

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Economics

School of Economics

3-2022

Occupation density estimation for noisy high-frequency data

Congshan ZHANG
Duke University

Jia LI
Singapore Management University, jjali@smu.edu.sg

Tim BOLLERSLEV
CREATES, Denmark

Follow this and additional works at: https://ink.library.smu.edu.sg/soe_research



Part of the [Econometrics Commons](#)

Citation

ZHANG, Congshan; LI, Jia; and BOLLERSLEV, Tim. Occupation density estimation for noisy high-frequency data. (2022). *Journal of Econometrics*. 227, (1), 189-211.

Available at: https://ink.library.smu.edu.sg/soe_research/2578

This Journal Article is brought to you for free and open access by the School of Economics at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Economics by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Occupation Density Estimation for Noisy High-Frequency Data

Congshan Zhang* Jia Li† Tim Bollerslev‡

September 12, 2019

Abstract

This paper studies the nonparametric estimation of occupation densities for semimartingale processes observed with noise. As leading examples we consider the stochastic volatility of a latent efficient price process, the volatility of the latent noise that separates the efficient price from the actually observed price, and nonlinear transformations of these processes. Our estimation methods are decidedly nonparametric and consist of two steps: the estimation of the spot price and noise volatility processes based on pre-averaging techniques and in-fill asymptotic arguments, followed by a kernel-type estimation of the occupation densities. Our spot volatility estimates attain the optimal rate of convergence, and are robust to leverage effects, price and volatility jumps, general forms of serial dependence in the noise, and random irregular sampling. The convergence rates of our occupation density estimates are directly related to that of the estimated spot volatilities and the smoothness of the true occupation densities. An empirical application involving high-frequency equity data illustrates the usefulness of the new methods in illuminating time-varying risks, market liquidity, and informational asymmetries across time and assets.

Keywords: high-frequency data; volatility; occupation density; microstructure noise; informed trading.

JEL: C14, C22, C58, G14.

*Department of Economics, Duke University; email: congshan.zhang@duke.edu.

†Department of Economics, Duke University; email: jl410@duke.edu.

‡Department of Economics, Duke University; NBER; CREATES; email: boller@duke.edu.

1 Introduction

Volatility is central to financial theory and decision making. Volatility is also inherently latent. We provide new econometric methods based on the notion of occupation densities for characterizing dynamic distributional features of the volatility of the latent efficient price of a financial asset, the volatility of the “noise” that arises from market microstructure frictions, and nonlinear transformations of these processes. Our theoretical results rely on infill asymptotic arguments in a possibly non-stationary non-ergodic setting, and are robust to leverage effects, price and volatility jumps, serial dependence in the noise, and general forms of random irregular sampling.

Methods for describing the distributions of financial market volatility have been the subject of a very large and still growing literature (see, e.g., the survey in Andersen, Bollerslev, and Diebold (2010)). Starting with the earliest work based on parametric ARCH/GARCH and stochastic volatility type models, to the more recent work relying on so-called realized volatilities constructed from high-frequency intraday data, our understanding of the latent volatility dynamics and distributional features of volatility has improved substantially. Still, most of the estimation methods hitherto used in the literature rely on their own set of assumptions, be it specific distributional assumptions used in the formulation of most ARCH/GARCH and stochastic volatility models, or the assumption of an efficient frictionless market that forms the foundation for the original realized volatility concept. We seek to remedy these limitations by providing an essentially unrestricted model-free approach that permits very general forms of market microstructure noise and irregularly sampled high-frequency observations.

Our estimation for the occupation densities is based on nonparametric estimators for the spot volatility of the efficient price and that of the microstructure noise. In the econometrics literature, the estimation of spot volatility traces back to Foster and Nelson (1996) and Comte and Renault (1998); see also Kristensen (2010). However, none of these studies allowed for price jumps or microstructure noise. Several jump-robust spot estimators have been proposed more recently based on the truncation technique of Mancini (2001); see, for example, Jacod and Protter (2012). There has also been a growing interest in the design of spot volatility estimators that are robust to microstructure noise, as exemplified by Zu and Boswijk (2014), Mancini, Mattiussi, and Renò (2015) and Bibinger, Hautsch, Malec, and Reiss (2019) among others. Nonetheless, the statistical settings in all of these papers are somewhat restrictive.¹

By comparison, we develop nonparametric spot estimators for the volatility of the price and

¹The theory of Mancini, Mattiussi, and Renò (2015), which include Zu and Boswijk (2014) as a special case, is restricted to a setting with independent noise and deterministic sampling. Bibinger, Hautsch, Malec, and Reiss (2019) do allow for finitely dependent noise, but their sampling scheme is essentially deterministic as well.

noise processes in the very general setting of Jacod, Li, and Zheng (2019, 2017), in which: (i) the latent price may have jumps; (ii) the noise may exhibit long-run serial dependence and possibly depend on both the price and the volatility of the price; and (iii) the sampling scheme may be irregular and stochastic, and possibly also depend on other processes, like the price and the volatility. In contrast to the estimators for the integrated volatility of the price and noise developed by Jacod, Li, and Zheng (2019, 2017), however, we propose new block-wise estimators for the spot price volatility and the spot volatility of the noise. While these estimators are explicitly designed for our main purpose of occupation density estimation, we also establish that the new spot volatility estimator for the price attains the optimal rate in this very general setting. In addition, we derive a type of uniform convergence for the spot estimators that we rely on in studying the asymptotic properties of the associated occupation density estimators.

Armed with the first-step nonparametric spot volatility estimates for the efficient price and the microstructure noise, we nonparametrically estimate the occupation densities of both processes, along with non-linear transformations thereof. Intuitively, the occupation density of a process measures the time that it spends in the vicinity of specific levels. As such, it may be seen as an “ex post” version of the usual probability density, in the same sense that the realized variance may be seen as an “ex post” analogue to the population variance.

Li, Todorov, and Tauchen (2013) have previously studied the nonparametric estimation of volatility occupation densities in a basic setting without microstructure noise or irregular sampling. The theory developed in the present paper is substantially more general than this prior work. First, we allow for general, and empirically more realistic, forms of microstructure noise and random irregular sampling. This in turn results in a series of highly nontrivial technical complications. Second, by explicitly accounting for the presence of noise, our setup allows us to also estimate the occupation densities for the volatility of the noise, as well as the ratio between the volatility of the noise and that of the efficient price. Each of these measures carries its own distinct economic interpretation and, hence, significantly broadens the empirical scope of the new procedures developed here. In particular, as discussed further below, the volatility of the noise is naturally interpreted as a measure of the effective spread or illiquidity (see, e.g., Roll (1984)), while the “noise-to-signal” ratio process serves as a succinct measure of the level of informed trading (see, e.g., Easley, Kiefer, O’Hara, and Paperman (1996)).

We illustrate the empirical usefulness of the new methods with ultra high-frequency (i.e., tick-by-tick) price data in a sample from 2007 to 2014, which includes both the height of the 2008 financial crisis and much more stable non-crisis periods before and after. We consider two individual stocks, Goldman Sachs (GS) and Starbucks (SBUX), as representatives of a financial and a non-

financial firm, respectively. We further complement the estimates for these individual stocks, with the analysis of prices at the aggregated level for the XLF financial sector exchange traded fund (ETF) and the SPY ETF for the S&P 500 market index. Comparing the occupation density estimates obtained for different episodes in our sample, we find that both the center and the dispersion of the distributions of the price volatility were higher during the crisis period than non-crisis periods. This, of course, is hardly surprising. Meanwhile, we also document the same pattern for the volatility of the microstructure noise. Taken together, these findings therefore corroborate the idea that not only did financial markets become more volatile during the crisis, they generally also became less liquid.

Interestingly, and more surprisingly, looking at the distributions of the *ratio* between the estimated noise and price volatility processes, we find that this measure of informational asymmetry concentrates at notably lower levels for financial firms than for non-financial firms during the crisis. As such, this suggests that the increased transparency and concerted governmental efforts aimed at stabilizing financial markets during the crisis may indeed have been successful in reducing informational asymmetries for the financial firms at the very center of the crisis. At a broader level this also directly highlights the usefulness of the new estimation theory for casting new light on questions of economic import.

The remainder of the paper is organized as follows. Section 2 introduces the theoretical setting. Section 3 develops the new nonparametric methods for spot volatility and occupation density estimation. Section 4 discusses the results from applying the new methods to high-frequency stock price data, and further connects the empirical findings to economic market microstructure theories. Section 5 concludes. Section 6 contains all proofs.

2 Theoretical setup and assumptions

2.1 Efficient and observed price processes

We assume that the efficient log-price X_t of an asset may be described as an Itô semimartingale defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$,

$$X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \sum_{s \leq t} \Delta X_s, \quad (2.1)$$

where the drift process b is locally bounded, the stochastic volatility process σ is càdlàg, W is a one-dimensional standard Brownian motion, and $\Delta X_t \equiv X_t - X_{t-}$ denotes the time- t (if any) jump in X . This setting is quite standard in the analysis of high-frequency financial data; see, for example, Jacod and Protter (2012) and Aït-Sahalia and Jacod (2014).

The continuous-time efficient price process X is not directly observable. Instead, we observe “noisy” prices at random (stopping) times over some fixed time interval $[0, T]$. Without loss of generality, we normalize $T = 1$, and denote the sampling times by $0 = T(n, 0) < T(n, 1) < \dots$. The actually observed price Y_i^n at each of the sampling times $T(n, i)$ is assumed to have a standard “signal-plus-noise” structure,

$$Y_i^n = X_{T(n,i)} + \epsilon_i^n, \quad (2.2)$$

where the observation error ϵ_i^n stems from various market microstructure frictions (e.g., discreteness, and bid-ask bounce effects). This term is commonly referred to as “noise” in the high-frequency econometrics literature.

The prior literature on realized volatility mainly focuses on the estimation of the volatility σ of the efficient price, treating the noise term as a statistical nuisance. However, the “noise” term is of central importance from an empirical market microstructure perspective (see, e.g., Hasbrouck (2007)). In particular, the volatility of the microstructure noise component ϵ_i^n may be interpreted as a measure of market illiquidity, and as such is of direct economic interest. This interpretation can be traced back at least to Roll (1984), who suggested that the volatility of ϵ_i^n may be seen as a proxy for the “effective spread,” or the cost associated with trading the asset. In actuality, the effective spread is typically smaller than the bid-ask spread observed in quote data, in that trades often occur within the quoted spreads due to hidden orders, dark pools, or the (slight) mismatch between the time stamps of transactions and quotes that invariably exists in most high-frequency databases.²

In our effort to further study the volatility of the noise as an illiquidity measure, we follow Jacod, Li, and Zheng (2019, 2017) and assume that it has a multiplicative structure,

$$\epsilon_i^n = \gamma_{T(n,i)} \cdot \chi_i. \quad (2.3)$$

Without loss of generality, we normalize the standard deviation of the χ_i shock to unity, so the γ process is directly interpretable as the volatility of the noise. We in turn use our estimate of γ as a measure of the effective spread in our empirical analysis. Importantly, since the γ process is generally stochastic and may be strongly persistent, our setup allows the ϵ_i^n noise term to be strongly serially dependent. Moreover, the χ_i term is also allowed to be serially correlated, thus permitting an extra layer of dependence. Formal regularity conditions on the different noise terms are given in Assumption 2 below.

²Although the effective spread measure extracted from the transaction prices are robust to the issue of within-spread trades, it needs to be estimated statistically, and hence is subject to estimation error. The actual quote prices, on the other hand, are directly observed. As such, these measures have their own distinct merits and should be regarded as complements rather than substitutes.

In addition to the volatility of the efficient price and the volatility of the noise, the corresponding “noise-to-signal” process,

$$\varphi_t = \gamma_t/\sigma_t, \tag{2.4}$$

is imbued with its own distinct economic interpretation. Indeed, using a version of the Glosten–Milgrom model (see Glosten and Milgrom (1985)), Hasbrouck (2007) shows that the bid-ask spread is proportional to the price volatility, and that the ratio between the two is an increasing function of the informational asymmetry between the market maker and the informed trader within the model.³ In a further extension of the Glosten–Milgrom model explicitly allowing for uncertain information events, Easley, Kiefer, O’Hara, and Paperman (1996) demonstrate that the spread-volatility ratio is also directly proportional to the probability of informed trading. As such, the φ_t ratio process is naturally interpreted as an omnibus measure of informational asymmetry, providing complementary information to that afforded by the more standard σ_t and γ_t price volatility and trading illiquidity measures.

2.2 Occupation densities

Our main econometric interest centers on characterizing *distributional* features of the price volatility σ , the noise volatility γ , and their ratio φ . As discussed above, these three different processes may be seen as proxies for the price risk, the market illiquidity, and the informational asymmetry of the asset, respectively.

Kernel density techniques are commonly used for empirically estimating distributional features of stochastic processes. However, the classical theory underlying kernel density estimation invariably relies on stationarity and weak-dependence type assumptions, and the exploitation of these regularities for characterizing the invariant distribution of the process of interest in a long-span asymptotic setting. By contrast, we are interested in estimating the distributional features of the σ , γ , and φ processes over relatively *short* calendar time-spans, so as to allow for temporal variation and comparisons of the estimates obtained in different economic environments. Since financial market volatility is well known to be highly persistent (see, e.g., Andersen, Bollerslev, Diebold, and Ebens (2001); Andersen, Bollerslev, Diebold, and Labys (2001)), arguing that standard stationarity and weak-dependence type assumptions have “kicked in” over short time spans simply is not tenable.

Instead, we depart from conventional kernel-based density estimation theory, and rely on the notion of occupation densities. Formally, for a generic stochastic process Z_t , its occupation time

³The level of information asymmetry is formally modeled as the proportion of informed traders in the population; see Section 5.2 in Hasbrouck (2007) for details.

is defined as,

$$F_Z(x) \equiv \int_0^1 1_{\{Z_s \leq x\}} ds, \quad x \in \mathbb{R}.$$

This is a random function corresponding to the proportion of time the process is below the level x (recall that the sample span T is normalized to unity). If the occupation time is pathwise differentiable with respect to the spatial variable x , the associated occupation density $f_Z(\cdot)$ is simply defined as its derivative,

$$f_Z(x) \equiv dF_Z(x)/dx \quad x \in \mathbb{R}.$$

The existence and basic properties of occupation densities for Markov and Gaussian processes have been studied by Geman and Horowitz (1980) and Marcus and Rosen (2006), among others. Additional results specifically pertaining to the jump-diffusion type processes widely used in analyzing financial data, as formally defined in (2.1), are available in Li, Todorov, and Tauchen (2016).

Intuitively, the occupation density measures how much time a process (e.g., σ , γ or φ) spends in the vicinity of specific levels. Li, Todorov, and Tauchen (2013) have previously proposed an estimator for the occupation density of the price volatility in a basic setting without microstructure noise (i.e., $\epsilon_i^n \equiv 0$). In the current paper, we consider a substantially wider class of processes, allowing for both serially dependent microstructure noise and irregular random sampling. Importantly, we also go beyond the estimation of the (price) volatility occupation density, by developing new methods for estimating the occupation densities for the noise volatility and the ratio processes, both of which are of independent economic interest.

