Journal of Financial Econometrics, 2017, 1–13
doi: 10.1093/jjfinec/nbx034
Correspondence

Comment on: Limit of Random Measures Associated with the Increments of a Brownian Semimartingale*

Jia Li[1] and Dacheng Xiu[2]

[1]Department of Economics, Duke University, Box 90097, Durham, NC, 27708, jl410@duke.edu and [2]Booth School of Business, University of Chicago Chicago, IL, 60637, dacheng.xiu@chicagobooth.edu

We thank the Editors' invitation for the opportunity of contributing to this special issue as a celebration of Professor Jean Jacod's seminal work originally written in 1994 (Jacod, 1994). This paper established general limit theorems for integrated volatility functionals, and provided theoretical tools that eventually changed the landscape of theoretical research concerning high-frequency data. This impact is also largely due to Professor Jacod's continuous contribution to a broad variety of challenging issues in the area of high-frequency financial econometrics, including volatility estimation, jumps, and microstructure noise, as well as a large body of mathematical results collected in Jacod and Shiryaev (2003), Jacod and Protter (2012), and Aït-Sahalia and Jacod (2014).

Professor Jacod's work has deeply influenced generations of researchers. We are grateful for having learnt and for continuing to learn from him. Perhaps the best way of celebrating the original contribution of Jacod (1994) is to review the progress in the past 20 years on integrated volatility estimation, to which results of this paper are directly applicable; and to discuss how this paper has influenced our thinking about econometric inference of stochastic volatility in general.

We thus organize our discussion as follows. Section 1 reviews the baseline problem of estimating integrated volatility. We discuss in Section 2 the efficient estimation of general integrated volatility functionals, and show in Section 3 how these functionals can be used to form integrated moment conditions for estimating asset pricing and market microstructure models. Section 4 contains some concluding remarks.

1 Integrated Volatility Estimation

The increasing availability of transaction-level data presents a unique opportunity and yet substantial challenges in risk measurement and management—an important agenda of the field of financial econometrics. Exploring this intraday dataset is certainly worthwhile, as it

* We thank the Editor (Federico Bandi) for inviting us to contribute to this special issue.

brings in feasible solutions to some problems that would otherwise be difficult to address with daily data. First, daily returns disguise informative intraday fluctuations. One striking example is the IBM stock on October 10, 2008, which ended slightly above 1% after having had a 10% rollercoaster ride within that day. Second, for stocks that are newly offered to the public, historical records are simply unavailable. Third, standard assumptions on stationarity, dependence, and heteroscedasticity in classic time series are neither essential nor relevant for intraday data. All these features make high-frequency data particularly attractive for measuring realizations of volatilities that are crucial for risk management in practice and empirical research in economics and finance.

The baseline model for volatility estimation assumes that the logarithm of the efficient price process X_t follows a continuous Itô semimartingale defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathrm{P})$, which satisfies

$$X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \tag{1.1}$$

where μ_t is the drift, σ_t is the stochastic volatility process, and W is a Brownian motion. A simple measure of volatility is the quadratic variation of X_t, namely, $\int_0^T c_s \mathrm{d}s$, where $c_t = \sigma_t^2$ is the instantaneous (i.e., spot) variance and $\{X_{i\Delta_n}\}_{i\geq 1}$ is a sequence of observations sampled regularly.

The most popular estimator, realized volatility, introduced to econometrics by Andersen *et al.* (2003) is simply the sum of squared returns:

$$\mathrm{RV}_n = \sum_{i=1}^{[T/\Delta_n]} \left(\Delta_i^n X\right)^2,$$

where $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$. To demonstrate the statistical properties of this estimator, it requires asymptotic analysis of the sum of squared increments of semimartingales. Jacod (1994) is the earliest work that formulates rigorously a theoretical foundation of this problem. More specifically, Theorem 6.1 of the paper immediately implies the following central limit result:

$$\sqrt{\Delta_n}\left(\mathrm{RV}_n - \int_0^T c_s \mathrm{d}s\right) \stackrel{\mathcal{L}-s}{\to} \sqrt{2}\int_0^T c_s \mathrm{d}B_s,$$

where $\stackrel{\mathcal{L}-s}{\to}$ denotes stable convergence in law, and B_s is a Brownian motion defined on an extension of the original probability space. To conduct inference, the same theorem yields a consistent estimator of the asymptotic variance:

$$\frac{2}{3}\sum_{i=1}^{[T/\Delta_n]} \left(\Delta_i^n X\right)^4 \stackrel{p}{\to} 2\int_0^T c_s^2 \mathrm{d}s.$$

