

Semi-Parametric Comparison of Stochastic Volatility Models using Realized Measures

VALENTINA CORRADI
Queen Mary, University of London

and

WALTER DISTASO
Imperial College London

First version received January 2004; final version accepted November 2005 (Eds.)

This paper proposes a procedure to test for the correct specification of the functional form of the volatility process within the class of eigenfunction stochastic volatility models. The procedure is based on the comparison of the moments of realized volatility measures with the corresponding ones of integrated volatility implied by the model under the null hypothesis.

We first provide primitive conditions on the measurement error associated with the realized measure, which allow to construct asymptotically valid specification tests.

Then we establish regularity conditions under which the considered realized measures, namely, realized volatility, bipower variation, and modified subsampled realized volatility, satisfy the given primitive assumptions.

Finally, we provide an empirical illustration based on three stocks from the Dow Jones Industrial Average.

1. INTRODUCTION

Estimation and testing of financial volatility models has received increased attention over the recent years, from both a theoretical and an empirical perspective. In fact, accurate specification of volatility is of crucial importance in financial risk management and in pricing and hedging derivative securities.

Realistic models describing the dynamics of asset prices are made of a continuous stochastic volatility component plus a jump component (see, for example, Andersen, Benzoni and Lund, 2002). These models are able to reproduce the main features of the observed implied volatility surfaces. Also, it is generally believed that the observed transaction prices can be decomposed into the unobservable efficient price plus an error component, called market microstructure noise, which can be ascribed to imperfections in the underlying trading mechanism (*e.g.* liquidity effects, bid–ask spreads, and informational asymmetries). Being able to disentangle and to separate the diffusive continuous part of the process from the jump component and from market microstructure noise is important in a number of different situations. First, jumps in the underlying price process are important ingredients for correctly pricing and hedging contingent claims and play a central role in financial risk management, which is typically based on some tail statistics. Second, if one is interested in studying the volatility of the efficient log-price process making use of the increased availability of high-frequency observations on transaction prices, then particular care should be devoted to handling the effect of microstructure noise.

In fact, the use of observations at the highest available frequency would produce an estimate of the variance of the microstructure noise, rather than of the object of interest, quadratic variation. Recent papers dealing with the separation of the Brownian component from the jump part include Aït-Sahalia (2004), Andersen, Bollerslev and Diebold (2005a), Bollerslev, Kretschmer, Pigorsch and Tauchen (2005), Huang and Tauchen (2005), and Tauchen and Zhou (2005), while the issue of the effect of microstructure noise on estimators of volatility has been analysed by Aït-Sahalia, Mykland and Zhang (2005a,b), Bandi and Russell (2006), and Zhang, Mykland and Aït-Sahalia (2005).

In the literature, several specification tests for stochastic volatility models have been proposed, based on different estimation methods. This is the case for various tests for the validity of overidentifying restrictions based on simulated generalized method of moment (GMM) estimators (e.g. Duffie and Singleton, 1993), indirect inference (Gouriéroux, Monfort and Renault, 1993), and efficient method of moments (see, for example, Gallant and Tauchen, 1996; Gallant, Hsieh and Tauchen, 1997; Chernov, Gallant, Ghysels and Tauchen, 2003). Recently, Fermanian and Salanié (2004) have suggested a simulated maximum likelihood estimator obtained by using a kernel density estimator of the simulated data; Altissimo and Mele (2005) have proposed a new estimator based on the minimization of the weighted distance between a kernel estimator of the actual data and of the simulated data. Specification tests based on these two estimators can be constructed accordingly. All the papers cited above require a complete specification of the drift and diffusion terms, as well as of the jump component. They do not require a characterization of the microstructure noise component; this is because the limiting distributions of the suggested test statistics require the time span to go to infinity, keeping the interval between successive observations fixed. Also, given the often substantial intra-day price movements, methods based on low frequency observations may well fail to exploit all the informational content of the data. Conversely, other specification tests for the functional form of the diffusion term of the process, such as Corradi and White (1999), Dette and von Lieres und Wilkau (2003), and Dette, Podolskij and Vetter (2004), based on infill asymptotics, require the interval between successive observations to shrink to 0 and so are implemented using high frequency observations. In these papers, the specification of the drift term is not required, since the drift term cannot be identified using observations over a fixed time span, but both the jump component and the microstructure error have to be fully specified.¹

The main objective of this paper is to provide semi-parametric statistical inference for stochastic volatility models. In fact, the paper provides specification tests for the functional form of the volatility process, which focus only on the diffusive part of the process and do not require a parametric specification of either the jump component or the microstructure noise. In other words, the main advantage of the testing procedures outlined below is that one can be (almost) completely silent about jumps and microstructure noise.