2.3 Regularity conditions

Our new estimation procedures naturally require some, albeit very mild, regularity conditions. We begin by discussing the conditions pertaining to the irregular sampling times, followed by the conditions for the market microstructure noise, and the latent price process itself.

Compared to a regular sampling scheme with $T(n, i) = i\Delta_n$ for some fixed sampling interval Δ_n , the present setting allows $T(n, i)$ to be random, irregularly spaced, and dependent on various underlying processes, including the latent price volatility process. Specifically, let $N_t^n \equiv \sum_{i \geq 1} 1_{\{T(n, i) \leq t\}}$ denote the number of returns observed up to time t . The sampling interval for the i th return is then,

$$\Delta(n, i) \equiv T(n, i) - T(n, i - 1).$$

We consider an infill asymptotic setting, in which these sampling intervals go to zero as $n \rightarrow \infty$ in an average sense. Below, we use Δ_n to denote a positive real sequence such that $\Delta_n = o(1)$ as $n \rightarrow \infty$. Formally, we maintain the following assumption.

Assumption 1. *The following conditions hold for some constant $\rho > 1/2$, a sequence $(\tau_m)_{m \geq 1}$ of stopping times increasing to ∞ , and a sequence $(K_m)_{m \geq 1}$ of strictly positive constants.*

(i) α is continuous and $\alpha_t > 0$ for all t .

(ii) $\Delta_n N_t^n \xrightarrow{\mathbb{P}} A_t \equiv \int_0^t \alpha_s ds$, for all t .

(iii) In restriction to $\{T(n, i-1) \leq \tau_m\}$, $|\mathbb{E}[\alpha_{T(n, i-1)} \Delta(n, i) | \mathcal{F}_{T(n, i-1)}] - \Delta_n| \leq K_m \Delta_n^{1+\rho}$ and $\mathbb{E}[|\alpha_{T(n, i-1)} \Delta(n, i)|^\kappa | \mathcal{F}_{T(n, i-1)}] \leq K_m \Delta_n^\kappa$ for all $\kappa \geq 2$.

The α process in Assumption 1 provides a summary measure of the sampling intensity. Condition (iii), in particular, suggests that $\Delta_{n, i} \approx \Delta_n / \alpha_{T(n, i-1)}$, so the sampling interval is shorter when the intensity is higher, and vice versa. As such, the α process is only identified up to scale (scaling Δ_n and α with the same constant leads to no change in the theory). Assumption 1 is inspired by Jacod, Li, and Zheng (2019, 2017), albeit slightly weaker than the conditions invoked in those papers.⁴ It accommodates all of the irregular sampling schemes commonly used in the literature including, for example, modulated Poisson schemes and time-changed regular sampling schemes (with $\rho = 1$), as well as modulated random walk schemes (with any $\rho > 0$); see Jacod, Li, and Zheng (2019, 2017) for more detailed discussions of these alternative schemes.

The two separate processes that combine to define the market microstructure noise term also need to satisfy some mild regularity conditions. The next assumption formalizes these conditions.

Assumption 2. *The ϵ_i^n noise has the form (2.3) and the following conditions hold for a sequence $(\tau_m)_{m \geq 1}$ of stopping times increasing to ∞ , and a sequence $(K_m)_{m \geq 1}$ of strictly positive constants.*

(i) The process γ is càdlàg and adapted, and $1/K_m \leq \gamma_{t-} \leq K_m$ in restriction to $\{t \leq \tau_m\}$.

(ii) The variables $(\chi_i)_{i \in \mathbb{Z}}$ form a stationary sequence, independent of $\mathcal{F}_\infty \equiv \bigvee_{t > 0} \mathcal{F}_t$, with zero mean, unit variance, and finite moments of all orders.

(iii) The variables $(\chi_i)_{i \in \mathbb{Z}}$ is ρ -mixing with mixing coefficient ρ_k satisfying $\rho_k = O(k^{-v})$ for some $v > 4$.

Importantly, these conditions allow for quite general dependencies in the noise term, including serial correlation.⁵ The serial dependence may arise through two separate channels: calendar-time persistence in the noise volatility process (γ_t) , and tick-time dependence among the shocks (χ_i) .

⁴See Assumption (O) in Jacod, Li, and Zheng (2017) and Assumption (O- ρ, ρ') in Jacod, Li, and Zheng (2019). These authors require stronger assumptions than ours so as to prove central limit theorems. The derivation of central limit theorems for volatility occupation densities, even in the case without microstructure noise (as in Li, Todorov, and Tauchen (2013)), remains a very open question, mainly due to the fact that the occupation density itself is a random function with very limited smoothness properties.

⁵The empirical analyses in Hansen and Lunde (2006) and Jacod, Li, and Zheng (2017) clearly point to the importance of allowing for high-frequency serially dependent noise.

Finally, we also require a “smoothness” type condition on the actual price processes. The following assumption formalizes this final requisite set of conditions.

Assumption 3. *The following conditions hold for a sequence of stopping times $(\tau_m)_{m \geq 1}$ increasing to ∞ and a sequence of positive constants $(K_m)_{m \geq 1}$.*

(i) *The efficient price process X satisfies (2.1), with its jump component given by*

$$\sum_{s \leq t} \Delta X_s = \int_0^t \int_{\mathbb{R}} \delta(s, z) \mu(ds, dz),$$

where μ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with a deterministic compensator $\nu(dt, dz) = dt \otimes \lambda(dz)$ for some σ -finite measure $\lambda(dz)$ on \mathbb{R} , and δ is a real-valued predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$. There exists a sequence $(\Gamma_m)_{m \geq 1}$ of deterministic nonnegative λ -integrable functions on \mathbb{R} such that $\|\tilde{\delta}(\omega, t, z)\|^r \wedge 1 \leq \Gamma_m(z)$ in restriction to $\{t \leq \tau_m\}$ for some $r \in (0, 1]$.

(ii) *Each $\psi \in \{\alpha, b, \sigma, \gamma\}$ is an Itô semimartingale of the form,*

$$\psi_t = \psi_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{\sigma}'_s dW'_s + \tilde{J}_t, \quad (2.5)$$

where the processes \tilde{b} , $\tilde{\sigma}$ and $\tilde{\sigma}'$ are locally bounded and adapted, W' is a Brownian motion orthogonal to W , \tilde{J}_t is a purely discontinuous process of the form,

$$\tilde{J}_t = \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s, z) \mathbf{1}_{\{\|\tilde{\delta}(s, z)\| \leq 1\}} (\mu - \nu)(ds, dz) + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s, z) \mathbf{1}_{\{\|\tilde{\delta}(s, z)\| > 1\}} \mu(ds, dz),$$

where $\tilde{\delta}$ is a real-valued predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$. There exists a sequence $(\Psi_m)_{m \geq 1}$ of deterministic nonnegative λ -integrable functions on \mathbb{R} , such that $\|\tilde{\delta}(\omega, t, z)\|^2 \wedge 1 \leq \Psi_m(z)$ in restriction to $\{t \leq \tau_m\}$.

(iii) *For each $m \geq 1$, $1/K_m \leq \sigma_{t-} \leq K_m$ when $t \leq \tau_m$.*

Assumption 3 is quite mild and is satisfied by most continuous-time models used in economics and finance. In particular, it allows both of the σ and γ volatility processes to exhibit intraday periodicity, very general forms of dynamic dependencies, and jumps of unrestricted activity.

3 Occupation density estimation

Our estimation of the occupation densities consists of two steps. In the first step, we nonparametrically estimate the price volatility σ_t and noise volatility γ_t processes, and then, in the second step, we use these estimates to form a kernel-type nonparametric estimator for the corresponding occupation densities. Below, we detail each of these two steps in turn.

3.1 Nonparametric spot estimators via pre-averaging

The spot estimators proposed below are essentially localized versions of the integrated variance estimators in Jacod, Li, and Zheng (2019, 2017). These estimators in turn are general versions of the pre-averaging estimators (see Jacod, Li, Mykland, Podolskij, and Vetter (2009); Jacod, Podolskij, and Vetter (2010)), extended to allow for price jumps, serially dependent noise, and random irregular sampling. The pre-averaging method is one of many strategies that are now available for estimating the integrated volatility of a latent efficient price process observed with noise (see, e.g., Zhang, Mykland, and Ait-Sahalia (2005), Zhang (2006), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), Xiu (2010) and Reiß (2011) for some of the other methods that have been proposed in the literature). We focus on the pre-averaging method mainly because of its generality and ease-of-implementation. However, other spot estimators could in principle be similarly adapted for the purpose of occupation density estimation, as further discussed in Section 3.2 below.

The key idea behind the pre-averaging method is to “smooth out” the noise, by appropriately averaging the noisy returns. To this end, we consider a weight function $w : \mathbb{R} \mapsto \mathbb{R}$ that is continuous, piecewise continuously differentiable with Lipschitz-continuous derivative, and further satisfies $w(s) = 0$ for $s \notin (0, 1)$ and $\int_0^1 w(s)^2 ds > 0$. In addition to the weight function w , the pre-averaging method also requires a smoothing parameter h_n that governs the scope of the local averaging. Denoting the i th noisy return by $\Delta_i^n Y \equiv Y_{T(n,i)} - Y_{T(n,i-1)}$, the pre-averaged returns are then simply defined by,

$$\tilde{Y}_i^n \equiv \sum_{j=1}^{h_n-1} w(j/h_n) \Delta_{i+j}^n Y, \quad (3.1)$$

where the tuning parameter h_n needs to grow “slowly” to infinity compared to the increasing number of high-frequency observations.

By construction, the variance of the pre-averaged return \tilde{Y}_i^n depends not only on the price volatility, but also on the variance and autocovariances of the noise. This induces a bias in the volatility estimation that needs to be corrected. In order to do so, we require an estimate of the spot autocovariances of the noise for a growing number of lags k'_n . As is common in the literature on spot volatility estimation, we consider a local window L_n , and divide the (normalized) sample period $[0, 1]$ into $M_n = \lfloor 1/L_n \rfloor$ non-overlapping blocks (each of length L_n in calendar time). Intuitively, we need $L_n \rightarrow 0$ in order to “kill” the bias in the nonparametric estimation, while at the same time requiring each of the local windows to contain a sufficiently large number of pre-averaged returns so as to attain statistical precision.

Specifically, let J_i denote the index of the first observation in the i th estimation block (i.e.,

$[(i-1)L_n, iL_n]$). For this block, we define the (unnormalized) spot autocovariance of the noise at lag l as,

$$U(l)_i^n \equiv \sum_{j=0}^{N_{iL_n}^n - N_{(i-1)L_n}^n - 5k_n} (Y_{T(n, J_i+j)} - \bar{Y}_{J_i+j+2k_n}) (Y_{T(n, J_i+j+l)} - \bar{Y}_{J_i+j+4k_n}), \quad (3.2)$$

where $\bar{Y}_j^n \equiv k_n^{-1} \sum_{k=0}^{k_n-1} Y_{T(n, j+k)}$. For the purpose of defining the spot volatility estimators, it is useful to further associate the weight function w and the smoothing parameter h_n with the following constants,

$$\begin{aligned} \phi_n &\equiv \frac{1}{h_n} \sum_{j \in \mathbb{Z}} w \left(\frac{j}{h_n} \right)^2, \\ \bar{\phi}_{n,l} &\equiv h_n \sum_{j \in \mathbb{Z}} \left(w \left(\frac{j+1}{h_n} \right) - w \left(\frac{j}{h_n} \right) \right) \left(w \left(\frac{j+1-l}{h_n} \right) - w \left(\frac{j-l}{h_n} \right) \right). \end{aligned} \quad (3.3)$$

Using this notation, our spot variance estimator of the efficient price on the i th block may be conveniently expressed as,

$$\begin{aligned} \hat{\sigma}_{(i-1)L_n}^2 &\equiv \frac{1}{L_n} \frac{1}{h_n \phi_n} \sum_{j=0}^{N_{iL_n}^n - N_{(i-1)L_n}^n - h_n} \left(\tilde{Y}_{J_i+j}^n \right)^2 1_{\{|\tilde{Y}_{J_i+j}^n| \leq u_n\}} \\ &\quad - \frac{1}{L_n} \frac{1}{h_n^2 \phi_n} \sum_{l=-k'_n}^{k'_n} \bar{\phi}_{n,l} U(|l|)_i^n, \end{aligned} \quad (3.4)$$

where the truncation threshold u_n is introduced to eliminate jumps (following the approach of Mancini (2001)), and the second term involving the growing number of k'_n autocovariances corrects for the bias induced by the possibly autocorrelated noise. Correspondingly, our estimator for the spot variance of the noise is given by,

$$\hat{\gamma}_{(i-1)L_n}^2 \equiv \frac{U(0)_i^n}{N_{iL_n}^n - N_{(i-1)L_n}^n}. \quad (3.5)$$

Finally, to extend the estimators on each of the blocks defined in (3.4) to all times in $[0, 1]$, we simply set,

$$\hat{\sigma}_t^2 \equiv \begin{cases} \hat{\sigma}_{(i-1)L_n}^2, & t \in [(i-1)L_n, iL_n], \quad 1 \leq i \leq M_n, \\ \hat{\sigma}_{(M_n-1)L_n}^2, & t \in [M_n L_n, 1], \end{cases} \quad (3.6)$$

with $\hat{\gamma}_t^2$ extended from (3.5) in a similar fashion.

The estimators defined above all rely on certain tuning parameters, as it is invariably the case with this type of nonparametric estimation in a general statistical setting. Since all of these tuning parameters are introduced in a standard fashion, the prior literature is also highly informative

about their choice and interpretation. The smoothing parameter h_n , in particular, is standard in all pre-averaging type estimators (see, e.g., Jacod, Li, Mykland, Podolskij, and Vetter (2009) and Jacod, Podolskij, and Vetter (2010)). It serves to reduce the bias resulting from microstructure noise. The thresholding technique and the use of the u_n truncation parameter is also standard for eliminating jumps, both in the cases with or without noise (see, e.g., Mancini (2001), Aït-Sahalia, Jacod, and Li (2012) and Li (2013)). Allowing the noise to have mixing-type serial dependence necessitates the additional tuning parameters k_n and k'_n for the purpose of estimating the “local long-run” variance of the noise (see Jacod, Li, and Zheng (2019, 2017)). This mirrors the heteroskedasticity and autocorrelation robust variance estimation in conventional long-span time-series settings. Finally, the use of a local window L_n is also a common feature of essentially all spot volatility estimators, with or without any noise (see, e.g., Jacod and Protter (2012)).