The realized volatility estimator, despite being simple, efficient, and tuning-free, is not robust to the microstructure noise, so that it can only be applied to relatively liquid stocks subsampled sparsely (say, every 5 min), at which frequencies the noise seems to disappear. Nonetheless, 5-min is a somewhat *ad hoc* choice not applicable for all assets with transaction records. Asserting the absence of noise at certain pre-specified frequency typically results in at least a substantial cost of efficiency, if not a loss of consistency. To test for the existence of the noise, Andersen *et al.* (2003) advocate the use of volatility signature plots; Aït-Sahalia and Xiu (2016) formalize this procedure and propose Hausman tests; Liu, Patton, and Sheppard (2015) suggest the use of a ranking method of Patton (2011).

Modeling the microstructure noise is the next step toward a solution to this problem. Motivated from the Roll (1984) model, the observed price can be decomposed as the efficient price plus noise. Earlier work along this line of research include Zhou (1996); Aït-Sahalia, Mykland, and Zhang (2005); Hansen and Lunde (2006); and Bandi and Russell (2008), which seek the optimal sampling frequency that trades off the bias and variance of the estimators. The first consistent estimator is proposed by Zhang, Mykland, and Aït-Sahalia (2005), which is robust to the i.i.d. white microstructure noise, despite a low $n^{1/6}$ convergence rate. Zhang (2006) later proposes an efficient multi-scale extension that has the optimal $n^{1/4}$ rate. Barndorff-Nielsen *et al.* (2008) propose a general class of realized kernel estimators that nest the two estimators above. Their kernel estimators, however, suffer from a border bias that requires an additional jittering step.

Jacod *et al.* (2009) and Jacod, Podolskij, and Vetter (2010) propose a pre-averaging estimator, that is effectively a bias-corrected realized volatility using returns that are averaged over a sequence of blocks. This new estimation strategy is simple, intuitive, and more importantly, sufficiently flexible that leads to estimators of other functionals of volatility. In particular, the pre-averaging approach has become the standard for estimating integrated quarticity, namely, $\int_0^T c_s^2 \mathrm{d}s$, which appears again in the asymptotic variance of the noisy case.

Aït-Sahalia, Mykland, and Zhang (2005) and Xiu (2010) propose a likelihood-based estimator, which is optimal under constant volatility. Gloter and Jacod (2001) prove the local asymptotic normality in this case, so that the minimal asymptotic variance achievable is indeed given by $8a\sigma^3/T$, where a is the standard error of the i.i.d. noise U. Reiß (2011) later on provides the minimax efficiency bound in the general stochastic volatility case.

The aforementioned papers effectively assume an i.i.d. microstructure noise when developing the theoretical properties of the proposed estimators. In practice, however, there is ample evidence of noise autocorrelations beyond the first lag, dating back to as early as Niederhoffer and Osborne (1966), and more recent work by Hasbrouck and Ho (1987) and Brogaard, Hendershott, and Riordan (2014). There are few exceptions in the literature that discuss general dependent noises. Jacod, Li, and Zheng (2017b) provide nonparametric estimators of the serial correlations and moments of the microstructure noise. Building on that, Jacod, Li, and Zheng (2017a) propose a pre-averaging estimator of volatility that is robust to dependent noises. Similarly, Varneskov (2016) proposes a flat-top realized kernel. Both estimators depend on three tuning parameters. By contrast, Da and Xiu (2017) develop an extension of the likelihood estimator, which is tuning-free barring order selection. Related work that discusses general noise processes also include Aït-Sahalia, Mykland, and Zhang (2005); Aït-Sahalia, Mykland, and Zhang (2011); Kalnina and Linton (2008); Bandi and Russell (2008); Hautsch and Podolskij (2013); Bibinger *et al.* (2015); and Li, Liu, and Xiu (2017).

Professor Jacod has developed many other refinements to the theory of volatility estimation, including the robustness to jumps, and robustness to irregular sampling schemes, etc. For reason of space, we do not discuss these important contributions here, but refer the readers to Jacod and Protter (2012) and Aït-Sahalia and Jacod (2014) for more details.