Given a specific stochastic volatility model, within the class of the eigenfunction stochastic volatility models of Meddahi (2001), which nests all the most popular stochastic volatility models as special cases, the first two moments and the covariance structure of the integrated, daily volatility process are derived.² Such moments can be either obtained in closed form or approximated by the sample moments of the simulated daily volatility process. Then our testing procedure is based on comparing a set of moments implied by a given model with the corresponding set of sample moments of daily integrated volatility. Of course, since the “true” integrated volatility process cannot be observed, in implementing our test we have to use the moments of

1. Indeed, in all three papers it has been assumed that there are no jumps and no microstructure noise.

2. If we denote σ_s^2 the instantaneous volatility at time s , then daily integrated volatility at day t , is defined as $\int_{t-1}^t \sigma_s^2 ds$.

some model-free estimator of quadratic variation. We call such an estimator, constructed using intra-daily observations, a realized measure. More precisely, we base our test on a series of T realized measures, where T denotes the number of days for which observations are available, and we use M intra-daily observations to construct each (daily) realized measure. Therefore, our set of moments, which is based on an unobservable variable (integrated volatility), is subject to a measurement error, since it is constructed using a realized measure. In the paper, we provide a set of sufficient conditions on the measurement error under which the test statistic constructed using sample moments of the unobservable integrated volatility and the corresponding one constructed using sample moments of the realized measure are asymptotically equivalent. These sufficient conditions impose some mild requirements on the moments and memory of the measurement error and on the relative rate of growth of M and T .

We adapt the sufficient conditions for the asymptotic negligibility of the measurement error to three popular realized measures, realized volatility, bipower variation, and subsampled realized volatility. Realized volatility has been introduced concurrently by Andersen, Bollerslev, Diebold and Labys (2001, 2003) and by Barndorff-Nielsen and Shephard (2001, 2002); the relevant limit theory and extensions to the multi-dimensional case are analysed by Jacod (1994), Jacod and Protter (1998), and Barndorff-Nielsen and Shephard (2004a). Assuming that we have M recorded intra-day observations for a given asset-price process, over a given day, realized volatility is computed by summing up the M -squared returns. If prices have continuous paths and are not contaminated by microstructure noise, then realized volatility is a consistent estimator of daily integrated volatility. Bipower variation, introduced by Barndorff-Nielsen and Shephard (2004b) and analysed by Barndorff-Nielsen, Graversen, Jacod, Podolskij and Shephard (2006) (see also Barndorff-Nielsen, Graversen, Jacod and Shephard, 2005a) is computed by summing up the products of absolute values of adjacent returns. Bipower variation retains consistency for integrated volatility when jumps are present in the log-price process. Finally, modified subsampled realized volatility proposed by Zhang *et al.* (2005) is a consistent estimator of integrated volatility when prices are contaminated by microstructure noise (see also related contributions by Zhang, 2004; Aït-Sahalia *et al.*, 2005b; Barndorff-Nielsen, Hansen, Lunde and Shephard, 2005b).

The idea of using moment conditions for estimating and testing stochastic volatility models using realized measures is not new. In fact, Bollerslev and Zhou (2002) have derived analytically the first two conditional moments of the latent volatility process for the class of affine stochastic volatility models. They then suggested a GMM estimator and an associated test for the validity of overidentifying restrictions based on the comparison between the analytical conditional moments of integrated volatility and the corresponding sample moments of realized volatility. Bollerslev and Zhou consider the case of the time span T approaching infinity, for a given number of intra-day observations M . The effects of various values of M on the properties of the test are analysed via a Monte Carlo simulation.

The present paper extends Bollerslev and Zhou's in four directions. First, we explicitly analyse our moment conditions in a measurement error context and provide sufficient conditions under which the latter can be asymptotically ignored. Second, we consider a double asymptotic theory in which both T and M approach infinity, and we provide regularity conditions on their relative rate of growth. Third, we also consider tests comparing (simulated) moments of integrated volatility with sample moments of bipower variation, thus allowing for possible jumps, and with sample moments of modified subsampled realized volatility, thus allowing for at least some classes of microstructure noise. Finally, we do not confine our attention to affine stochastic volatility models, but we consider the class of eigenfunction stochastic volatility models of Meddahi (2001), where the latent volatility process is modelled as a linear combination of the eigenfunctions associated with the infinitesimal generator of the diffusion driving the volatility process.