The following theorem formally spells out the consistency of the spot estimators and their associated rates of convergence subject to standard choices of the different tuning parameters.⁶

Theorem 1. *Suppose that (i) Assumptions 1, 2 and 3 hold; (ii) $h_n \asymp \Delta_n^{-1/2}$, $k_n \asymp \Delta_n^{-1/5}$, $k'_n \asymp \Delta_n^{-1/8}$, $L_n \asymp \Delta_n^\tau$ with $0 < \tau < \frac{1}{3}$, $u_n \asymp (h_n \Delta_n)^\varpi$ and*

$$r < \frac{2 \lfloor v \rfloor - 4}{2 \lfloor v \rfloor - 3} \quad \text{and} \quad \frac{1}{4 - 2r} < \varpi < \frac{2 \lfloor v \rfloor - 3}{4 \lfloor v \rfloor - 4}.$$

Then, for each $t \in [0, 1]$,

$$\begin{aligned} \hat{\sigma}_t^2 - \sigma_t^2 &= O_p \left(\Delta_n^{\tau/2} \vee \Delta_n^{1/4 - \tau/2} \right), \\ \hat{\gamma}_t^2 - \gamma_t^2 &= O_p \left(\Delta_n^{\tau/2} \vee \Delta_n^{3/10 - \tau/2} \right). \end{aligned}$$

If, in addition (iii) $0 < \tau < 11/60$, then

$$\begin{aligned} \sup_{1 \leq i \leq M_n} \left| \hat{\sigma}_{(i-1)L_n}^2 - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds \right| &= O_p \left(\Delta_n^{1/4 - \tau} \vee \Delta_n^{11/40 - 3\tau/2} \right), \\ \sup_{1 \leq i \leq M_n} \left| \hat{\gamma}_{(i-1)L_n}^2 - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \gamma_s^2 ds \right| &= O_p \left(\Delta_n^{2/5 - 3\tau/2} \vee \Delta_n^{\tau/2} \log(\Delta_n^{-1})^{1/2} \right). \end{aligned}$$

Theorem 1 provides upper bounds for both pointwise and uniform convergence rates of the spot variance estimator of the efficient price and that of the microstructure noise. The former, in particular, attains the $\Delta_n^{-1/8}$ pointwise convergence rate for $\tau = 1/4$.⁷ Bibinger, Hautsch, Malec,

⁶ These new spot estimators could in principle be used in the construction of reduced-form forecasting models for price volatility and market illiquidity. We leave further work along those lines for future research.

⁷This optimal rate is attained by judiciously balancing the estimation bias, which is of order $O_p(\Delta_n^{\tau/2})$, and the statistical error, which is of order $O_p(\Delta_n^{1/4 - \tau/2})$.

and Reiss (2019) have recently shown that in a setting with finitely-dependent noise and a random independent sampling scheme, the optimal convergence rate for spot volatility estimation is $\Delta_n^{-1/8}$. Since the present statistical setting is more general, our spot volatility estimator thus achieves rate optimality as well.

The theorem also establishes the convergence rate at which $\hat{\sigma}_{(i-1)L_n}^2$ and $\hat{\gamma}_{(i-1)L_n}^2$ provide uniform approximations to the local averages of σ^2 and γ^2 within the $[(i-1)L_n, iL_n)$ window. If the σ_t and γ_t processes are both continuous, this result further implies that $\hat{\sigma}_t$ and $\hat{\gamma}_t$ are uniformly consistent estimators of σ_t and γ_t for $t \in [0, 1]$. However, this uniform convergence is not generally attainable when σ and/or γ contain jumps.

3.2 Occupation density estimators

Equipped with the nonparametric spot estimators, the occupation densities may be estimated by “plug-in” kernel smoothing type methods. More specifically, we are interested in the occupation density of $Z = g(\sigma^2, \gamma^2)$ for some smooth transform $g(\cdot)$. For example, by setting $g(\sigma^2, \gamma^2) = \sigma$, γ or γ/σ , we can study the occupation density of the price volatility, the noise volatility or their ratio, respectively.

More specifically, let the corresponding spot estimator of Z be denoted by $\hat{Z} = g(\hat{\sigma}^2, \hat{\gamma}^2)$. Then, in parallel to conventional probability density estimation, we estimate the occupation density $f_Z(\cdot)$ using,

$$\hat{f}_Z(x) \equiv \int_0^1 \frac{1}{\delta_n} H\left(\frac{\hat{Z}_s - x}{\delta_n}\right) ds, \quad (3.7)$$

where δ_n denotes a bandwidth sequence, and the kernel function $H(\cdot)$ is assumed to be bounded and Lipschitz continuous satisfying $\int_{-\infty}^{\infty} H(s) ds = 1$. To analyze the asymptotic property of this estimator, we impose the following additional assumption.

Assumption 4. *The occupation density $f_Z(\cdot)$ of Z exists. Moreover, for some constant $\beta \in (0, 1]$ and any compact set $\mathcal{K} \subset (0, \infty)$, there exists a constant $K > 0$, such that for all $a, b \in \mathcal{K}$, $\mathbb{E} |f_Z(a) - f_Z(b)| \leq K |a - b|^\beta$.*

Assumption 4 stipulates that the occupation density is β -Hölder continuous under the L_1 -norm for some $\beta \in (0, 1]$. Unlike conventional density estimation problems, in which the density function is smooth (e.g., differentiable up to a certain order), the occupation density of stochastic processes are generally not very smooth. For example, the occupation density for a one-dimensional Brownian motion satisfies Assumption 4 with $\beta = 1/2$ (see Exercise VI.1.33 in Revuz and Yor (1999)), while for more general jump-diffusion models the condition may be verified using Lemma 2.1 in Li, Todorov, and Tauchen (2016).

The following theorem establishes the consistency, along with the rate of convergence, for the occupation density estimator $\hat{f}_Z(\cdot)$ under this assumption. As alluded to above, our estimator for the occupation density is not specific to the pre-averaging spot estimator developed in Section 3.1. To make this point explicit, we state the theorem under a high-level condition on the convergence rate of the spot estimators $\hat{\sigma}^2$ and $\hat{\gamma}^2$ (see condition (ii)), while remaining agnostic about their exact constructions. Meanwhile, Theorem 1 may be used to verify this high-level condition for our proposed pre-averaging spot estimators in the general setting with serial dependent noise and random irregular sampling.

Theorem 2. *Suppose that (i) Assumption 4 holds; (ii) for $\theta = (\sigma^2, \gamma^2)$, $\sup_{1 \leq i \leq M_n} \|\hat{\theta}_{(i-1)L_n} - L_n^{-1} \int_{(i-1)L_n}^{iL_n} \theta_s ds\| = O_p(a_n)$ for some positive sequence $a_n = o(1)$; (iii) the kernel function H satisfies $\int H(z) |z|^\beta dz < \infty$; (iv) the processes σ^2 and γ^2 are, locally in time, bounded and bounded away from zero; (v) the transform $g(\cdot)$ is Lipschitz on each compact subset of $(0, \infty) \times (0, \infty)$. Then, for each fixed $x > 0$,*

$$\hat{f}_Z(x) - f_Z(x) = O_p(\delta_n^{-2} \bar{a}_n \vee \delta_n^\beta),$$

where $\bar{a}_n \equiv a_n \vee \Delta_n^{\tau/2}$.

Theorem 2 readily implies the consistency of $\hat{f}_Z(x)$, as long as the bandwidth δ_n converges to zero sufficiently “slowly,” namely $\delta_n \rightarrow 0$ and $\delta_n^{-2} \bar{a}_n \rightarrow 0$. Not surprisingly, the convergence rate of the occupation density estimator is closely connected to the uniform rate of convergence of the spot estimator and the smoothness of the occupation density. That being said, whether the estimator attains the optimal rate is unclear. In fact, characterizing the optimal rate for the nonparametric estimation of occupation densities remains an open question in need of future research more generally.

Importantly, condition (ii) only requires that the spot estimator $\hat{\theta}_{(i-1)L_n}$ uniformly approximates the local average $L_n^{-1} \int_{(i-1)L_n}^{iL_n} \theta_s ds$. As shown in Theorem 1, this condition can be verified in general settings, including jumps in the $\theta = (\sigma^2, \gamma^2)$ process. Instead, a theory build on the “seemingly natural” uniform approximation framed in terms of $\sup_{t \in [0,1]} \|\hat{\theta}_t - \theta_t\|$ would not allow for the presence of jumps.

Note also that condition (v) only requires $g(\cdot)$ to be Lipschitz on compact sets. As such, it allows $g(\cdot)$ to have “explosive growth” at zero, which is the case for the $g(\sigma^2, \gamma^2) = \sigma, \gamma$ and γ/σ transforms studied empirically below.⁸

The estimation of occupation densities for volatility processes have previously been studied by Li, Todorov, and Tauchen (2013). Our results extend this prior work in two important directions.

⁸The transforms $g(\sigma^2, \gamma^2) = \sigma$ and γ have unbounded first derivative at zero, while γ/σ is unbounded at $\sigma = 0$.

First, Li, Todorov, and Tauchen (2013) only consider a basic setting with regularly sampled prices observed without any noise. By contrast, we allow for very general forms of market microstructure noise and irregular random sampling. Second, Li, Todorov, and Tauchen (2013) only study the occupation density of the volatility of the (directly observed) efficient price itself. However, our methods allow for the estimation of the occupation densities of the γ and φ processes as well, thereby considerably broadening the scope of the analysis, and affording new economic insights related to liquidity and informational asymmetries in financial markets. The empirical application, to which we now turn, directly illustrates this point.

4 An empirical application

We apply the new methods to the estimation of the occupation densities for two individual stocks and two exchange trade funds (ETFs). Specifically, we consider Goldman Sachs (GS) and Starbucks (SBUX), as representatives of similar sized financial and non-financial firms, respectively, together with the financial sector ETF (XLF) and the S&P 500 ETF (SPY). We intentionally pick these specific assets in an effort to highlight the differences in volatility, liquidity and informational asymmetries between firms in the financial and non-financial sectors.

Our data is obtained from the Trade and Quote (TAQ) database from 2007 to 2014. The raw data consists of tick-by-tick transaction prices for all regular trading days. We remove “overnight” returns, keeping only the returns observed during the active part of the trading day from 9:30 to 16:00.

In contrast to more traditional probability density estimation techniques based on time series data, which rely on an increasing sample span and strong stationarity assumptions, our infill asymptotic theory remains formally valid and empirically reliable over relatively short fixed calendar time spans (e.g., 1 quarter or 1 year). In estimating the relevant spot processes, we further set $L_n = 1/13$, corresponding to 30-minute estimation blocks. We adaptively choose $h_n = \lfloor 0.5N^{1/2} \rfloor$, $k_n = \lfloor N^{1/5} \rfloor$, and $k'_n = \lfloor N^{1/8} \rfloor$, where N denotes the number of trades within an estimation block of length L_n ; these choices directly mirror Jacod, Li, and Zheng (2019). For simplicity, we do not impose any threshold truncation for jumps (the actual empirical estimates obtained with truncation are almost identical to the ones discussed below).

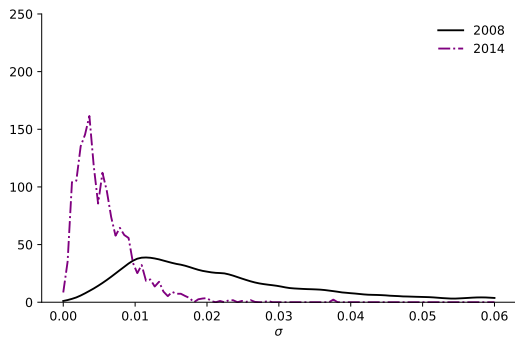
Our second stage estimation of the occupation densities relies on the widely used Epanechnikov kernel $H(u) = 0.75 \max\{1 - u^2, 0\}$, with the corresponding bandwidth parameter δ_n chosen by cross-validation. We do not claim any optimality of cross-validation, but merely consider it a sensible way of choosing the bandwidth parameter. The formal theoretical analysis of cross-

validation in the present context remains an open (and difficult) question for at least two reasons. First, our estimation is “doubly” nonparametric, with the second-stage kernel estimation relying on first-stage nonparametric high-frequency spot estimators. Secondly, the occupation densities themselves are stochastic processes with limited smoothness, rendering them much more difficult to analyze theoretically than the probability density functions studied in the more conventional nonparametric estimation literature.

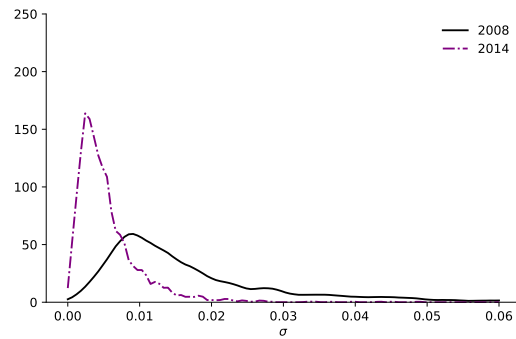
Turning to the actual empirical results, we begin by discussing the estimates from two years, 2008 and 2014, so as to make a direct comparison of the estimates obtained at the height of the financial crisis with the estimates during a more recent and stable post-crisis period. Figure 1 shows the occupation density estimates for the price volatility σ , the illiquidity measure γ , and the informational asymmetry measure φ , with the estimates for Goldman Sachs reported in the left three panels and the estimates for Starbucks in the right three panels. Looking first at the estimates for the volatility in panels (a) and (b) in the top row, we see that, not surprisingly, for both of the stocks the levels are clearly higher in 2008 than in 2014. At the same time, the volatility occupation density estimates reveal not only a shift in the level between the crisis and non-crisis years, but also in the general shapes of the distributions. The higher dispersion, in particular, that is evident for both of the stocks in 2008, implies a much greater volatility risk, or “volatility-of-volatility,” during the crisis. This pattern is especially pronounced for Goldman Sachs, for which the 2008 distribution manifests a heavy right tail. By comparison, the 2008 volatility density estimate for Starbucks is much closer to being symmetric, suggestive of an interesting contrast in the price volatility risk between financial and non-financial sector stocks during the crisis.

The occupation density estimates for the γ illiquidity process in panels (c) and (d) also clearly reveal both higher centered and more dispersed distributions for 2008 compared to 2014. This finding underscores the close relationship that generally exists between illiquidity and volatility. In contrast to the volatility occupation density estimates in panels (a) and (b), however, the estimated illiquidity occupation densities for the two different stocks appear quite similar during the crisis. This differential pattern therefore suggests that illiquidity is not entirely driven by volatility, and that the illiquidity-volatility relationship may be time-varying and differ across assets depending on the prevailing informational asymmetries at play.