2 Efficient Estimation of Integrated Volatility Functionals

Besides estimating volatility, Jacod (1994) considered general estimators formed as the Riemann sum of nonlinear transformations of normalized returns, namely,

$$V_n(g) \equiv \Delta_n \sum_{i=1}^{[T/\Delta_n]} g\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}\right). \tag{2.1}$$

Heuristically, because of the approximate local Gaussianity of Brownian martingales, the normalized returns $\left(\Delta_i^n X/\sqrt{\Delta_n}\right)_{i\geq 1}$ form approximately an i.n.i.d. array with marginal distribution $\mathcal{N}\left(0, c_{(i-1)\Delta_n}\right)$. Under this heuristic, one can see that

$$V_n(g) \xrightarrow{\mathbb{P}} \int_0^T \rho(c_s; g)\mathrm{d}s, \tag{2.2}$$

where the mapping $c \mapsto \rho(c; g)$ is defined as the expectation of $g(\mathcal{N}(0, c))$.

The limiting variable in Equation (2.2) forms an *integrated volatility functional* for the nonlinear transform $c \mapsto \rho(c; g)$. One class of example is with $g(x) = |x|^p$ which corresponds to $\rho(c; g) = m_p c^{p/2}$ for some constant m_p. The resulting functionals are integrated volatility polynomials. Another example is considered by Todorov and Tauchen (2012) with $g(x) = \cos\left(\sqrt{2u}x\right)$. The associated estimator $V_n(g)$ is the real part of the empirical characteristic function of the high-frequency returns. Its limit is given by $\int_0^T \exp\left(-uc_s\right)\mathrm{d}s$, that is, the Laplace transform of the volatility path.

The estimator $V_n(g)$, however, puts implicit restrictions on the type of integrated volatility functionals that can be estimated. Indeed, the nonlinear transform needs to have the form $\rho(\cdot; g)$. A further question is whether estimators can be constructed for general integrated volatility functionals with the form

$$S(h) = \int_0^T h(c_s)\mathrm{d}s$$

for a large class of test functions $h(\cdot)$.

A natural approach is to consider a plug-in estimator by replacing the latent volatility path with a nonparametric estimator for it. One of such nonparametric estimator can be formed by "localizing" the RV estimator over a "short (i.e., asymptotically shrinking)" time window. More generally, one can consider kernel-based nonparametric estimators for the spot volatility process (Kristensen, 2010). With $\hat{c}_s$ denoting the estimator of c_s, the plug-in estimator is formed as

$$\hat{S}(h) = \int_0^T h(\hat{c}_s)\mathrm{d}s.$$

Under mild regularity conditions, $\hat{S}(h)$ is a consistent estimator for $S(h)$.

In order to make inference, we need central limit theorems for characterizing the asymptotic distributions of these estimators. This higher-order calculation depends crucially on assumptions on the smoothness of the volatility process. It is instructive to put some additional structures on the plug-in estimator. To fix idea, suppose that the nonparametric estimator $\hat{c}_s$ is piece-wise constant on local windows of length $k_n\Delta_n$. Then, $\hat{S}(h)$ can be written as a Riemann sum

$$\hat{S}(h) = k_n\Delta_n \sum_{i=0}^{[T/(k_n\Delta_n)]} h(\hat{c}_{i\Delta_n}). \tag{2.3}$$

By a second-order expansion,

$$\begin{aligned}\hat{S}(h) \approx & \underbrace{k_n\Delta_n \sum_{i=0}^{[T/k_n\Delta_n]} h(c_{i\Delta_n})}_{\text{Leading term}} \\ & \underbrace{+ k_n\Delta_n \sum_{i=0}^{[T/k_n\Delta_n]} \partial h(c_{i\Delta_n})(\hat{c}_{i\Delta_n} - c_{i\Delta_n})}_{\text{Central limit theorem term}} \\ & \underbrace{+ \frac{k_n\Delta_n}{2} \sum_{i=0}^{[T/k_n\Delta_n]} \partial^2 h(c_{i\Delta_n})(\hat{c}_{i\Delta_n} - c_{i\Delta_n})^2}_{\text{Nonlinearity bias term}}.\end{aligned}$$

Kristensen (2010) considers a setting in which the volatility process has smooth (at least differentiable) sample paths. When the volatility path is smooth, it can be recovered at a relatively fast rate of convergence via kernel-based estimators. By properly choosing the bandwidth parameter in the spot volatility estimation, the nonlinearity bias term can be made asymptotically negligible. A central limit theorem for $\hat{S}_n(h)$ is then driven by the sampling error in the spot volatility estimation through the second term on the right hand of the above display.

However, volatility paths in typical stochastic volatility models are generally not very smooth. For example, in Heston's model the volatility path is non-differentiable because it is driven by a Brownian motion. More generally, volatility can also "jump" in response to major news arrivals, rendering its paths discontinuous. These realistic features of stochastic volatility set a $n^{1/4}$ upper bound for the rate of convergence of the spot volatility estimator. As a result, the nonlinearity bias term is of (sharp) order $O_p(n^{-1/2})$ and, hence, is no longer negligible for deriving a central limit theorem.