The main reason why we focus on Meddahi's eigenfunction stochastic volatility class is that it ensures that the integrated volatility process has a memory decaying at a geometric rate and has an $\text{ARMA}(p, p)$ structure when the number of eigenfunctions, p , is finite (see Barndorff-Nielsen and Shephard, 2001, 2002; Andersen, Bollerslev and Meddahi, 2004, 2005b) and that the measurement error associated with the realized measures has a memory decaying at a fast enough rate. These features are crucial, as in our context both T and M approach infinity.

Also, the limit theorems of Jacod (1994), Jacod and Protter (1998), Barndorff-Nielsen *et al.* (2006), and Zhang *et al.* (2005) for the measurement error associated with all the three realized measures, which hold for a very general class of semi-martingale processes, are derived in the fixed time span case. Therefore, in that context there is no need to impose restrictions on the degree of memory of the volatility process. Of course, if one wishes to construct a testing procedure based on a finite time span, there is no need to consider a specific class of models, and then he or she can benefit from the generality of the results above.

The proposed testing procedure is applied to the three most liquid stocks included in the Dow Jones Industrial Average, namely General Electric, Intel, and Microsoft. In particular, we conduct a test on the popular square root model, proposed by Heston (1993). The results of the empirical analysis show that Heston's model is a reasonably good candidate to explain the behaviour of stochastic volatility and, interestingly, highlight some different outcomes of the tests, according to the specific realized measure used.

This paper is organized as follows. Section 2 describes the set-up. Section 3 provides primitive conditions on the measurement error, in terms of its first two moments and autocorrelation structure, which allow to construct tests for overidentifying restrictions, based on the comparison between sample moments of the realized measure and analytical moments of integrated volatility, when the latter are known in closed form. In particular, we provide conditions on the rate at which the time span can approach infinity, in relation to the rate at which the moments of measurement error approach 0. Section 4 considers the case in which there is no explicit closed form for the moments of integrated volatility. For this case, we propose a simulated version of the test based on the comparison of the sample moments of realized measures and sample moments of the simulated integrated volatility process. We also discuss the possibility of constructing a test based on the comparison of sample moments of actual and simulated realized measure for fixed M . Section 5 provides conditions under which realized volatility, bipower variation, and modified subsampled realized volatility satisfy the primitive conditions on the measurement error. In particular, it is emphasized that the rate at which T can grow, relatively to M , differs across the three realized measures. Section 6 provides an empirical illustration of the suggested procedure, based on data on different stocks of the Dow Jones Industrial Average. Finally, Section 7 concludes. Technical lemmas and proofs are gathered in the Appendix.

2. THE MODEL

In this section, we will briefly introduce the class of eigenfunction stochastic volatility models and describe the realized measures used to derive the moment conditions. Let the observable state variable, $Y_t = \log S_t$, where S_t denotes the price of a financial asset or the exchange rate between two currencies, be modelled as a jump diffusion process with a constant drift term. According to the eigenfunction stochastic volatility class, the variance term is modelled as a measurable function of a latent factor, f_t , which is also generated by a diffusion process. Thus,

$$dY_t = mdt + dz_t + \sqrt{\sigma_t^2} \left(\sqrt{1 - \rho^2} dW_{1,t} + \rho dW_{2,t} \right) \quad (1)$$

and

$$\sigma_t^2 = \psi(f_t) = \sum_{i=0}^p a_i P_i(f_t) \quad (2)$$

$$df_t = \mu(f_t, \theta)dt + \sigma(f_t, \theta)dW_{2,t}, \quad (3)$$

for some $\theta \in \Theta \in \mathbb{R}^{2p+1}$, where $W_{1,t}$ and $W_{2,t}$ refer to two independent Brownian motions, $\rho \in (-1, 1)$ and $P_i(f_t)$ denotes the i -th eigenfunction of the infinitesimal generator \mathcal{A} associated with the unobservable state variable f_t .³

Here, equation (1) describes the behaviour of the log-price process. Following the discussion given in the Introduction, we model the log-price process as the sum of a diffuse component characterized by a constant drift term and the variance term and a jump component. Notice that we specify explicitly the possibility of a leverage effect through the inclusion of the parameter ρ . According to the eigenfunction stochastic volatility class of models, the instantaneous variance, σ_t^2 , is expressed in (2) as a linear combination of the first p eigenfunctions associated with the stationary diffusion process f_t , characterized by (3).