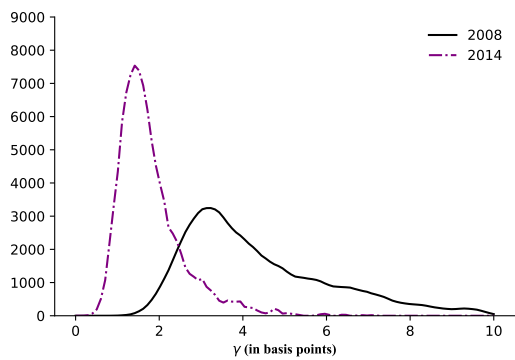
Economic market microstructure theory provides a useful guide for better understanding these issues. In particular, as previously noted, building on insights from the Glosten–Milgrom equilibrium model (see, e.g., Hasbrouck (2007); Easley, Kiefer, O’Hara, and Paperman (1996)), the ratio $\varphi = \gamma/\sigma$ may be seen as a monotonically increasing function of the level of informational asymmetry. This asymmetry naturally varies, both over time and across assets. During the financial



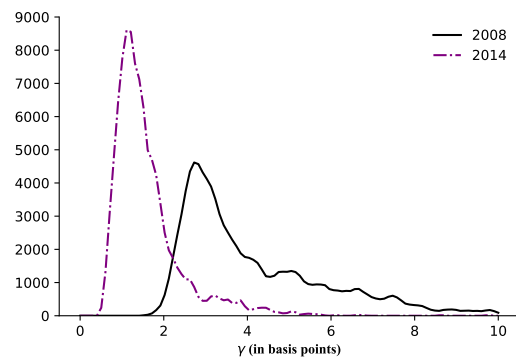
(a) Price Volatility, GS



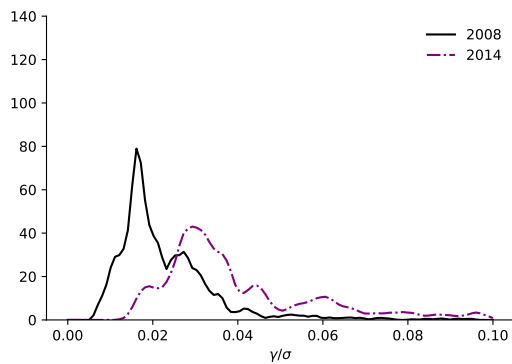
(b) Price Volatility, SBUX



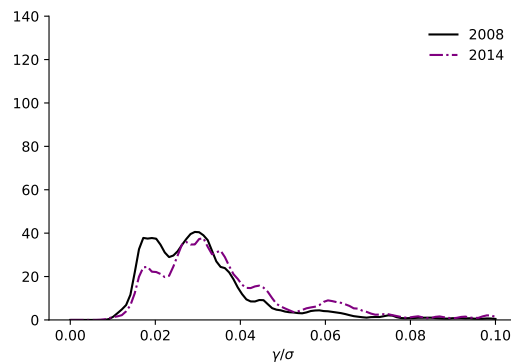
(c) Illiquidity, GS



(d) Illiquidity, SBUX



(e) Information Asymmetry, GS



(f) Information Asymmetry, SBUX

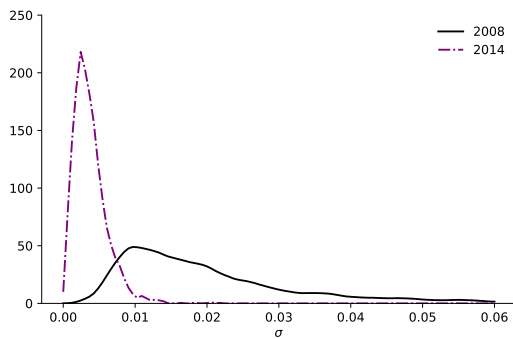
Figure 1: Occupation density estimates for Goldman Sachs (GS) and Starbucks (SBUX) individual stocks.

crisis, a tremendous amount of information gathering about financial sector firms and the financial sector as a whole was undertaken by the government. These governmental efforts were further accompanied by extensive analysis by regulators and researchers in academia and industry alike. Hence, in spite of the very high level of risk observed for a financial firm like Goldman Sachs, the level of informational asymmetry among market participants may actually have been *lower* during the crisis. Correspondingly, trading in Goldman Sachs stock may have been driven more by liquidity/rebalancing needs, as opposed to informed trading, during the crisis. On the other hand, the increased information gathering about the financial sector that occurred during the crisis should be much less relevant for a company like Starbucks that primarily derives its revenues from selling coffee.

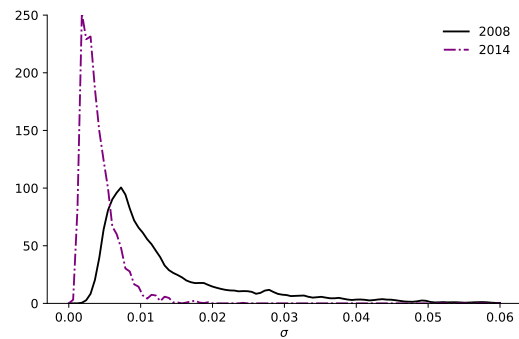
To examine this conjecture, the bottom two panels (e) and (f) in Figure 1 plot the occupation density estimates for the ratio process $\varphi = \gamma/\sigma$. Consistent with the implications from the economic reasoning above, the density estimate for Goldman Sachs concentrates around a much lower level in 2008 than in 2014, suggesting that the informational advantage of “informed traders” were indeed diluted by the increased awareness and level of public information related to the financial sector as a whole during the financial crisis. By comparison, the estimated occupation densities of the φ ratio process for Starbucks are quite similar for 2008 and 2014, indicative of much more stable through-time informational asymmetries.

The economic logic behind this “tale-of-two-firms” is not specific to Goldman Sachs and Starbucks, but holds true for other financial and non-financial firms, and accordingly diversified portfolios. To further corroborate this, Figure 2 shows the occupation density estimates for the XLF financial sector and SPY S&P 500 ETFs. The XLF is comprised of a wide range of stocks in the financial services industry, including insurance, banks, as well as consumer and mortgage finance, while the SPY is comprised of 500 large company stocks across all industries. As such, these two ETFs may naturally be seen as a “Wall Street” versus “Main Street” portfolio parallel to the Goldman Sachs versus Starbucks “tale-of-two-firms.”

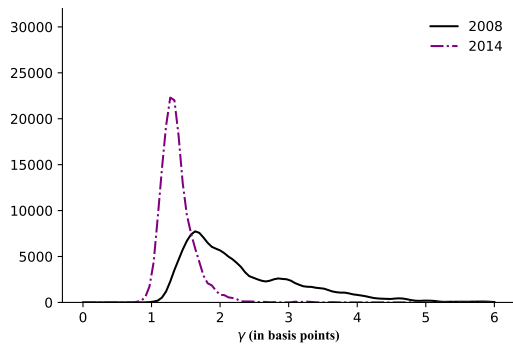
Looking at the actual estimation results also reveal qualitatively very similar patterns to those observed for the two individual stocks. In particular, the density estimates for the volatility and illiquidity processes, σ and γ , are both at higher overall levels and more dispersed in 2008 compared to 2014. At the same time, the estimates for the ratio process $\varphi = \gamma/\sigma$ for the SPY ETF shown in the bottom right panel (f) are fairly similar for each of the two years. By contrast, the estimates for the ratio process φ for the XLF financial sector ETF in panel (e) are clearly different between the crisis and non-crisis years. This difference again suggests that the large amount of public information gathering pertaining to the financial services industry that occurred during the crisis



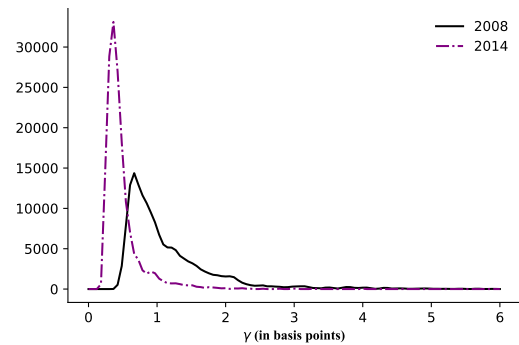
(a) Price Volatility, XLF



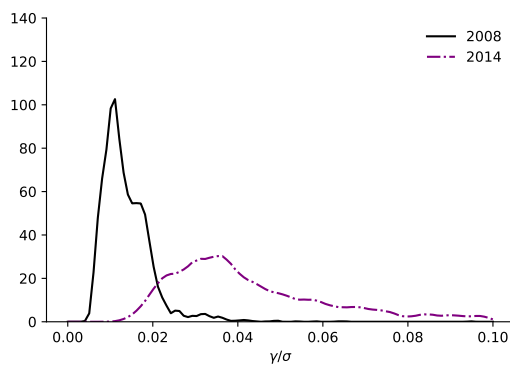
(b) Price Volatility, SPY



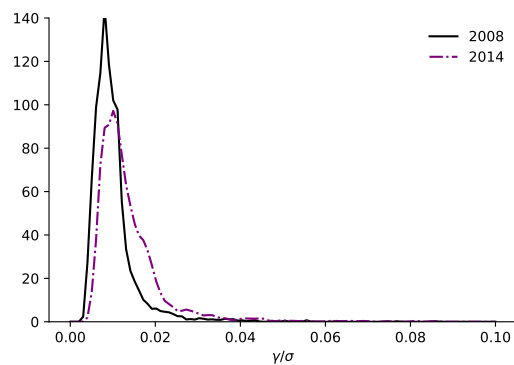
(c) Illiquidity, XLF



(d) Illiquidity, SPY



(e) Information Asymmetry, XLF



(f) Information Asymmetry, SPY

Figure 2: Occupation density estimates for the financial sector (XLF) and S&P 500 (SPY) exchange traded funds (ETFs).

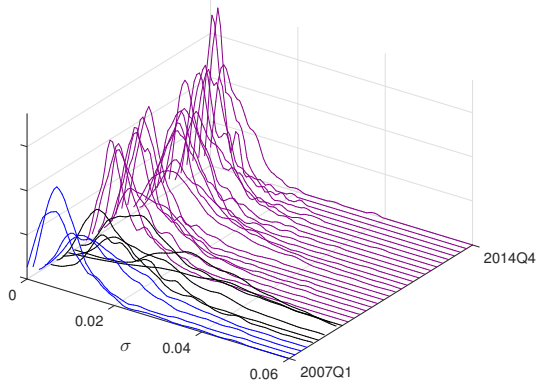
was effective in reducing informational asymmetries for the sector as a whole. It also indirectly suggests that without these concerted efforts, financial sector firms would likely have been even more costly to trade during the crisis, and as such these efforts might have helped mitigate liquidity spiral type effects (see, e.g., the discussion in Brunnermeier and Pedersen (2009)).

Finally, to help visualize the time-varying nature of risk in volatility, illiquidity, and asymmetric information more generally, we estimate the respective occupation densities for each quarter in our full 2007–2014 sample. The resulting estimates for the two individual stocks and the two ETFs shown in Figures 3 and 4, respectively, clearly demonstrate that the occupation densities do vary over time. They also make it clear that the more detailed results for 2008 and 2014 discussed above may indeed be seen as representative for crisis and non-crisis periods. Specifically, the occupation densities of price volatility and illiquidity both became noticeable more dispersed during the great recession, tightened afterwards, while the estimated densities pertaining to asymmetric information varied in the opposite direction.

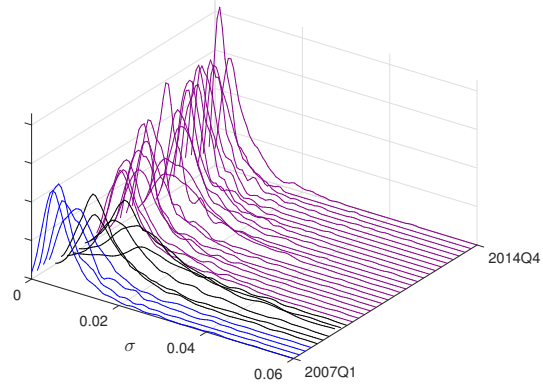
5 Conclusion

This paper develops new methods for the nonparametric estimation of occupation densities of the volatility of a latent efficient price process, the volatility of the noise that separates the efficient price from the actually observed price process, and nonlinear transformations of these processes. The new methods are valid in general high-frequency statistical settings. In parallel to conventional probability density estimators, the occupation density estimators naturally characterize the distributional features of the underlying processes. Unlike conventional theory for density estimation, however, the new theory underlying the estimation of the occupation densities does not require stationarity or weak-dependence assumptions on the processes of interest. This allows implementation of the new methods over relatively short calendar time-spans, thereby providing a useful complement to more traditional nonparametric density estimation techniques.

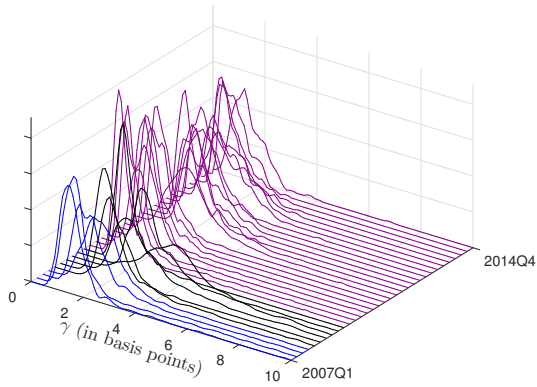
An empirical application based on ultra high-frequency equity returns illustrates the usefulness of the new methods for delineating information about time-varying risks, liquidity, and informational asymmetries. Consistent with prior empirical evidence based on alternative procedures, we document systematically higher levels of volatility and illiquidity during the financial crisis compared to a more recent non-crisis period, along with more dispersed distributions of these same quantities. In new empirical findings, we further document a clear distributional shift in the illiquidity/volatility ratio, indicative of lower informational asymmetries for financial sector firms during the financial crisis. By comparison, the estimated ratio processes and informational