This complication is addressed in several recent work. Jacod and Rosenbaum (2013) propose a bias-correction for the "raw" estimator as follows:

$$\hat{S}(h) - \frac{1}{k_n}\hat{B}_n(h), \tag{2.4}$$

where the correction term $\hat{B}_n(h)$ is given by

$$\hat{B}_n(h) = k_n\Delta_n \sum_{i=1}^{[T/k_n\Delta_n]} \partial^2 h\left(c_{(i-1)\Delta_n}\right)\hat{c}^2_{(i-1)\Delta_n}. \tag{2.5}$$

Jacod and Rosenbaum (2013) show that this correction effectively eliminates the nonlinearity bias and the resulted bias-corrected estimator admits a feasible central limit theorem.

Compared with the estimator (2.1), the plug-in estimator (2.4) can be used to make inference for a much larger class of test functions. More precisely, Jacod and Rosenbaum (2013) allow for a class of test functions that are $\mathcal{C}^3$ (i.e., three-time continuously differentiable) with polynomial growth. The latter condition means that the test function h is bounded by a polynomial function near zero and infinity. This condition is restrictive for empirical applications but it turns out to be a technical one that can be relaxed by using the spatial localization argument later developed in Li, Todorov, and Tauchen (2016a). Inference tools for integrated volatility functionals with general $\mathcal{C}^3$ test functions are now available in the arsenal of financial econometricians. Using these tools, Aït-Sahalia and Xiu (2015) propose principal component analysis of high-frequency data; Li, Todorov, and Tauchen (2016b) estimate volatility functional dependencies; and Kalnina and Xiu (2017) estimate the leverage effect. Whether the class of "admissible" functions can be further extended is an interesting open question.

Besides accommodating a larger class of test functions, an additional advantage of the bias-corrected plug-in estimator (2.4) is that it is semiparametrically efficient. In particular, its asymptotic variance is generally strictly smaller than that of Equation (2.2). Semiparametric efficiency is the relevant notion of efficiency in this type of settings because we are interested in estimating a finite-dimensional quantity (i.e., the integrated volatility functional), without making parametric assumptions on the data generating process. We refer the readers to Bickel *et al*. (1998) for general discussions on semiparametric efficiency. Roughly speaking, the semiparametric efficiency bound depicts the best estimation accuracy in the "worst-case" parametric submodel. In some specific settings, Clément, Delattre, and Gloter (2013) and Renault, Sarisoy, and Werker (2016) have established the semiparametric efficiency bound for estimating integrated volatility functionals, which are attained by the estimator (2.4).

As discussed above, the bias-corrected estimator (2.4) has both broad scope in applications and high statistical efficiency. That said, the bias correction can be difficult to calculate in multivariate applications because one needs to calculate a large number of second-order derivatives (see Equation (2.5)). In a recent work, Li, Liu and Xiu (2017) propose an alternative way of bias correction using multiscale jackknife. The idea is to form "raw" estimators using multiple local windows (i.e., k_n) for the spot volatility estimation. Then, by properly forming a linear combination of these raw estimators, the bias terms can be implicitly eliminated. This jackknife estimator also attains the semiparametric efficiency bound, but is typically much easier to implement in practice. In addition, Li, Liu and Xiu (2017) show that the jackknife estimator is able to correct biases resulted from volatility jumps and volatility-of-volatility, which are not explicitly accounted for by the bias-correction in Equation (2.4). As a result, the jackknife estimator is valid for a broader range of choices of bandwidth parameters, which shows a type of robustness. The jackknife estimator shall provide an easy-to-implement alternative for making inference of general integrated volatility functionals.

From the above discussion, we see that much progress on the inference for integrated volatility functionals has been made since the seminal work of Jacod (1994). This progress not only deepens our understanding about the statistics in volatility estimation, but also offers new opportunities for various econometric applications, to which we now turn.

3 Semiparametric Estimation Based on Integrated Moments

In this section, we discuss an estimation strategy based on integrated moment conditions. Since the seminal work of Hansen (1982), moment-based estimation has been *the* standard tool in various areas of applied econometrics. Nonetheless, the high-frequency econometrics literature has been considered "nonstandard." The fundamental reason is that the "population" quantities in the high-frequency infill asymptotic setting are sample paths of economic variables, instead of their joint distributions as in conventional statistical settings. Apparently, the lack of the concept of "moments" in the high-frequency setting restricts the use of GMM.