The pure jump process dz_t specified in (1) is such that

$$Y_t = mt + \int_0^t \sqrt{\sigma_s^2} \left(\sqrt{1 - \rho^2} dW_{1,s} + \rho dW_{2,s} \right) + \sum_{i=1}^{J_t} c_i,$$

where J_t is a finite activity counting process, and c_i is a non-zero i.i.d. random variable, independent of J_t . As J_t is a finite activity counting process, we confine our attention to models characterized by a finite number of jumps over any fixed time span.

As customary in the literature on stochastic volatility models, the volatility process is assumed to be driven by (a function of) the unobservable state variable f_t . Rather than assuming an ad hoc function for $\psi(\cdot)$, the eigenfunction stochastic volatility model adopts a more flexible approach. In fact $\psi(\cdot)$ is modelled as a linear combination of the eigenfunctions of \mathcal{A} associated with f_t . Notice that the a_i 's are real numbers and that p may be infinite. Also, for normalization purposes, it is further assumed that $P_0(f_t) = 1$ and that $\text{var}(P_i(f_t)) = 1$, for any $i \neq 0$. When p is infinite, we also require that $\sum_{i=0}^{\infty} a_i^2 < \infty$. The generality and embedding nature of the approach just outlined stems from the fact that any square integrable function $\psi(f_t)$ can be written as a linear combination of the eigenfunctions associated with the state variable f_t . As a result, most of the widely used stochastic volatility models can be derived as special cases of the general eigenfunction stochastic volatility model. For more details on the properties of these models, see Meddahi (2001) and Andersen *et al.* (2004).

Finally, notice that we have assumed a constant drift term. This is in line with Bollerslev and Zhou (2002), who assume a zero drift term and justify this with the fact that there is very little predictive variation in the mean of high frequency returns, as supported by the empirical findings of Andersen and Bollerslev (1997).

3. The infinitesimal generator \mathcal{A} associated with f_t is defined by

$$\mathcal{A}\phi(f_t) \equiv \mu(f_t)\phi'(f_t) + \frac{\sigma^2(f_t)}{2}\phi''(f_t)$$

for any square integrable and twice differentiable function $\phi(\cdot)$. The corresponding eigenfunctions $P_i(f_t)$ and eigenvalues $-\lambda_i$ are given by $\mathcal{A}P_i(f_t) = -\lambda_i P_i(f_t)$. For a detailed discussion and analysis on infinitesimal generators and spectral decompositions, see Aït-Sahalia, Hansen and Scheinkman (2004).

Following the widespread consensus that transaction data occurring in financial markets are often contaminated by measurement errors, we assume to have a total of MT observations, consisting of M intra-daily observations for T days, for

$$X_{t+j/M} = Y_{t+j/M} + \epsilon_{t+j/M}, \quad t = 1, \dots, T \quad \text{and} \quad j = 1, \dots, M,$$

where

$$\epsilon_{t+j/M} \sim \text{i.i.d. } (0, \nu) \quad \text{and} \quad E(\epsilon_{t+j/M} Y_{s+i/M}) = 0 \quad \text{for all } t, s, j, i. \quad (4)$$

Thus, we allow for the possibility that the observed transaction price can be decomposed into the efficient one plus a “noise” due to measurement error, which captures generic microstructure effects.

The microstructure noise is assumed to be identically and independently distributed and to be independent of the underlying prices. This is consistent with the model considered by Bandi and Russell (2003, 2006), Aït-Sahalia *et al.* (2005a), and Zhang *et al.* (2005).⁴ Needless to say, when $\nu = 0$ then $\epsilon_{t+j/M} = 0$ (almost surely, a.s.) and therefore $X_{t+j/M} = Y_{t+j/M}$ (a.s.).

