M., 2020. Mismeasured and unobserved variables. In: Durlauf, S.N., Hansen, L.P., Heckman, J.J., Matzkin, R.L. (Eds.), Handbook of Econometrics, Vol. 7A. Elsevier, pp. 487–565.
Shaffer, J.P., 1991. The Gauss-Markov theorem and random regressors. Amer. Statist. 45 (4), 269–273.
Van der Vaart, A.W., 1998. Asymptotic Statistics. Cambridge University Press.
Zhang, L., Mykland, P.A., Aït-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. J. Amer. Statist. Assoc. 100 (472), 1394–1411.