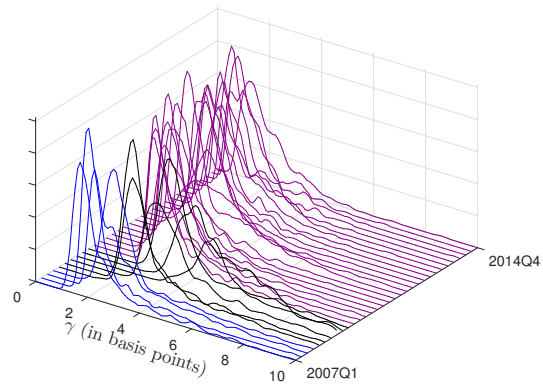
(a) Price Volatility, GS



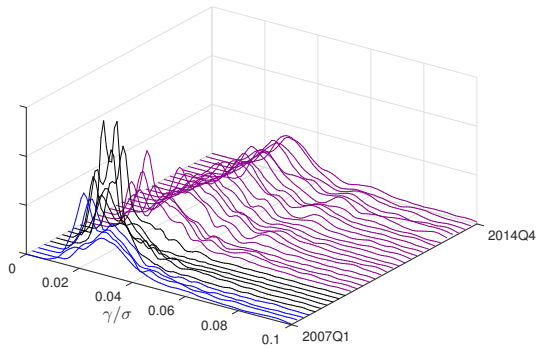
(b) Price Volatility, SBUX



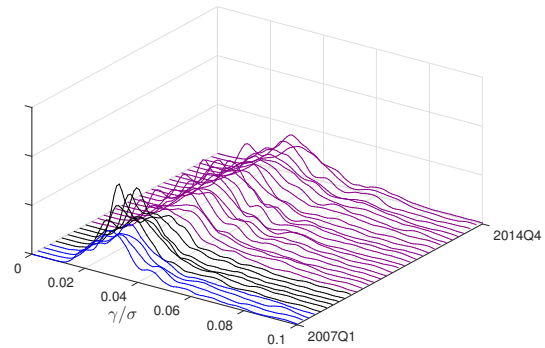
(c) Illiquidity, GS



(d) Illiquidity, SBUX

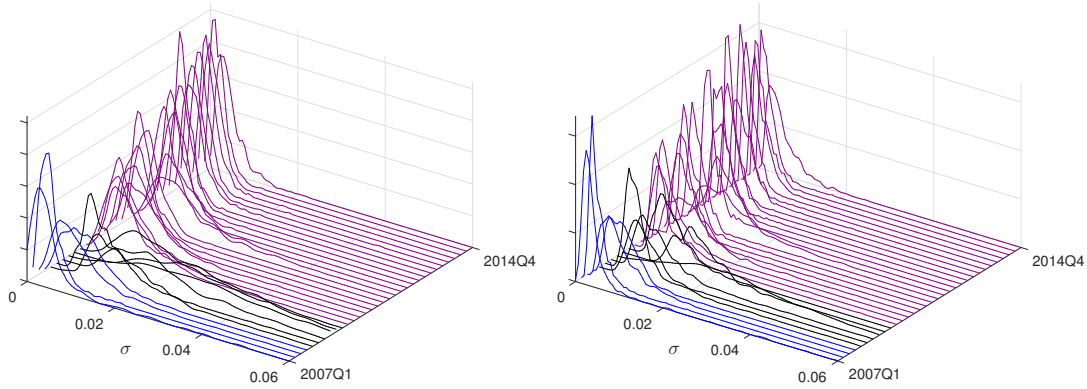


(e) Information Asymmetry, GS



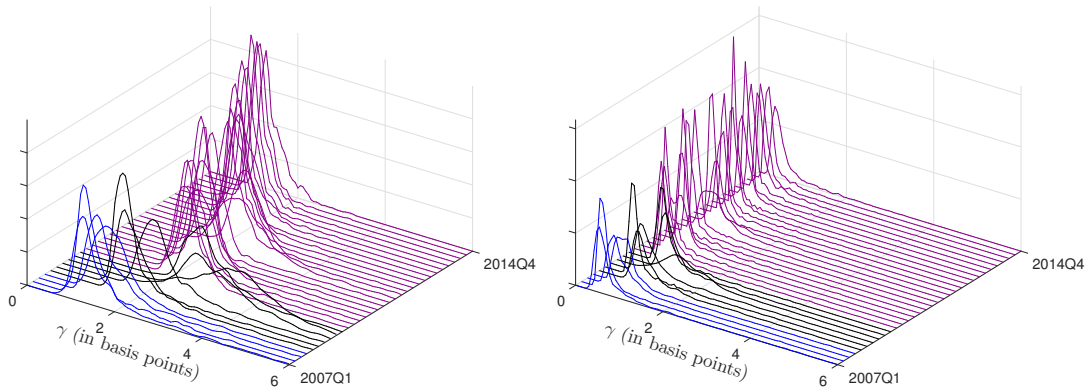
(f) Information Asymmetry, SBUX

Figure 3: Quarterly occupation density estimates for Goldman Sachs (GS) and Starbucks (SBUX) individual stocks spanning 2007–2014. The dark color highlights the six quarters from 2008Q1 to 2009Q2.



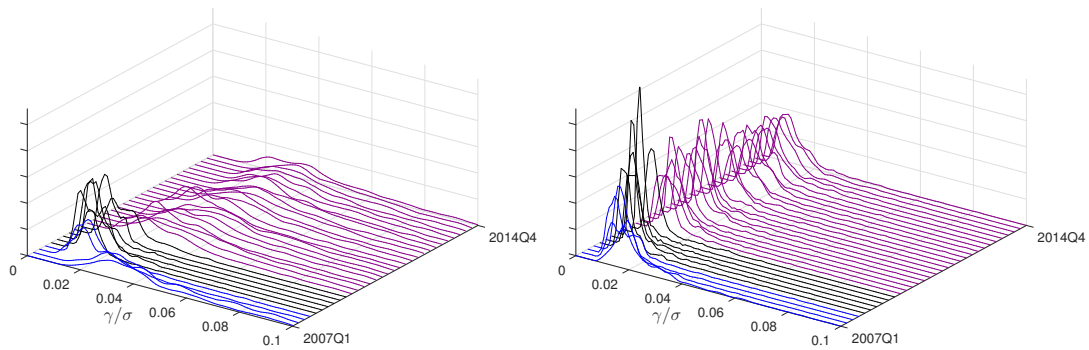
(a) Price Volatility, XLF

(b) Price Volatility, SPY



(c) Illiquidity, XLF

(d) Illiquidity, SPY



(e) Information Asymmetry, XLF

(f) Information Asymmetry, SPY

Figure 4: Quarterly occupation density estimates for the financial sector (XLF) and S&P 500 (SPY) exchange traded funds (ETFs) spanning 2007–2014. The dark color highlights the six quarters from 2008Q1 to 2009Q2.

asymmetries pertaining to non-financial firms appear remarkably stable across crisis and non-crisis periods.

6 Proofs

Throughout the proofs, we use K to denote a generic constant that may change from line to line. By a classical localization argument, we impose the following stronger assumption without loss of generality.

Assumption 5. *We have Assumptions 1–3 with $\tau_1 = \infty$. Moreover, the functions δ , $\tilde{\delta}$ and the processes $b, \sigma, 1/\sigma, \alpha, 1/\alpha, \gamma, 1/\gamma, X$ are bounded and for some constant $K \geq 1$, we have,*

$$N_t^n \leq \frac{Kt}{\Delta_n}, \quad (6.1)$$

$$\left| \mathbb{E} \left[\Delta(n, i) - \frac{\Delta_n}{\alpha_{T(n, i-1)}} \middle| \mathcal{F}_{T(n, i-1)} \right] \right| \leq K \Delta_n^{1+\rho}, \quad (6.2)$$

$$\mathbb{E} [\Delta(n, i)^\kappa] \leq K \Delta_n^\kappa. \quad (6.3)$$

We recall some classical estimates for Itô semimartingales that are used repeatedly in the sequel. If Z is an Itô semimartingale that satisfies Assumption 5, then for any two stopping times $u \leq v$ and a constant $q \geq 2$,

$$\begin{aligned} |\mathbb{E} [Z_v - Z_u | \mathcal{F}_u]| &\leq K \mathbb{E} [v - u | \mathcal{F}_u], \\ \mathbb{E} \left[\sup_{s \in [u, v]} |Z_s - Z_u|^q \middle| \mathcal{F}_u \right] &\leq K \mathbb{E} [(v - u)^q | \mathcal{F}_u] + K \mathbb{E} [v - u | \mathcal{F}_u]. \end{aligned}$$

Since $\mathbb{E} [\Delta(n, i)^\kappa] \leq K \Delta_n^\kappa$ (Assumption 5), we further have for any $i, j \geq 0$, $s \geq 0$ and $q \in [2, \kappa]$,

$$\begin{aligned} |\mathbb{E} [Z_{(T(n, i)+s) \wedge T(n, i+j)} - Z_{T(n, i)} | \mathcal{F}_{T(n, i)}]| &\leq Kj \Delta_n, \\ \mathbb{E} \left[\sup_{s \in [T(n, i), T(n, i+j)]} |Z_s - Z_{T(n, i)}|^q \middle| \mathcal{F}_{T(n, i)} \right] &\leq K ((j \Delta_n)^q + j \Delta_n). \end{aligned} \quad (6.4)$$

6.1 Proof of Theorem 1

We start with introducing some additional notation and preliminary estimates. For simplicity, we set $w_j^n \equiv w(j/h_n)$ and $\bar{w}_j^n \equiv w((j+1)/h_n) - w(j/h_n)$. For a generic process Z , we denote $Z_i^n = Z_{T(n, i)}$. Similarly, we write $\mathcal{F}_i^n = \mathcal{F}_{T(n, i)}$. We consider the following σ -fields,

$$\mathcal{G}_j \equiv \sigma(\chi_i : i \leq j), \quad \mathcal{G} \equiv \sigma(\chi_i : i \in \mathbb{Z}), \quad \mathcal{K}_j^n = \mathcal{F}_j^n \otimes \mathcal{G}_{j-h_n}.$$

In addition, we let X^c be the continuous part of X , $\Delta N(L_n)_i^n \equiv N_{iL_n}^n - N_{(i-1)L_n}^n$, $r(j) \equiv \mathbb{E}[\chi_0 \chi_j]$, $\tilde{r}_{n,0} \equiv \mathbb{E}[(\tilde{\chi}_0^n)^2]$ and

$$\left\{ \begin{array}{l} \bar{V}_t \equiv \int_0^t \sigma_s^2 ds, \quad \iota_j^n \equiv \sum_{u=1}^{h_n-1} (w_u^n)^2 \Delta_{j+u}^n \bar{V}, \\ \hat{\epsilon}_j^n \equiv (\tilde{\epsilon}_j^n)^2 - (\gamma_j^n)^2 \tilde{r}_{n,0}, \quad \hat{X}_j^{c,n} \equiv (\tilde{X}_j^{c,n})^2 - \iota_j^n, \\ Y_j^{c,n} \equiv X_{T(n,j)}^c + \epsilon_j^n, \quad \tilde{Y}_j^{c,n} \equiv - \sum_{u=0}^{h_n-1} \bar{w}_u^n Y_{j+u}^{c,n}, \\ \Xi_j^n \equiv (\tilde{Y}_j^{c,n})^2 - \iota_j^n - (\gamma_j^n)^2 \tilde{r}_{n,0} = \hat{X}_j^{c,n} + \hat{\epsilon}_j^n + 2\tilde{X}_j^{c,n} \tilde{\epsilon}_j^n, \end{array} \right.$$

where the pre-averaged values $\tilde{\chi}_0^n$, $\tilde{\epsilon}_j^n$, $\tilde{X}_j^{c,n}$ are defined in exactly the same way as in (3.1).

Lemma 1. *Let Z be a generic bounded càdlàg adapted process that satisfies (the localized version of) Assumption 3(ii). Suppose (i) the conditions of Theorem 1 hold; (ii) (m_n) is a sequence of positive integers satisfying $m_n \Delta_n^{1-\tau} \rightarrow 0$. Then we have*

$$(a) \quad \mathbb{E} \left[\left| \frac{\Delta_n}{L_n} \sum_{j=0}^{(\Delta N(L_n)_i^n - m_n) \vee 0} Z_{T(n, J_i+j)} - \alpha_{(i-1)L_n} Z_{(i-1)L_n} \right| \right] \leq K \left(\Delta_n^{\tau/2} \vee m_n \Delta_n^{1-\tau} \right);$$

$$(b) \quad \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{\Delta_n}{L_n} \sum_{j=0}^{(\Delta N(L_n)_i^n - m_n) \vee 0} Z_{T(n, J_i+j)} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \alpha_s Z_s ds \right| \right] \leq K \left(\Delta_n^{1/2-2\tau} \vee m_n \Delta_n^{1-\tau} \right);$$

(c) *Moreover, $A_{n,i} \equiv \{\Delta N(L_n)_i^n \leq KL_n/\Delta_n \text{ for some } K > 0\}$ satisfies $\mathbb{P}(A_{n,i}) \rightarrow 1$ and, when $\tau \in (0, 1/4)$, $A_n \equiv \{\Delta N(L_n)_i^n \leq KL_n/\Delta_n, \text{ for all } i \text{ and some } K > 0\}$ satisfies $\mathbb{P}(A_n) \rightarrow 1$.*

Proof. Parts (a) and (b). We decompose

$$\frac{\Delta_n}{L_n} \sum_{j=0}^{(\Delta N(L_n)_i^n - m_n) \vee 0} Z_{T(n, J_i+j)} - \alpha_{(i-1)L_n} Z_{(i-1)L_n} = \xi_i^n + \xi_i^m, \quad (6.5)$$

where

$$\xi_i^n \equiv \frac{\Delta_n}{L_n} \sum_{j=0}^{(\Delta N(L_n)_i^n - m_n) \vee 0} Z_{T(n, J_i+j)} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \alpha_s Z_s ds,$$

$$\xi_i^m \equiv \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \alpha_s Z_s ds - \alpha_{(i-1)L_n} Z_{(i-1)L_n}.$$

With an appeal to classic estimates for Itô semimartingales together with the assumption $L_n \asymp \Delta_n^\tau$, we deduce that

$$\mathbb{E} [|\xi_i^m|] \leq KL_n^{1/2} \leq K \Delta_n^{\tau/2}. \quad (6.6)$$

It remains to derive bounds for the term ξ_i^n on the right-hand side of (6.5). To this end, we decompose $\xi_i^n = \sum_{k=1}^4 \zeta_i^{n,k}$, where

$$\begin{aligned}\zeta_i^{n,1} &\equiv \frac{1}{L_n} \sum_{j=0}^{\Delta N(L_n)_i^n} Z_{T(n, J_i+j)} \left\{ \Delta_n - \mathbb{E} \left[\alpha_{T(n, J_i+j)} \Delta(n, J_i+j+1) \mid \mathcal{F}_{T(n, J_i+j)} \right] \right\}, \\ \zeta_i^{n,2} &\equiv \frac{1}{L_n} \sum_{j=0}^{\Delta N(L_n)_i^n} Z_{T(n, J_i+j)} \left\{ \mathbb{E} \left[\alpha_{T(n, J_i+j)} \Delta(n, J_i+j+1) \mid \mathcal{F}_{T(n, J_i+j)} \right] \right. \\ &\quad \left. - \alpha_{T(n, J_i+j)} \Delta(n, J_i+j+1) \right\}, \\ \zeta_i^{n,3} &\equiv \frac{1}{L_n} \sum_{j=0}^{\Delta N(L_n)_i^n} \left(Z_{T(n, J_i+j)} \alpha_{T(n, J_i+j)} \Delta(n, J_i+j+1) - \int_{T(n, J_i+j)}^{T(n, J_i+j+1)} Z_s \alpha_s ds \right), \\ \zeta_i^{n,4} &\equiv \frac{1}{L_n} \int_{iL_n}^{T(n, J_i + \Delta N(L_n)_i^n + 1)} Z_s \alpha_s ds \\ &\quad - \frac{1}{L_n} \int_{(i-1)L_n}^{T(n, J_i)} Z_s \alpha_s ds - \frac{1}{L_n} \sum_{j=(\Delta N(L_n)_i^n - m_n) \vee 0}^{\Delta N(L_n)_i^n} Z_{T(n, J_i+j)} \Delta_n.\end{aligned}$$

To prove (a), it is enough to bound $\mathbb{E}[|\zeta_i^{n,k}|]$, for $k = 1, \dots, 4$. To prove (b), we need bounds for $\mathbb{E}[\sup_{1 \leq i \leq M_n} |\zeta_i^{n,k}|]$, for $k = 1, \dots, 4$. We consider these cases in turn.