The daily integrated volatility process at day t is defined as

$$IV_t = \int_{t-1}^t \sigma_s^2 ds, \quad (5)$$

where σ_s^2 denotes the instantaneous volatility at time s . Proposition 4.1 in Andersen *et al.* (2004) gives the complete moment structure of integrated volatility

$$\begin{aligned} E(IV_t(\boldsymbol{\theta})) &= a_0 \\ \text{var}(IV_t(\boldsymbol{\theta})) &= 2 \sum_{i=1}^p \frac{a_i^2}{\lambda_i^2} (\exp(-\lambda_i) + \lambda_i - 1) \\ \text{cov}(IV_t(\boldsymbol{\theta}), IV_{t-k}(\boldsymbol{\theta})) &= \sum_{i=1}^p a_i^2 \exp(-\lambda_i(k-1)) \frac{(1 - \exp(-\lambda_i))^2}{\lambda_i^2}. \end{aligned} \quad (6)$$

This set of moments provides the basis for the testing procedure derived in the following sections. In particular, since IV_t is not observable, different realized measures, based on the sample $X_{t+j/M}$, $t = 1, \dots, T$ and $j = 1, \dots, M$, are used as proxies for it. The realized measure, say $RM_{t,M}$, is a noisy measure of the true integrated volatility process; in fact

$$RM_{t,M} = IV_t + N_{t,M},$$

where $N_{t,M}$ denotes the measurement error associated with the realized measure $RM_{t,M}$. Note that, in the case where $\nu > 0$, any realized measure of integrated volatility is contaminated by two sources of measurement errors, given that it is constructed using contaminated data.

Our objective is to compare the moment structure of the chosen realized measure $RM_{t,M}$ with that of IV_t given in (6). Note that when $p = 1$, $\text{cov}(IV_t(\boldsymbol{\theta}), IV_{t-k_1}(\boldsymbol{\theta}))/\text{cov}(IV_t(\boldsymbol{\theta}), IV_{t-k_2}(\boldsymbol{\theta})) = \exp(-\lambda_1(k_1 - k_2))$, so that, by using mean, variance, and two autocovariances of $IV_t(\boldsymbol{\theta})$, we obtain one overidentifying restriction. Analogously, when $p = 2$, we shall be using

4. Recently, Hansen and Lunde (2005) have addressed the issue of time dependence in the microstructure noise, while Awartani, Corradi and Distaso (2005) allow for correlation between the underlying price and the market microstructure noise.

four autocovariances, as well as mean and variance, in such a way to obtain one overidentifying restriction.⁵ In order to test the correct specification of a given eigenfunction volatility model, we impose the particular parametrization implied by the model under the null hypothesis.

In the sequel, we will first provide primitive conditions on the measurement error $N_{t,M}$, in terms of its moments and memory structure, for the asymptotic validity of tests based on the comparison of the moments of $RM_{t,M}$ with those of IV_t . Then, we shall adapt the given primitive conditions on $N_{t,M}$ to the three considered realized measures of integrated volatility, namely,

(a) realized volatility, defined as

$$RV_{t,M} = \sum_{j=1}^{M-1} (X_{t+(j+1)/M} - X_{t+j/M})^2, \quad (7)$$

(b) normalized bipower variation, defined as

$$(\mu_1)^{-2} BV_{t,M} = (\mu_1)^{-2} \frac{M}{M-1} \sum_{j=2}^{M-1} |X_{t+(j+1)/M} - X_{t+j/M}| |X_{t+j/M} - X_{t+(j-1)/M}| \quad (8)$$

where $\mu_1 = E|Z| = 2^{1/2} \Gamma(1) / \Gamma(1/2)$ and Z is a standard normal distribution,

(c) modified subsampled realized volatility, defined as

$$\widehat{RV}_{t,l,M}^u = RV_{t,l,M}^{\text{avg}} - 2l\widehat{v}_{t,M}, \quad (9)$$

where

$$\begin{aligned} \widehat{v}_{t,M} &= \frac{RV_{t,M}}{2M} = \frac{1}{2M} \sum_{j=1}^{M-1} (X_{t+(j+1)/M} - X_{t+j/M})^2, \\ RV_{t,l,M}^{\text{avg}} &= \frac{1}{B} \sum_{b=1}^B RV_{t,l,M}^b = \frac{1}{B} \sum_{b=1}^B \sum_{j=1}^{l-1} (X_{t+(jB+b)/M} - X_{t+((j-1)B+b)/M})^2, \end{aligned} \quad (10)$$

and $Bl \cong M$; l denotes the subsample size and B the number of subsamples.