We argue that this conceptual gap may be reconciled by viewing quantities like the integrated volatility functional as a type of moment. In fact, as discussed in Li, Todorov, and Tauchen (2013), the integral $\int_0^T h(c_s)\mathrm{d}s$ is the moment of the function $h(\cdot)$ under the occupation measure induced by the process c over the time interval $[0, T]$. More precisely, the occupation measure $\mathbb{F}$ evaluated on a measurable set B is defined as the amount of time when the process c stay in the set B, that is,

$$\mathbb{F}(B) = \int_0^T 1\{c_s \in B\}\mathrm{d}s.$$

It is easy to see that

$$\int_0^T h(c_s)\mathrm{d}s = \int h(x)\mathbb{F}(\mathrm{d}x).$$

From this viewpoint, the integrated volatility functional is simply a moment under the occupation measure. As in moment-based estimation problems, integrated volatility functionals (or occupational moments more generally) can be used to estimate model parameters. In order to streamline ideas, we start with the simplest example that concerns the estimation of beta. We consider the following bivariate model of log asset returns:

$$\mathrm{d}Y_t = \beta_t \mathrm{d}X_t + \mathrm{d}Y'_t,$$

where $\mathrm{d}X_t$ denotes the diffusive market returns, $\mathrm{d}Y_t$ is the stock return, and $\mathrm{d}\tilde{Y}_t$ is an idiosyncratic component that is orthogonal to the market returns. The spot beta β_t can be identified from the covariance matrix process as

$$\beta_t = c_{XY,t}/c_{XX,t},$$

where $c_{XX,t}$ (respectively, $c_{XY,t}$) is the spot variance of X (respectively, covariance between X and Y), respectively. In empirical asset pricing, it is often implicitly assumed that the beta is constant over a certain sample period. This amounts to impose a parametric restriction on the spot covariance matrix c_t, namely, for some constant β,

$$c_{XY,t} - \beta c_{XX,t} = 0, \quad \text{for all } t \in [0, T]. \tag{3.1}$$

While this restriction is unlikely to hold over long samples, it may provide an adequate description of the relationship between asset returns when the sample span T is relatively short.

The model restriction (3.1) apparently concerns stochastic processes. However, we stress that it can also be understood as an instantaneous conditional moment restriction. Indeed, the (unobserved) spot covariance matrix c_t is exactly the instantaneous second moment of the return vector (dY_t, dX_t). Analogously to classic moment-based estimation problems, we can "instrument" Equation (3.1) so as to obtain unconditional (occupational) moment conditions. That is, for any process w_t, Equation (3.1) implies the following integrated moment condition:

$$\int_0^T \left(c_{XY,s} - \beta c_{XX,s}\right) w_s \mathrm{d}s = 0, \tag{3.2}$$

which yields

$$\beta = \frac{\int_0^T c_{XY,s}\ w_s \mathrm{d}s}{\int_0^T c_{XX,s}\ w_s \mathrm{d}s}. \tag{3.3}$$

Of course, if the model (3.1) is indeed correctly specified, Equation (3.3) defined with different weights should all coincide with the true beta. However, under misspecification, the expression in Equation (3.3) should be interpreted as a pseudo-true parameter. Regardless of whether the model is correctly specified or not, the (pseudo) true parameter is a transformation of integrated volatility functional. Hence, the limit theory described in the previous section can be used to make inference for it.

The above framework follows closely the econometric tradition in the GMM literature. An immediate payoff of this conceptualization is that it offers some econometric discipline for thinking about the notion of "integrated beta." For example, if the weight function is $w_t = 1$, then the pseudo-true beta parameter is

$$\beta(1) = \frac{\int_0^T c_{XY,s}\ \mathrm{d}s}{\int_0^T c_{XX,s}\ \mathrm{d}s}, \tag{3.4}$$

which is the integrated beta proposed by Barndorff-Nielsen and Shephard (2004). On the other hand, if we set the weight $w_t = c_{XX,t}^{-1}$, then Equation (3.3) becomes

$$\beta\left(c_{XX}^{-1}\right) = \frac{1}{T}\int_0^T \beta_s \mathrm{d}s. \tag{3.5}$$