Case $k = 1$. Due to the boundedness of Z , the fact that $\Delta N(L_n)_i^n \leq N_1^n \leq K/\Delta_n$, (6.2) in Assumption 5 and $L_n \asymp \Delta_n^\tau$, we have

$$\begin{cases} |\zeta_i^{n,1}| \leq K \Delta_n^\rho / L_n \leq K \Delta_n^{\rho-\tau}, \\ \sup_{1 \leq i \leq M_n} |\zeta_i^{n,1}| \leq K \Delta_n^{\rho-\tau}. \end{cases} \quad (6.7)$$

Case $k = 2$. Let $\eta_j^{n,i}$ denote the summand in the definition of $\zeta_i^{n,2}$, so that we can rewrite $\zeta_i^{n,2} = L_n^{-1} \sum_{j=0}^{\Delta N(L_n)_i^n} \eta_j^{n,i}$. Note that the summands $\eta_j^{n,i}$ form a martingale difference array; hence, we have

$$\mathbb{E} \left[\left| \zeta_i^{n,2} \right|^2 \right] \leq \frac{1}{L_n^2} \mathbb{E} \left[\left(\sup_{0 \leq q \leq K/\Delta_n} \left| \sum_{j=0}^q \eta_j^{n,i} \right| \right)^2 \right] \leq K \Delta_n / L_n^2, \quad (6.8)$$

where the first inequality follows from the fact that $\Delta N(L_n)_i^n \leq K/\Delta_n$; the second inequality follows from Doob's inequality and (6.3) in Assumption 5. A maximal inequality further implies (with $\|\cdot\|_2$ denoting the L_2 norm)

$$\left\| \sup_{1 \leq i \leq M_n} |\zeta_i^{n,2}| \right\|_2 \leq M_n^{1/2} \sup_{1 \leq i \leq M_n} \left(\mathbb{E} \left[\left| \zeta_i^{n,2} \right|^2 \right] \right)^{1/2} \leq K \Delta_n^{1/2} M_n^{1/2} / L_n. \quad (6.9)$$

Since $L_n \asymp \Delta_n^\tau$ and $M_n \leq K/L_n$ by definition, we deduce from (6.8), (6.9) and Jensen's inequality that

$$\begin{cases} \mathbb{E} \left[\left| \zeta_i^{n,2} \right| \right] \leq K \Delta_n^{1/2-\tau}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,2} \right| \right] \leq K \Delta_n^{1/2-3\tau/2}. \end{cases} \quad (6.10)$$

Case $k = 3$. We observe that

$$\begin{aligned} \mathbb{E} \left[\left| \zeta_i^{n,3} \right| \right] &\leq \frac{K}{L_n} \mathbb{E} \left[\sum_{j=0}^{K/\Delta_n} \int_{T(n, J_i+j)}^{T(n, J_i+j+1)} \left| Z_{T(n, J_i+j)} \alpha_{T(n, J_i+j)} - Z_s \alpha_s \right| ds \right] \\ &\leq \frac{K}{L_n} \sum_{j=0}^{K/\Delta_n} \mathbb{E} \left[\sup_{s \in [T(n, J_i+j), T(n, J_i+j+1)]} \left| Z_{T(n, J_i+j)} \alpha_{T(n, J_i+j)} - Z_s \alpha_s \right| \Delta(n, J_i + j + 1) \right] \\ &\leq K \Delta_n^{1/2} / L_n, \end{aligned} \quad (6.11)$$

where the first inequality is by the triangle inequality; the second inequality is obvious; the third inequality follows from the Cauchy–Schwarz inequality, (6.3) in Assumption 5 and (6.4). By a maximal inequality, it follows that

$$\mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,3} \right| \right] \leq K M_n \sup_{1 \leq i \leq M_n} \mathbb{E} \left[\left| \zeta_i^{n,3} \right| \right] \leq K \Delta_n^{1/2} M_n / L_n. \quad (6.12)$$

Since $L_n \asymp \Delta_n^\tau$ and $M_n \leq K/L_n$, we deduce from (6.11) and (6.12) that

$$\begin{cases} \mathbb{E} \left[\left| \zeta_i^{n,3} \right| \right] \leq K \Delta_n^{1/2-\tau}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,3} \right| \right] \leq K \Delta_n^{1/2-2\tau}. \end{cases} \quad (6.13)$$

Case $k = 4$. The expectations of the absolute values of the first two terms in the definition of $\zeta_i^{n,4}$ are bounded by $K \Delta_n / L_n$. In addition, the absolute value of the third term there is bounded by $K \Delta_n m_n / L_n$. Under our maintained assumptions on L_n , we obtain

$$\begin{cases} \mathbb{E} \left[\left| \zeta_i^{n,4} \right| \right] \leq K m_n \Delta_n^{1-\tau}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,4} \right| \right] \leq K \left(\Delta_n^{1-2\tau} + m_n \Delta_n^{1-\tau} \right). \end{cases} \quad (6.14)$$

We are now ready to prove the assertions in parts (a) and (b). Under the assumption that $\rho > 1/2$, combining (6.7), (6.10), (6.13) and (6.14), we deduce that $\mathbb{E} \left[\left| \xi_i^n \right| \right] \leq K \left(\Delta_n^{1/2-\tau} \vee m_n \Delta_n^{1-\tau} \right)$ and $\mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \xi_i^n \right| \right] \leq K \left(\Delta_n^{1/2-2\tau} \vee m_n \Delta_n^{1-\tau} \right)$, which proves part (b) of the lemma. The assertion in (a) then follows from this estimate of $\mathbb{E} \left[\left| \xi_i^n \right| \right]$, the assumption $0 < \tau < 1/3$, and (6.6).

(c) Given $0 < \tau < 1/4$, applying part (b) with $m_n = 0$ and $Z_t = 1$ identically, we deduce

$$\sup_{1 \leq i \leq M_n} \left| \Delta N(L_n)_i^n \Delta_n / L_n - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \alpha_s ds \right| = o_p(1).$$

Since the process α is bounded, the variables $(\Delta N(L_n)_i^n \Delta_n / L_n)_{1 \leq i \leq M_n}$ are uniformly bounded with probability approaching one. That is, $\mathbb{P}(A_n) \rightarrow 1$. Similarly, using the bound in part (a), we deduce $\mathbb{P}(A_{n,i}) \rightarrow 1$, for each fixed i . Q.E.D.

Lemma 2. *Suppose that the conditions in Theorem 1 hold. Then for some constant $\eta > 0$ and an array $(\Phi_{i,j}^n)$ of \mathcal{G} -measurable random variables satisfying $\mathbb{E}[(\Phi_{i,j}^n)^2] \leq 1$, the following statements hold:*

- (a) $\left| \mathbb{E} \left[\Xi_{J_i+j}^n | \mathcal{K}_{J_i+j}^n \right] \right| \leq K \Phi_{i,j}^n \Delta_n$;
- (b) $\mathbb{E} \left[\left| \Xi_{J_i+j}^n \right|^4 \right] \leq K (h_n^{-v \wedge 4} + \Delta_n^2)$;
- (c) $\mathbb{E} \left[\left| \left(\tilde{Y}_{J_i+j}^n \right)^2 1_{\{|\tilde{Y}_{J_i+j}^n| \leq u_n\}} - \left(\tilde{Y}_{J_i+j}^{c,n} \right)^2 \right| \right] \leq K (h_n \Delta_n)^{3/2+\eta}$;
- (d) for each $l \in \{1, \dots, k_n\}$,

$$\begin{cases} \mathbb{E} \left[\left| U(l)_i^n - r(l) \sum_{j=0}^{\Delta N(L_n)_i^n - 5k_n} \left(\gamma_{J_i+j}^n \right)^2 \right| 1_{A_{n,i}} \right] \leq K L_n^{1/2} k_n^{1/2} / \Delta_n^{1/2}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| U(l)_i^n - r(l) \sum_{j=0}^{\Delta N(L_n)_i^n - 5k_n} \left(\gamma_{J_i+j}^n \right)^2 \right| 1_{A_n} \right] \leq K k_n^{1/2} / L_n^{1/2} \Delta_n^{1/2}. \end{cases}$$

Proof. To see part (a), we note by the triangle inequality that

$$\begin{aligned} \left| \mathbb{E} \left[\Xi_{J_i+j}^n | \mathcal{K}_{J_i+j}^n \right] \right| &\leq K \left(\left| \mathbb{E} \left[\hat{X}_{J_i+j}^{c,n} | \mathcal{K}_{J_i+j}^n \right] \right| + \left| \mathbb{E} \left[\hat{\epsilon}_{J_i+j}^n | \mathcal{K}_{J_i+j}^n \right] \right| + \left| \mathbb{E} \left[\tilde{X}_{J_i+j}^{c,n} \tilde{\epsilon}_{J_i+j}^n | \mathcal{K}_{J_i+j}^n \right] \right| \right) \\ &\leq K \Phi_{i,j}^n \left(h_n \Delta_n^{3/2} + \Delta_n + \frac{\Delta_n^{1/2}}{h_n^{v-1/2}} \right), \end{aligned}$$

where the second inequality follows from Lemma A2 and Lemma A5 in Jacod, Li, and Zheng (2019). The assertion of part (a) then follows from $h_n \asymp \Delta_n^{-1/2}$ and $v > 4$. Part (b) can be proved similarly by noting

$$\begin{aligned} \mathbb{E} \left[\left| \Xi_{J_i+j}^n \right|^4 \right] &\leq K \left(\mathbb{E} \left[\left| \hat{X}_{J_i+j}^{c,n} \right|^4 \right] + \mathbb{E} \left[\left| \hat{\epsilon}_{J_i+j}^n \right|^4 \right] + \mathbb{E} \left[\left| \tilde{X}_{J_i+j}^{c,n} \tilde{\epsilon}_{J_i+j}^n \right|^4 \right] \right) \\ &\leq K \left((h_n \Delta_n)^4 + h_n^{-v \wedge 4} + \Delta_n^2 \right). \end{aligned}$$

Part (c) is directly from Lemma A.6 of Jacod, Li, and Zheng (2019). We now turn to part (d). In restriction to $A_{n,i} \equiv \{\Delta N(L_n)_i^n \leq K L_n / \Delta_n \text{ for some } K > 0\}$, we can adapt the proof of Lemma A.7 of Jacod, Li, and Zheng (2019) with $N_t^n \leq K t / \Delta_n$ replaced by $\Delta N(L_n)_i^n \leq K L_n / \Delta_n$ and deduce the first assertion in part (d). The second assertion of part (d) then follows from the first assertion of part (d), a maximal inequality and the fact that $M_n \leq K / L_n$. Q.E.D.

The following lemma establishes the convergence rate of the spot estimator $\hat{\sigma}^2$.

Lemma 3. Under conditions (i,ii) of Theorem 1, we have

- (a) $\hat{\sigma}_{(i-1)L_n}^2 - \sigma_{(i-1)L_n}^2 = O_p\left(\Delta_n^{\tau/2} \vee \Delta_n^{1/4-\tau/2}\right)$.
(b) $\sup_{1 \leq i \leq M_n} \left| \hat{\sigma}_{(i-1)L_n}^2 - L_n^{-1} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds \right| = O_p\left(\Delta_n^{1/4-\tau} \vee \Delta_n^{11/40-3\tau/2}\right)$ if $\tau \in (0, 11/60)$.

Proof. We decompose $\hat{\sigma}_{(i-1)L_n}^2 - \sigma_{(i-1)L_n}^2 = \sum_{k=1}^5 \xi_i^{n,k}$ where

$$\begin{aligned} \xi_i^{n,1} &\equiv \frac{1}{L_n} \frac{1}{h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} \iota_{J_i+j}^n - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds, \\ \xi_i^{n,2} &\equiv \frac{1}{L_n} \frac{1}{h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} \left(\left(\tilde{Y}_{J_i+j}^n \right)^2 \mathbf{1}_{\{|\tilde{Y}_{J_i+j}^n| \leq u_n\}} - \left(\tilde{Y}_{J_i+j}^{c,n} \right)^2 \right), \\ \xi_i^{n,3} &\equiv \frac{1}{L_n} \left(\frac{\tilde{r}_{n,0}}{h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} \left(\gamma_{J_i+j}^n \right)^2 - \frac{1}{h_n^2 \phi_n} \sum_{l=-k'_n}^{k'_n} \bar{\phi}_{n,l} U(|l|)_i^n \right), \\ \xi_i^{n,4} &\equiv \frac{1}{L_n} \frac{1}{h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} \Xi_{J_i+j}^n, \\ \xi_i^{n,5} &\equiv \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds - \sigma_{(i-1)L_n}^2. \end{aligned}$$

We also note that $\hat{\sigma}_{(i-1)L_n}^2 - L_n^{-1} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds = \sum_{k=1}^4 \xi_i^{n,k}$. Below, we provide estimates for $\xi_i^{n,k}$, $1 \leq k \leq 5$.

Case $k = 1$. By definition, we can rewrite $\xi_i^{n,1}$ as

$$\xi_i^{n,1} = \frac{1}{L_n} \frac{1}{h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} \sum_{u=1}^{h_n-1} (w_u^n)^2 \Delta_{J_i+j+u}^n \bar{V} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds.$$

Changing the order of the two summations yields

$$\begin{aligned} \xi_i^{n,1} &= \frac{1}{L_n} \frac{1}{h_n \phi_n} \sum_{v=1}^{\Delta N(L_n)_i^n - 1} \Delta_{J_i+v}^n \bar{V} \sum_{l=1 \vee (v+h_n - \Delta N(L_n)_i^n)}^{v \wedge (h_n-1)} (w_l^n)^2 - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds \\ &= \frac{1}{L_n} \sum_{v=h_n-1}^{\Delta N(L_n)_i^n - h_n + 1} \Delta_{J_i+v}^n \bar{V} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds + e_i^n, \end{aligned} \quad (6.15)$$

where the boundary term e_i^n is given by

$$\begin{aligned} e_i^n &= \frac{1}{L_n} \frac{1}{h_n \phi_n} \left(\sum_{v=1}^{h_n-2} \Delta_{J_i+v}^n \bar{V} \sum_{l=1}^v (w_l^n)^2 + \sum_{v=\Delta N(L_n)_i^n - h_n + 2}^{\Delta N(L_n)_i^n - 1} \Delta_{J_i+v}^n \bar{V} \sum_{l=v+h_n - \Delta N(L_n)_i^n}^{h_n-1} (w_l^n)^2 \right) \\ &\leq \frac{K}{L_n} \left(\sum_{v=1}^{h_n-2} \Delta_{J_i+v}^n \bar{V} + \sum_{v=\Delta N(L_n)_i^n - h_n + 2}^{\Delta N(L_n)_i^n - 1} \Delta_{J_i+v}^n \bar{V} \right). \end{aligned}$$

Due to the boundedness of σ^2 , (6.3) and our maintained assumptions on L_n and h_n , we have

$$\mathbb{E} [|e_i^n|] \leq Kh_n \Delta_n / L_n \leq K \Delta_n^{1/2-\tau}. \quad (6.16)$$

Similarly, we get

$$\mathbb{E} \left[\frac{1}{L_n} \left| \sum_{v=h_n-1}^{\Delta N(L_n)_i^n - h_n + 1} \int_{T(n, J_i+v-1)}^{T(n, J_i+v)} \sigma_s^2 ds - \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds \right| \right] \leq K \Delta_n^{1/2-\tau}. \quad (6.17)$$

Combining (6.16) and (6.17) yields

$$\begin{cases} \mathbb{E} \left[|\xi_i^{n,1}| \right] \leq K \Delta_n^{1/2-\tau}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} |\xi_i^{n,1}| \right] \leq K M_n \Delta_n^{1/2-\tau} \leq K \Delta_n^{1/2-2\tau}, \end{cases} \quad (6.18)$$

where we used a maximal inequality and the fact $M_n \leq K/L_n$ for deriving the second (uniform) estimate in (6.18).