In particular, for each considered realized measure, we will provide in Section 5 the specific regularity conditions for the validity of the proposed specification tests, in terms of the relative rate of growth of T , M , and l .

In the remainder of the paper, two main cases will be considered. The first is when explicit formulae for the moments of the integrated volatility are available, and so the map between the parameters $(a_0, \dots, a_p, \lambda_1, \dots, \lambda_p)$ and the parameters describing the volatility diffusion in (2) is known in closed form; the second case is when explicit formulae for the moments of the integrated volatility are not available. As detailed in the following section, in the first case the parameters of the model will be estimated with a GMM estimator, while in the second case a simulated method of moments estimator will be employed.

3. THE CASE WHERE THE MOMENTS ARE KNOWN EXPLICITLY

When explicit formulae for the moments of the integrated volatility are available, and so is the map between the parameters $(a_0, \dots, a_p, \lambda_1, \dots, \lambda_p)$ and the parameters describing the volatility

5. However, note that when $p = 2$, in the case of Ornstein–Uhlenbeck and affine processes, $\lambda_2 = 2\lambda_1$. Thus, in this case we have one less parameter to estimate but also one less identifying restriction.

diffusion in (2), we can immediately write the set of moment conditions as

$$\begin{aligned} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}) &= \frac{1}{T} \sum_{t=1}^T \mathbf{g}_{t,M}(\boldsymbol{\theta}) \\ &= \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T \text{RM}_{t,M} - \text{E}(\text{IV}_1(\boldsymbol{\theta})) \\ \frac{1}{T} \sum_{t=1}^T (\text{RM}_{t,M} - \bar{\text{RM}}_{T,M})^2 - \text{var}(\text{IV}_1(\boldsymbol{\theta})) \\ \frac{1}{T} \sum_{t=1}^T (\text{RM}_{t,M} - \bar{\text{RM}}_{T,M})(\text{RM}_{t-1,M} - \bar{\text{RM}}_{T,M}) - \text{cov}(\text{IV}_1(\boldsymbol{\theta}), \text{IV}_2(\boldsymbol{\theta})) \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T (\text{RM}_{t,M} - \bar{\text{RM}}_{T,M})(\text{RM}_{t-k,M} - \bar{\text{RM}}_{T,M}) - \text{cov}(\text{IV}_1(\boldsymbol{\theta}), \text{IV}_{k+1}(\boldsymbol{\theta})) \end{pmatrix}, \end{aligned} \quad (11)$$

where $\bar{\text{RM}}_{T,M} = T^{-1} \sum_{t=1}^T \text{RM}_{t,M}$ and the moments of integrated volatility are computed under the volatility model implied by the null hypothesis. The GMM estimator can be defined as the minimizer of the quadratic form

$$\hat{\boldsymbol{\theta}}_{T,M} = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta})' \mathbf{W}_{T,M}^{-1} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}). \quad (12)$$

The weighting matrix in (12) is given by

$$\begin{aligned} \mathbf{W}_{T,M} &= \frac{1}{T} \sum_{t=1}^T (\mathbf{g}_{t,M}^* - \bar{\mathbf{g}}_{T,M}^*) (\mathbf{g}_{t,M}^* - \bar{\mathbf{g}}_{T,M}^*)' \\ &\quad + \frac{2}{T} \sum_{v=1}^{pT} w_v \sum_{t=v+1}^T (\mathbf{g}_{t,M}^* - \bar{\mathbf{g}}_{T,M}^*) (\mathbf{g}_{t-v,M}^* - \bar{\mathbf{g}}_{T,M}^*)', \end{aligned} \quad (13)$$

where $w_v = 1 - \frac{v}{pT-1}$, pT denotes the lag truncation parameter, $\bar{\mathbf{g}}_{T,M}^* = T^{-1} \sum_{t=1}^T \mathbf{g}_{t,M}^*$, and

$$\mathbf{g}_{t,M}^* = \begin{pmatrix} \text{RM}_{t,M} \\ (\text{RM}_{t,M} - \bar{\text{RM}}_{T,M})^2 \\ (\text{RM}_{t,M} - \bar{\text{RM}}_{T,M})(\text{RM}_{t-1,M} - \bar{\text{RM}}_{T,M}) \\ \vdots \\ (\text{RM}_{t,M} - \bar{\text{RM}}_{T,M})(\text{RM}_{t-k,M} - \bar{\text{RM}}_{T,M}) \end{pmatrix}. \quad (14)$$

Note that the vector $\bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta})$ is $(2p+2) \times 1$, while the parameter space $\boldsymbol{\Theta} \in \mathbb{R}^{2p+1}$; therefore, the use of $\bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta})$ in estimating $\boldsymbol{\theta}$ imposes one overidentifying restriction.