The two versions of "integrated betas" defined in Equations (3.4) and (3.5) both have some intuitive appeal. Econometrically, they can both be interpreted as pseudo-true parameters in the constant beta model, but associated with different instruments. Generally speaking, the estimators associated with these pseudo-true parameters are not directly comparable because they estimate different quantities. That said, it is meaningful to compare their statistical efficiency if the constant beta model indeed holds. In that case, it turns out that the above *ad hoc* choices of weights are not efficient. As shown in Li, Todorov, and Tauchen (2016a), the optimal weight for estimating the constant beta is

$$w_t = \left(c_{YY,t} - c^2_{XY,t}/c_{XX,t}\right)^{-1},$$

that is, the inverse of the spot idiosyncratic variance of asset Y. This is reminiscent of Robinson's efficient estimation under unknown form of heteroskedasticity (Robinson, 1987). The resulting estimator is not only semiparametrically efficient, but actually adaptive with respect to the nonparametric nuisance process c_{XX}.

The above discussion highlights the importance of knowing whether certain parametric assumption holds or not, because such knowledge dictates how to interpret estimators and how to construct efficient estimators that fully exploit the information content underlying the model restrictions. In classic moment-based estimation problems, specification test for conditional moment equalities can be carried out by examining whether a continuum of unconditional moment conditions hold or not (Bierens, 1982). The same idea can be implemented in the high-frequency setting. Indeed, it can be shown that the constant beta restriction (3.1) holds if and only if

$$\int_0^T \left(c_{XY,s} - \beta c_{XX,s}\right) w(\tau s) \mathrm{d}s = 0,$$

for all τ in a bounded interval with positive length for some weight function $w(\cdot)$ that is properly chosen [e.g., $w(\tau s) = \cos(\tau s) + \sin(\tau s)$]. This amounts to making inference for a continuum of integrated functionals associated with the class of functions

$$h_{\beta,\tau}(c) = (c_{XY} - \beta c_{XX}) w(\tau s).$$

Li, Todorov, and Tauchen (2016b) develop an empirical process theory that extends the finite-dimensional central limit theory of Jacod and Rosenbaum (2013) and Li, Todorov, and Tauchen (2016a) so as to address this functional inference problem. The above strategy for specification testing is clearly not restricted for testing constant beta, but can be applied to test any parametric restrictions on a multivariate covariance process; see Li, Todorov, and Tauchen (2016b) for the general discussion.

So far, we have considered applications that concern the spot covariance matrix process itself. However, economic theory often have implications on the relationship between price volatility and other economic variables. One example is the classic study of the volume–volatility relationship such as the mixed distribution hypothesis (MDH); see Clark (1973) and Tauchen and Pitts (1983). Under the MDH, both the volume and return variance are driven by a latent information flow, which results in a linear volume–variance relationship. Another example is from option pricing, in which a pricing theory depicts a mapping from state variables (such as stock price and volatility) to option prices.

These applications can be cast into an econometric model with the form

$$Y_{i\Delta_n} = f(Z_{i\Delta_n}, c_{i\Delta_n}; \theta) + \epsilon_{i\Delta_n},$$

where $Y_{i\Delta_n}$ denotes an economic variable such as volume or option price, the function f depicts a relationship implied by the economic theory which is known up to the parameter θ, Z is some state variable, and ϵ is a disturbance term (e.g., option pricing error or noise trading volume). The econometric inference is about making inference for θ. Like in the

GMM, it is natural to construct estimators based on instrumented sample moments. With an instrument of the form $w(Z_{i\Delta_n}, c_{i\Delta_n}; \theta)$ for some smooth function $w(\cdot)$, the "raw" sample moment condition is given by

$$G_n(\theta) = \Delta_n \sum_{i=0}^{[T/\Delta_n]} (Y_{i\Delta_n} - f(Z_{i\Delta_n}, \hat{c}_{i\Delta_n}; \theta)) w(Z_{i\Delta_n}, \hat{c}_{i\Delta_n}; \theta),$$

where we have replaced the latent spot covariance matrix process c with its nonparametric estimate. This sample moment condition is more complicated than the estimator (2.3) because it also involves the variables $Y_{i\Delta_n}$ and $Z_{i\Delta_n}$, where the former is modeled as a noisy semimartingale. Moreover, for the same reason discussed in Section 2, we also need to correct a nonlinearity bias that arises from the estimation of the spot covariance matrix by using the bias-corrected moment condition given by

$$G_n^* \equiv G_n(\theta) - \frac{\Delta_n}{k_n} \sum_{i=0}^{[T/\Delta_n]} \partial_c^2 [(Y_{i\Delta_n} - f(Z_{i\Delta_n}, \hat{c}_{i\Delta_n}; \theta)) w(Z_{i\Delta_n}, \hat{c}_{i\Delta_n}; \theta)] \hat{c}_{i\Delta_n}^2.$$

Based on the bias-corrected moment condition, the parameter θ can then be estimated like in the GMM via the following minimization:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmin}} \; G_n^*(\theta)^\top W_n G_n^*(\theta).$$

In Li and Xiu (2016), we propose the above estimator for θ and showed that $\hat{\theta}_n$ is asymptotically mixed Gaussian. We also propose a Bierens-type specification test in the spirit of the constant-beta test discussed above.