Case $k = 2$. Recall the definitions of $A_{n,i}$ and A_n from Lemma 1. In view of Lemma 2(c) and our maintained assumptions on M_n , L_n and h_n , we derive using the triangle inequality and a maximal inequality that, for some $\eta > 0$,

$$\begin{cases} \mathbb{E} \left[|\xi_i^{n,2}| 1_{A_{n,i}} \right] \leq K \Delta_n^{1/4+\eta}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} |\xi_i^{n,2}| 1_{A_n} \right] \leq K M_n \Delta_n^{1/4+\eta} \leq K \Delta_n^{1/4+\eta-\tau}. \end{cases} \quad (6.19)$$

Case $k = 3$. Let $\lambda_i^n \equiv \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} (\gamma_{J_i+j}^n)^2$. Note that $\tilde{r}_{n,0} = h_n^{-1} \sum_{|l| < \infty} r(l) \bar{\phi}_{n,l}$. Hence, we can decompose $\xi_i^{n,3} = \Theta_i^n + \Theta_i^m$, where

$$\Theta_i^n \equiv \frac{\lambda_i^n}{L_n h_n^2 \phi_n} \sum_{|l| > k'_n} \bar{\phi}_{n,l} r(l) \quad \text{and} \quad \Theta_i^m \equiv \frac{1}{L_n h_n^2 \phi_n} \sum_{|l| \leq k'_n} \bar{\phi}_{n,l} (r(l) \lambda_i^n - U(|l|)_i^n).$$

Due to the boundedness of $\bar{\phi}_{n,l}$ and γ , and the fact that $\sum_{|l| > k'_n} |r(l)| \leq K/k_n'^{v-1}$, we deduce

$$|\Theta_i^n| 1_{A_n} \leq |\Theta_i^n| 1_{A_{n,i}} \leq K/k_n'^{v-1} \leq K \Delta_n^{(v-1)/8}. \quad (6.20)$$

Next, we further decompose $\Theta_i^m = \Upsilon_i^{n,1} + \Upsilon_i^{n,2}$, where

$$\begin{aligned} \Upsilon_i^{n,1} &\equiv \frac{1}{L_n h_n^2 \phi_n} \sum_{|l| \leq k'_n} \bar{\phi}_{n,l} \left(r(l) \sum_{j=0}^{\Delta N(L_n)_i^n - 5k_n} (\gamma_{J_i+j}^n)^2 - U(|l|)_i^n \right), \\ \Upsilon_i^{n,2} &\equiv -\frac{1}{L_n h_n^2 \phi_n} \sum_{|l| \leq k'_n} \bar{\phi}_{n,l} r(l) \sum_{\Delta N(L_n)_i^n - h_n}^{\Delta N(L_n)_i^n - 5k_n} (\gamma_{J_i+j}^n)^2. \end{aligned}$$

Note that $\sum_{l \in \mathbb{Z}} |r(l)| < \infty$. Hence, uniformly in i ,

$$|\Upsilon_i^{n,2}| \leq K/(L_n h_n) = K \Delta_n^{1/2-\tau}. \quad (6.21)$$

In view of Lemma 2(d), as well as the assumptions on k_n, k'_n, L_n and h_n , we deduce

$$\begin{aligned} \mathbb{E} \left[\left| \Upsilon_i^{n,1} \right| 1_{A_{n,i}} \right] &\leq K \frac{k'_n k_n^{1/2}}{L_n^{1/2} h_n^2 \Delta_n^{1/2}} = K \Delta_n^{11/40-\tau/2}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \Upsilon_i^{n,1} \right| 1_{A_n} \right] &\leq K M_n \sup_{1 \leq i \leq M_n} \mathbb{E} \left[\left| \Upsilon_i^{n,1} \right| 1_{A_n} \right] \leq K \Delta_n^{11/40-3\tau/2}. \end{aligned} \quad (6.22)$$

Combining (6.21) and (6.22), we deduce (recall that $\tau \in (0, 1/3)$)

$$\mathbb{E} \left[\left| \Theta_i^{n,1} \right| 1_{A_{n,i}} \right] \leq K \Delta_n^{11/40-\tau/2}, \quad \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \Theta_i^{n,1} \right| 1_{A_n} \right] \leq K \Delta_n^{11/40-3\tau/2}.$$

This estimate, together with (6.20) and $v > 4$, implies

$$\begin{cases} \mathbb{E} \left[\left| \xi_i^{n,3} \right| 1_{A_{n,i}} \right] \leq K \Delta_n^{11/40-\tau/2}; \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \xi_i^{n,3} \right| 1_{A_n} \right] \leq K \Delta_n^{11/40-3\tau/2}. \end{cases} \quad (6.23)$$

Case $k = 4$. We decompose $\xi_i^{n,4} = \zeta_i^{n,1} + \zeta_i^{n,2}$, where

$$\begin{aligned} \zeta_i^{n,1} &\equiv \frac{1}{L_n h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} (\Xi_{J_i+j}^n - \mathbb{E} [\Xi_{J_i+j}^n | \mathcal{K}_{J_i+j}^n]), \\ \zeta_i^{n,2} &\equiv \frac{1}{L_n h_n \phi_n} \sum_{j=0}^{\Delta N(L_n)_i^n - h_n} \mathbb{E} [\Xi_{J_i+j}^n | \mathcal{K}_{J_i+j}^n]. \end{aligned}$$

By Lemma 2(a), we deduce

$$\mathbb{E} \left[\left| \zeta_i^{n,2} \right| 1_{A_{n,i}} \right] \leq K \Delta_n^{1/2}, \quad \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,2} \right| 1_{A_n} \right] \leq K \Delta_n^{1/2-\tau}. \quad (6.24)$$

Note that $\Xi_{J_i+j}^n - \mathbb{E}[\Xi_{J_i+j}^n | \mathcal{K}_{J_i+j}^n]$ is $\mathcal{K}_{J_i+j}^n$ -conditionally mean zero and $\mathcal{K}_{J_i+j+2h_n}^n$ -measurable. Hence, we can use the Cauchy–Schwarz inequality to deduce that

$$\begin{aligned} \mathbb{E} \left[\left| \zeta_i^{n,1} \right|^2 1_{A_{n,i}} \right] &\leq \frac{K}{L_n^2 h_n} \sum_{j=0}^{KL_n/\Delta_n - h_n} \mathbb{E} \left[(\Xi_{J_i+j}^n - \mathbb{E} [\Xi_{J_i+j}^n | \mathcal{K}_{J_i+j}^n])^2 \right] \\ &\leq \frac{K}{L_n h_n \Delta_n} \left(\Delta_n + \frac{1}{h_n^{v/2 \wedge 2}} \right) \leq K \Delta_n^{1/2-\tau}, \end{aligned}$$

where the second line is implied by Lemma 2(b) and $v > 0$. Then, a maximal inequality implies $\left\| \sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,1} \right| 1_{A_n} \right\|_2 \leq K \Delta_n^{1/4-\tau}$, where $\|\cdot\|_2$ denotes the L_2 norm. From these estimates, it readily follows that

$$\mathbb{E} \left[\left| \zeta_i^{n,1} \right| 1_{A_{n,i}} \right] \leq K \Delta_n^{1/4-\tau/2}, \quad \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \zeta_i^{n,1} \right| 1_{A_n} \right] \leq K \Delta_n^{1/4-\tau}. \quad (6.25)$$

Combining (6.24) and (6.25), we deduce

$$\begin{cases} \mathbb{E} \left[\left| \xi_i^{n,4} \right| 1_{A_{n,i}} \right] \leq K \Delta_n^{1/4-\tau/2}, \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \xi_i^{n,4} \right| 1_{A_n} \right] \leq K \Delta_n^{1/4-\tau}. \end{cases} \quad (6.26)$$

Case $k = 5$. Due to classic estimates of Itô semimartingales,

$$\mathbb{E} \left[\left| \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \left(\sigma_s^2 - \sigma_{(i-1)L_n}^2 \right) ds \right| \right] \leq K L_n^{1/2} \leq K \Delta_n^{\tau/2}. \quad (6.27)$$

We are now ready to prove the assertions of the lemma. From (6.18), (6.19), (6.23), (6.26) and (6.27), we deduce

$$\begin{cases} \mathbb{E} \left[\left| \hat{\sigma}_{(i-1)L_n}^2 - \sigma_{(i-1)L_n}^2 \right| 1_{A_{n,i}} \right] \leq K \left(\Delta_n^{\tau/2} \vee \Delta_n^{1/4-\tau/2} \right), \\ \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \hat{\sigma}_{(i-1)L_n}^2 - (1/L_n) \int_{(i-1)L_n}^{iL_n} \sigma_s^2 ds \right| 1_{A_n} \right] \leq K \left(\Delta_n^{1/4-\tau} \vee \Delta_n^{11/40-3\tau/2} \right). \end{cases}$$

By Lemma 1(c), $\mathbb{P}(A_{n,i}) \rightarrow 1$ and $\mathbb{P}(A_n) \rightarrow 1$. The assertions of the lemma then readily follows from the inequalities displayed above. *Q.E.D.*

The following lemma establishes the convergence rate of the spot estimator $\hat{\gamma}^2$ for the variance of microstructure noise.

Lemma 4. *Under the assumptions of Theorem 1, we have (a)*

$$\hat{\gamma}_{(i-1)L_n}^2 - \gamma_{(i-1)L_n}^2 = O_p \left(\Delta_n^{\tau/2} \vee \Delta_n^{(3-5\tau)/10} \right),$$

and (b)

$$\begin{aligned} & \sup_{1 \leq i \leq M_n} \left| \hat{\gamma}_{(i-1)L_n}^2 - L_n^{-1} \int_{(i-1)L_n}^{iL_n} \gamma_s^2 ds \right| \\ &= O_p \left(\Delta_n^{1/2-2\tau} \vee \Delta_n^{2/5-3\tau/2} \vee \log(\Delta_n^{-1})^{1/2} \Delta_n^{\tau/2} \right). \end{aligned}$$

Proof. Recall the definitions of $A_{n,i}$ and A_n from Lemma 1. Let

$$B_{n,i} \equiv \left\{ \frac{\Delta_n \Delta N(L_n)_i^n}{L_n} > \eta, \text{ for some } \eta \right\}, \quad B_n \equiv \left\{ \frac{\Delta_n \Delta N(L_n)_i^n}{L_n} > \eta, \text{ for some } \eta \text{ and all } i \right\}.$$

Applying Lemma 1 with $m_n = 0$ and $Z = 1$ implies that

$$\begin{cases} |\Delta_n \Delta N(L_n)_i^n / L_n - \alpha_{(i-1)L_n}| = o_p(1) & \text{for } \tau \in (0, 1/3), \\ \sup_{1 \leq i \leq M_n} \left| \frac{\Delta_n \Delta N(L_n)_i^n}{L_n} - (1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s ds \right| = o_p(1) & \text{for } \tau \in (0, 1/4). \end{cases}$$

Since α is bounded away from zero, we must have $\mathbb{P}(B_{n,i}) \rightarrow 1$ for $0 < \tau < 1/3$ and $\mathbb{P}(B_n) \rightarrow 1$ for $0 < \tau < 1/4$. Hence, to prove part (a) and part (b), we can restrict our calculations to $A_{n,i} \cap B_{n,i}$ and $A_n \cap B_n$, respectively, without loss of generality.

(a) We decompose $\hat{\gamma}_{(i-1)L_n}^2 - \gamma_{(i-1)L_n}^2 = \xi_i^n + \xi_i^m$, where

$$\begin{aligned} \xi_i^n &\equiv \frac{L_n}{\Delta_n \Delta N(L_n)_i^n} \left(\frac{\Delta_n}{L_n} U(0)_i^n - \alpha_{(i-1)L_n} \gamma_{(i-1)L_n}^2 \right), \\ \xi_i^m &\equiv -\frac{L_n \gamma_{(i-1)L_n}^2}{\Delta_n \Delta N(L_n)_i^n} \left(\frac{\Delta_n \Delta N(L_n)_i^n}{L_n} - \alpha_{(i-1)L_n} \right). \end{aligned}$$

Lemma 1(a) applied with $m_n = 5k_n$ and $Z = \gamma^2$, coupled with the restrictions $k_n \asymp \Delta_n^{-1/5}$ and $0 < \tau < 1/3$, implies

$$\mathbb{E} \left[\left| \frac{\Delta_n}{L_n} \sum_{j=0}^{\Delta N(L_n)_i^n - 5k_n} (\gamma_{J_i+j}^n)^2 - \alpha_{(i-1)L_n} \gamma_{(i-1)L_n}^2 \right| \right] \leq K \Delta_n^{\tau/2}. \quad (6.28)$$

Using Lemma 2(d) with $l = 0$, we derive

$$\mathbb{E} \left[\left| \frac{\Delta_n}{L_n} U(0)_i^n - \frac{\Delta_n}{L_n} \sum_{j=0}^{\Delta N(L_n)_i^n - 5k_n} (\gamma_{J_i+j}^n)^2 \right| 1_{A_{n,i}} \right] \leq K \frac{\Delta_n^{1/2} k_n^{1/2}}{L_n^{1/2}} \leq K \Delta_n^{(3-5\tau)/10}. \quad (6.29)$$

Combining (6.28) and (6.29), we deduce

$$\mathbb{E} \left[\left| \frac{\Delta_n}{L_n} U(0)_i^n - \alpha_{(i-1)L_n} \gamma_{(i-1)L_n}^2 \right| 1_{A_{n,i}} \right] \leq K \left(\Delta_n^{\tau/2} \vee \Delta_n^{(3-5\tau)/10} \right).$$

Note that, in restriction to $B_{n,i}$, $L_n / (\Delta_n \Delta N(L_n)_i^n)$ is bounded. Hence, the displayed estimate above further implies

$$\mathbb{E} [|\xi_i^n| 1_{A_{n,i} \cap B_{n,i}}] \leq K \mathbb{E} \left[\left| \frac{\Delta_n}{L_n} U(0)_i^n - \alpha_{(i-1)L_n} \gamma_{(i-1)L_n}^2 \right| 1_{A_{n,i}} \right] \leq K \left(\Delta_n^{\tau/2} \vee \Delta_n^{(3-5\tau)/10} \right). \quad (6.30)$$

Due to the boundedness of γ and Lemma 1(a) applied with $m_n = 0$ and $Z = 1$, we deduce that

$$\mathbb{E} [|\xi_i^m| 1_{A_{n,i} \cap B_{n,i}}] \leq K \mathbb{E} \left[\left| \frac{\Delta_n \Delta N(L_n)_i^n}{L_n} - \alpha_{(i-1)L_n} \right| \right] \leq K \Delta_n^{\tau/2}. \quad (6.31)$$

The assertion of part (a) then follows from (6.30) and (6.31).

(b) We observe

$$\begin{aligned}
& \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \hat{\gamma}_{(i-1)L_n}^2 - L_n^{-1} \int_{(i-1)L_n}^{iL_n} \gamma_s^2 ds \right| 1_{A_n \cap B_n} \right] \\
& \leq \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{U(0)_i^n}{\Delta N(L_n)_i^n} - \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} \right| 1_{A_n \cap B_n} \right] \\
& \quad + \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \gamma_s^2 ds \right| 1_{A_n \cap B_n} \right].
\end{aligned} \tag{6.32}$$

It remains to bound the two terms on the majorant side of (6.32).