4 Concluding Remarks

The important work of Jacod (1994) established the inference theory for a large class of integrated volatility functionals and opened the door for a lot interesting research. Much progress has been made in the past 20 years. As discussed above, integrated volatility functionals is not only useful for measuring various aspect of volatility risk, but can also serve as integrated (or occupational) moment conditions for estimating economic models and testing economic theory. Empirical applications using these novel tools should shed new light on classic problems like the volume–volatility relationship, option pricing, or perhaps more excitingly, new empirical questions about which high-frequency data are informative. Besides the many opportunities for new empirical research, much theoretical work remains to be done. This includes the efficient estimation of integrated volatility functionals for noisy high-frequency data, high-dimensional factor models, semiparametrically efficient estimation in empirical finance models, and bootstrap inference for such procedures, to name but a few.

References

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

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang. 2005. How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise. *Review of Financial Studies* 18: 351–416.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang. 2011. Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise. *Journal of Econometrics* 160: 160–175.

Aït-Sahalia, Y., and D. Xiu. 2015. Principal Component Analysis of High Frequency Data. Technical Report, Princeton University and the University of Chicago.

Aït-Sahalia, Y., and D. Xiu. 2016. A Hausman Test for the Presence of Market Microstructure Noise in High Frequency Data. Technical Report, Princeton University and the University of Chicago.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2003. Modeling and Forecasting Realized Volatility. *Econometrica* 71: 579–625.

Bandi, F. M., and J. R. Russell. 2008. Microstructure Noise, Realized Volatility and Optimal Sampling. *Review of Economic Studies* 75: 339–369.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard. 2008. Designing Realized Kernels to Measure Ex-Post Variation of Equity Prices in the Presence of Noise. *Econometrica* 76: 1481–1536.

Barndorff-Nielsen, O. E., and N. Shephard. 2004. Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics. *Econometrica* 72: 885–925.

Bibinger, M., N. Hautsch, P. Malec, and M. Reiß. 2015. Estimating the Spot Covariation of Asset Prices—Statistical Theory and Empirical Evidence. Technical Report, University of Mannheim.

Bickel, P. J., C. A. J. Klaassen, Y. Ritov, and J. A. Wellner. 1998. *Efficient and Adaptive Estimation for Semiparametric Models*. New York: Springer-Verlag.

Bierens, H. 1982. Consistent Model Specification Tests. *Journal of Econometrics* 20: 105–134.

Brogaard, J. A., T. Hendershott, and R. Riordan. 2014. High Frequency Trading and Price Discovery. *Review of Financial Studies* 27: 2267–2306.

Clark, P. K. 1973. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41: 135–155.

Clément, E., S. Delattre, and A. Gloter. 2013. An Infinite Dimensional Convolution Theorem with Applications to the Efficient Estimation of the Integrated Volatility. *Stochastic Processes and their Applications* 123: 2500–2521.

Da, R., and D. Xiu. 2017. When Moving-Average Models Meet High-Frequency Data: Uniform Inference on Volatility. Technical Report, University of Chicago.

Gloter, A., and J. Jacod. 2001. Diffusions with Measurement Errors: I. Local Asymptotic Normality. *ESAIM: Probability and Statistics* 5: 225–242.

Hansen, L. P. 1982. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* 50: 1029–1054.

Hansen, P. R., and A. Lunde. 2006. Realized Variance and Market Microstructure Noise. *Journal of Business and Economic Statistics* 24: 127–161.

Hasbrouck, J., and T. S. Y. Ho. 1987. Order Arrival, Quote Behavior, and the Return-Generating Process. *Journal of Finance* 42: 1035–1048.

Hautsch, N., and M. Podolskij. 2013. Pre-averaging Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence. *Journal of Business & Economic Statistics* 31: 165–183.

Jacod, J. 1994. Limit of Random Measures Associated with the Increments off a Brownian Semimartingale. *Journal of Financial Econometrics*, forthcoming.

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: 2249–2276.

Jacod, J., Y. Li, and X. Zheng. 2017a. Estimating the Integrated Volatility with Tick Observations. Technical Report, Universite Paris VI.

Jacod, J., Y. Li, and X. Zheng. 2017b. Statistical Properties of Microstructure Noise. *Econometrica*, 85: 1133–1174.

Jacod, J., M. Podolskij, and M. Vetter. 2010. Limit Theorems for Moving Averages of Discretized Processes Plus Noise. *Annals of Statistics* 38: 1478–1545.

Jacod, J., and P. Protter. 2012. *Discretization of Processes*. Berlin Heidelberg: Springer.

Jacod, J., and M. Rosenbaum. 2013. Quarticity and Other Functionals of Volatility: Efficient Estimation. *Annals of Statistics* 41: 1462–1484.

Jacod, J., and A. N. Shiryaev. 2003. *Limit Theorems for Stochastic Processes*. 2nd edn. Berlin Heidelberg: Springer-Verlag.

Kalnina, I., and O. Linton. 2008. Estimating Quadratic Variation Consistently in the Presence of Endogenous and Diurnal Measurement Error. *Journal of Econometrics* 147: 47–59.

Kalnina, I., and D. Xiu. 2017. Nonparametric Estimation of the Leverage Effect: A Trade-off between Robustness and Efficiency. *Journal of American Statistical Association* 112: 384–396.

Kristensen, D. 2010. Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach. *Econometric Theory* 26: 60–93.

Li, J., Y. Liu, and D. Xiu. 2017. Efficient Estimation of Integrated Volatility Functionals via Multiscale Jackknife. Technical Report, Duke University and University of Chicago.

Li, J., V. Todorov, and G. Tauchen. 2013. Volatility Occupation Times. Annals of Statistics, 41: 1865–1891.

Li, J., V. Todorov, and G. Tauchen. 2016a. Adaptive Estimation of Continuous-Time Regression Models using High-Frequency Data. *Journal of Econometrics*, 200: 36–47.

Li, J., V. Todorov, and G. Tauchen. 2016b. Inference Theory on Volatility Functional Dependencies. *Journal of Econometrics* 193: 17–34.

Li, J., and D. Xiu. 2016. Generalized Method of Integrated Moments for High-Frequency Data. *Econometrica* 84: 1613–1633.

Li, Z. M., R. J. Laeven, and M. H. Vellekoop. 2017. Dependent Microstructure Noise and Integrated Volatility Estimation from High-Frequency Data. Technical Report, University of Amsterdam.

Liu, L. Y., A. J. Patton, and K. Sheppard. 2015. Does Anything Beat 5-Minute RV? A Comparison of Realized Measures across Multiple Asset Classes. *Journal of Econometrics* 187: 293–311.

Niederhoffer, V., and M. Osborne. 1966. Market Making and Reversal on the Stock Exchange. *Journal of American Statistical Association* 61: 897–916.

Patton, A. J. 2011. Data-Based Ranking of Realised Volatility Estimators. *Journal of Econometrics* 161: 284–303.

Reiß, M. 2011. Asymptotic Equivalence for Inference on the Volatility from Noisy Observations. *Annals of Statistics* 39: 772–802.

Renault, E., C. Sarisoy, and B. J. Werker. 2016. Efficient Estimation of Integrated Volatility and Related Processes. *Econometric Theory* 33: 439–478.

Robinson, P. M. 1987. Asymptotically Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form. *Econometrica* 55: 875–891.

Roll, R. 1984. A Simple Model of the Implicit Bid–Ask Spread in an Efficient Market. *The Journal of Finance* 39: 1127–1139.

Tauchen, G., and M. Pitts. 1983. The Price Variability–Volume Relationship on Speculative Markets. *Econometrica* 51: 485–505.

Todorov, V., and G. Tauchen. 2012. The Realized Laplace Transform of Volatility. *Econometrica* 80: 1105–1127.

Varneskov, R. T. 2016. Estimating the Quadratic Variation Spectrum of Noisy Asset Prices using Generalized Flat-Top Realized Kernels. *Journal of Business & Economic Statistics*, 34: 1–22.

Xiu, D. 2010. Quasi-Maximum Likelihood Estimation of Volatility with High Frequency Data. *Journal of Econometrics* 159: 235–250.

Zhang, L. 2006. Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-scale Approach. *Bernoulli* 12: 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: 1394–1411.

Zhou, B. 1996. High-Frequency Data and Volatility in Foreign-Exchange Rates. *Journal of Business & Economic Statistics* 14: 45–52.