For the first term on the right-hand side of (6.32), we note that

$$\begin{aligned}
& \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{U(0)_i^n}{\Delta N(L_n)_i^n} - \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} \right| 1_{A_n \cap B_n} \right] \\
& \leq \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{U(0)_i^n}{\Delta N(L_n)_i^n} - \frac{(1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\Delta_n \Delta N(L_n)_i^n / L_n} \right| 1_{A_n \cap B_n} \right] \\
& \quad + \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{(1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\Delta_n \Delta N(L_n)_i^n / L_n} - \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} \right| 1_{A_n \cap B_n} \right].
\end{aligned} \tag{6.33}$$

Recall that, in restriction to B_n , $L_n / (\Delta_n \Delta N(L_n)_i^n) \leq K$. Hence,

$$\begin{aligned}
& \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{U(0)_i^n}{\Delta N(L_n)_i^n} - \frac{(1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\Delta_n \Delta N(L_n)_i^n / L_n} \right| 1_{A_n \cap B_n} \right] \\
& \leq K \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{\Delta_n}{L_n} U(0)_i^n - (1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds \right| 1_{A_n \cap B_n} \right] \\
& \leq K \left(\Delta_n^{1/2-2\tau} \vee \Delta_n^{2/5-3\tau/2} \right),
\end{aligned} \tag{6.34}$$

where the second inequality follows from Lemma 1(b) and Lemma 2(d). In addition, we note that

$$\begin{aligned}
& \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{(1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\Delta_n \Delta N(L_n)_i^n / L_n} - \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} \right| 1_{A_n \cap B_n} \right] \\
& \leq K \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{\Delta_n \Delta N(L_n)_i^n}{L_n} - (1/L_n) \int_{(i-1)L_n}^{iL_n} \alpha_s ds \right| 1_{A_n \cap B_n} \right] \\
& \leq K \Delta_n^{1/2-2\tau},
\end{aligned} \tag{6.35}$$

where the first inequality holds because the processes α , γ , $1/\alpha$ are bounded and $L_n / (\Delta_n \Delta N(L_n)_i^n) \leq K$ in restriction to B_n ; and the second inequality follows from Lemma 1(b). Combining (6.33),

(6.34) and (6.35), we deduce

$$\mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{U(0)_i^n}{\Delta N(L_n)_i^n} - \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} \right| 1_{A_n \cap B_n} \right] \leq K \left(\Delta_n^{1/2-2\tau} \vee \Delta_n^{2/5-3\tau/2} \right). \quad (6.36)$$

Finally, we consider the second term on the right-hand side of (6.32). Denote $\bar{\alpha}_i^n \equiv L_n^{-1} \int_{(i-1)L_n}^{iL_n} \alpha_s ds$. We observe

$$\begin{aligned} & \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{\int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds}{\int_{(i-1)L_n}^{iL_n} \alpha_s ds} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \gamma_s^2 ds \right| 1_{A_n \cap B_n} \right] \\ & \leq \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \alpha_s \gamma_s^2 ds - \left(\frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \gamma_s^2 ds \right) \bar{\alpha}_i^n \right| \right] \\ & \leq \mathbb{E} \left[\sup_{1 \leq i \leq M_n} \left| \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} (\alpha_s - \bar{\alpha}_i^n) \gamma_s^2 ds \right| \right] \\ & \leq K \mathbb{E} \left[\sup_{s,t \in [0,1], |s-t| \leq L_n} |\alpha_s - \alpha_t| \right] \leq K (L_n \log(L_n^{-1}))^{1/2}, \end{aligned}$$

where the first inequality is due to the fact α_t is bounded away from zero; the second inequality is obvious; and the last line follows from the boundedness of γ and Theorem 1 in Fischer and Nappo (2009). The assertion of part (b) then follows from this estimate and (6.36). *Q.E.D.*

We are now ready to prove Theorem 1.

Proof of Theorem 1. By Lemma 3 and Lemma 4, for each t ,

$$\begin{aligned} |\hat{\sigma}_t^2 - \sigma_t^2| & \leq \left| \hat{\sigma}_t^2 - \sigma_{(i-1)L_n}^2 \right| + \left| \sigma_{(i-1)L_n}^2 - \sigma_t^2 \right| = O_p \left(\Delta_n^{\tau/2} \vee \Delta_n^{1/4-\tau/2} \right), \\ |\hat{\gamma}_t^2 - \gamma_t^2| & \leq \left| \hat{\gamma}_t^2 - \gamma_{(i-1)L_n}^2 \right| + \left| \gamma_{(i-1)L_n}^2 - \gamma_t^2 \right| = O_p \left(\Delta_n^{\tau/2} \vee \Delta_n^{(3-5\tau)/10} \right), \end{aligned}$$

where $i = \max\{j : 1 \leq j \leq M_n, (j-1)L_n \leq t\}$. This proves the first assertion of the theorem. The second assertion of the theorem follows directly from the uniform bounds in Lemmas 3 and 4, under the condition $0 < \tau < 11/60$. *Q.E.D.*

6.2 Proof of Theorem 2

Proof of Theorem 2. By localization, we can strengthen condition (iv) by assuming that the processes (σ_t^2, γ_t^2) take values in $\mathcal{K} \times \mathcal{K}$ for some bounded closed interval $\mathcal{K} \subset (0, \infty)$. Consequently, Z also takes values in a compact set and $f_Z(\cdot)$ is compactly supported. Under condition (ii),

$$\sup_{1 \leq i \leq M_n} \left\| \hat{\theta}_{(i-1)L_n} - L_n^{-1} \int_{(i-1)L_n}^{iL_n} \theta_s ds \right\| = o_p(1).$$

By enlarging \mathcal{K} slightly if necessary, we further deduce that $(\hat{\sigma}_{(i-1)L_n}^2, \hat{\gamma}_{(i-1)L_n}^2)_{1 \leq i \leq M_n}$ takes value in $\mathcal{K} \times \mathcal{K}$ with probability approaching one. By condition (v), $g(\cdot)$ is Lipschitz on $\mathcal{K} \times \mathcal{K}$. Hence,

$$\left| \hat{Z}_s - Z_s \right| = \left\| g(\hat{\theta}_s) - g(\theta_s) \right\| \leq K \left\| \hat{\theta}_s - \theta_s \right\|. \quad (6.37)$$

For each x , we set

$$f_{Z,n}(x) \equiv \int_0^1 \frac{1}{\delta_n} H \left(\frac{Z_s - x}{\delta_n} \right) ds.$$

We observe that

$$\begin{aligned} & \left| \hat{f}_Z(x) - f_{Z,n}(x) \right| \\ & \leq \frac{1}{\delta_n} \int_0^1 \left| H \left(\frac{\hat{Z}_s - x}{\delta_n} \right) - H \left(\frac{Z_s - x}{\delta_n} \right) \right| ds \leq \frac{K}{\delta_n^2} \int_0^1 \left| \hat{Z}_s - Z_s \right| ds \\ & \leq \frac{K}{\delta_n^2} \int_0^1 \left\| \hat{\theta}_s - \theta_s \right\| ds \\ & \leq \frac{K}{\delta_n^2} \left\{ \sum_{i=1}^{M_n} \int_{(i-1)L_n}^{iL_n} \left(\left\| \hat{\theta}_s - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \theta_u du \right\| + \left\| \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \theta_u du - \theta_s \right\| \right) ds \right. \\ & \quad \left. + \int_{M_n L_n}^1 \left\| \hat{\theta}_s - \theta_s \right\| ds \right\} \\ & \leq \frac{K}{\delta_n^2} \sup_{1 \leq i \leq M_n} \left\| \hat{\theta}_{(i-1)L_n} - \frac{1}{L_n} \int_{(i-1)L_n}^{iL_n} \theta_s du \right\| \\ & \quad + \frac{K}{\delta_n^2 L_n} \sum_{i=1}^{M_n} \int_{(i-1)L_n}^{iL_n} \int_{(i-1)L_n}^{iL_n} \left\| \theta_u - \theta_s \right\| dud s + \frac{K}{\delta_n^2} \int_{M_n L_n}^1 \left\| \hat{\theta}_s - \theta_s \right\| ds, \end{aligned} \quad (6.38)$$

where the first inequality is obvious; the second inequality follows from the smoothness requirement on the kernel function $H(\cdot)$; the third inequality is by (6.37); the fourth inequality follows from the triangle inequality; and the last inequality is obvious.

By condition (ii), the first term on the majorant side of (6.38) is $O_p(a_n \delta_n^{-2})$. The second term is $O_p(\Delta_n^{\tau/2} \delta_n^{-2})$, which follows from $\mathbb{E}[\|\theta_u - \theta_s\|] \leq K L_n^{1/2}$, $M_n L_n = O(1)$ and $L_n \asymp \Delta_n^\tau$. Since $\hat{\theta}$ and θ are uniformly bounded with probability approaching one, the third term on the majorant side is $O_p(\Delta_n^\tau \delta_n^{-2})$. In view of these estimates, we see that (6.38) further implies (recalling $\bar{a}_n = a_n \vee \Delta_n^{\tau/2}$)

$$\hat{f}_Z(x) - f_{Z,n}(x) = O_p(\delta_n^{-2} \bar{a}_n). \quad (6.39)$$

Finally, we provide bound for the difference $f_{Z,n}(x) - f_Z(x)$. By the definition of occupation density and a change of variable, we can rewrite

$$f_{Z,n}(x) = \int_{\mathbb{R}} \frac{1}{\delta_n} H \left(\frac{v - x}{\delta_n} \right) f_Z(v) dv = \int_{\mathbb{R}} H(z) f_Z(x + z \delta_n) dz.$$

Therefore, we get

$$\begin{aligned}\mathbb{E} [|f_{Z,n}(x) - f_Z(x)|] &\leq \int_{\mathbb{R}} H(z) \mathbb{E} [|f_Z(x + z\delta_n) - f_Z(x)|] dz \\ &\leq K\delta_n^\beta \int_{\mathbb{R}} H(z) |z|^\beta dz \leq K\delta_n^\beta.\end{aligned}$$

Combining this estimate with (6.39), we deduce the assertion of the theorem.

Q.E.D.

References

- AÏT-SAHALIA, Y., AND J. JACOD (2014): *High-Frequency Financial Econometrics*. Princeton University Press.
- AÏT-SAHALIA, Y., J. JACOD, AND J. LI (2012): “Testing for jumps in noisy high frequency data,” *Journal of Econometrics*, 168(2), 207–222.
- ANDERSEN, T. G., T. BOLLERSLEV, AND F. X. DIEBOLD (2010): “Parametric and nonparametric volatility measurement,” in *Handbook of Financial Econometrics*, ed. by Y. Aït-Sahalia, and L. P. Hansen, chap. 2, pp. 67–137. Elsevier, Amsterdam.
- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND H. EBENS (2001): “The distribution of realized stock return volatility,” *Journal of Financial Economics*, 61(1), 43–76.
- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2001): “The distribution of realized exchange rate volatility,” *Journal of the American Statistical Association*, 96(453), 42–55.
- BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): “Designing realized kernels to measure the ex post variation of equity prices in the presence of noise,” *Econometrica*, 76(6), 1481–1536.
- BIBINGER, M., N. HAUTSCH, P. MALEC, AND M. REISS (2019): “Estimating the spot covariation of asset prices—Statistical theory and empirical evidence,” *Journal of Business & Economic Statistics*, 37(3), 419–435.
- BRUNNERMEIER, M. K., AND L. H. PEDERSEN (2009): “Market liquidity and funding liquidity,” *Review of Financial Studies*, 22(6), 2201–2238.
- COMTE, F., AND E. RENAULT (1998): “Long memory in continuous time stochastic volatility models,” *Mathematical Finance*, 8(4), 291–323.
- EASLEY, D., N. M. KIEFER, M. O’HARA, AND J. B. PAPERMAN (1996): “Liquidity, information, and infrequently traded stocks,” *Journal of Finance*, 51(4), 1405–1436.
- FISCHER, M., AND G. NAPPO (2009): “On the moments of the modulus of continuity of Itô processes,” *Stochastic Analysis and Applications*, 28(1), 103–122.

- FOSTER, D. P., AND D. B. NELSON (1996): “Continuous record asymptotics for rolling sample variance estimators,” *Econometrica*, 64(1), 139–174.
- GEMAN, D., AND J. HOROWITZ (1980): “Occupation densities,” *Annals of Probability*, 8(1), 1–67.
- GLOSTEN, L. R., AND P. R. MILGROM (1985): “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of Financial Economics*, 14(1), 71–100.
- HANSEN, P. R., AND A. LUNDE (2006): “Realized variance and market microstructure noise,” *Journal of Business & Economic Statistics*, 24(2), 127–161.
- HASBROUCK, J. (2007): *Empirical Market Microstructure*. Oxford University Press New York.
- JACOD, J., Y. LI, P. A. MYKLAND, M. PODOLSKIJ, AND M. VETTER (2009): “Microstructure noise in the continuous case: The pre-averaging approach,” *Stochastic Processes and Their Applications*, 119(7), 2249–2276.
- JACOD, J., Y. LI, AND X. ZHENG (2017): “Statistical properties of microstructure noise,” *Econometrica*, 85(4), 1133–1174.
- (2019): “Estimating the integrated volatility with tick observations,” *Journal of Econometrics*, 208(1), 80–100.
- JACOD, J., M. PODOLSKIJ, AND M. VETTER (2010): “Limit theorems for moving averages of discretized processes plus noise,” *Annals of Statistics*, 38(3), 1478–1545.
- JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Springer, Heidelberg.
- KRISTENSEN, D. (2010): “Nonparametric filtering of the realized spot volatility: A kernel-based approach,” *Econometric Theory*, 26(1), 60–93.
- LI, J. (2013): “Robust estimation and inference for jumps in noisy high-frequency data: A local-to-continuity theory for the pre-averaging method,” *Econometrica*, 81(4), 1673–1693.
- LI, J., V. TODOROV, AND G. TAUCHEN (2013): “Volatility occupation times,” *Annals of Statistics*, 41(4), 1865–1891.
- (2016): “Estimating the volatility occupation time via regularized Laplace inversion,” *Econometric Theory*, 32(5), 1253–1288.
- MANCINI, C. (2001): “Disentangling the jumps of the diffusion in a Geometric jumping Brownian Motion,” *Giornale dell’Istituto Italiano degli Attuari*, LXIV, 19–47.
- MANCINI, C., V. MATTIUSI, AND R. RENÒ (2015): “Spot volatility estimation using delta sequences,” *Finance and Stochastics*, 19(2), 261–293.
- MARCUS, M. B., AND J. ROSEN (2006): *Markov Processes, Gaussian Processes, and Local Times*. Cambridge University Press.

- REISS, M. (2011): “Asymptotic equivalence for inference on the volatility from noisy observations,” *Annals of Statistics*, 39(2), 772–802.
- REVUZ, D., AND M. YOR (1999): *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, Germany.
- ROLL, R. (1984): “A simple implicit measure of the effective bid-ask spread in an efficient market,” *Journal of Finance*, 39(4), 1127–1139.
- XIU, D. (2010): “Quasi-maximum likelihood estimation of volatility with high-frequency data,” *Journal of Econometrics*, 159(1), 235–250.
- ZHANG, L. (2006): “Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach,” *Bernoulli*, 12(6), 1019–1043.
- ZHANG, L., P. A. MYKLAND, AND Y. AÏT-SAHALIA (2005): “A tale of two time scales: Determining integrated volatility with noisy high-frequency data,” *Journal of the American Statistical Association*, 100(472), 1394–1411.
- ZU, Y., AND H. P. BOSWIJK (2014): “Estimating spot volatility with high-frequency financial data,” *Journal of Econometrics*, 181(2), 117–135.