Indeed, GMM is not the only available estimation procedure. For example, Barndorff-Nielsen and Shephard (2002) suggested a quasi maximum likelihood estimator (QMLE) using a state-space approach, based on the series of realized volatilities. Thus, QMLE explicitly takes into account the measurement error between realized and integrated volatility. In the present context, we limit our attention to (simulated) GMM, as our objective is to provide a specification test based on the validity of overidentifying restrictions.

We can define the minimizer of the limiting quadratic form

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta})' \mathbf{W}_{\infty}^{-1} \bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta}), \quad (15)$$

where $\bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta})$ and \mathbf{W}_{∞}^{-1} are the probability limits, as T and M go to infinity, of $\bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta})$ and $\mathbf{W}_{T,M}^{-1}$, respectively.

Hereafter, we shall test the following hypothesis

$$H_0 : \bar{\mathbf{g}}_\infty(\boldsymbol{\theta}^*) = \mathbf{0} \quad \text{versus} \quad H_A : \bar{\mathbf{g}}_\infty(\boldsymbol{\theta}^*) \neq \mathbf{0}. \quad (16)$$

In the sequel, we shall need the following set of assumptions.

Assumption 1. *There is a sequence b_M , with $b_M \rightarrow \infty$ as $M \rightarrow \infty$, such that, uniformly in t ,*

- (i) $E(N_{t,M}) = O(b_M^{-1})$,
- (ii) $E(N_{t,M}^2) = O(b_M^{-1})$,
- (iii) *either*
 - (a) $N_{t,M}$ is strong mixing with size $-2r/(r-2)$, where $r > 2$, or
 - (b) $E(N_{t,M}^4) = O(b_M^{-2})$.

Assumption 2. f_t is a time reversible process.

Assumption 3. *The spectrum of the infinitesimal generator operator \mathcal{A} of f_t is discrete, and denoted by $\lambda_0 = 0 < \lambda_1 < \dots < \lambda_i < \lambda_{i+1}$, where $i \in \mathbb{N}$ and $-\lambda_i$ is the eigenvalue associated with the i -th eigenfunction $P_i(f_t)$.*

Assumption 4. $\boldsymbol{\Theta}$ is a compact set of \mathbb{R}^{2p+1} , with p finite.

- Assumption 5.**
- (i) $\hat{\boldsymbol{\theta}}_{T,M}$ and $\boldsymbol{\theta}^*$ are in the interior of $\boldsymbol{\Theta}$,
 - (ii) $E(\partial \mathbf{g}_{t,M}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} | \boldsymbol{\theta}^*)$ is of full rank,
 - (iii) $\bar{\mathbf{g}}_\infty(\boldsymbol{\theta})' \mathbf{W}_\infty^{-1} \bar{\mathbf{g}}_\infty(\boldsymbol{\theta})$ has a unique minimizer.

Assumption 1 states some primitive conditions on the measurement error $N_{t,M}$. Basically, it requires that its first, second, and fourth moments approach 0 at a fast enough rate as $M \rightarrow \infty$, and that $E(N_{t,M} N_{t-k,M})$ declines to 0 fast enough as both $|k|, M \rightarrow \infty$. As we shall see in Section 5, the rate at which b_M^{-1} declines to 0 depends on the specific realized measure we use. Assumptions 2 and 3 are the assumptions used by Meddahi (2001, 2002b) and by Andersen *et al.* (2004) for the moments and covariance structure of $IV_t(\boldsymbol{\theta})$. One-dimensional diffusions are stationary and ergodic if Assumption 2 is satisfied (see, for example, Hansen, Scheinkman and Touzi, 1998) while Assumption 3 holds providing that the infinitesimal generator operator is compact (Hansen *et al.*, 1998) and is satisfied, for example, in the square root or the log-normal volatility models.⁶ Note that correct specification of the integrated volatility process implies the satisfaction of the null hypothesis.

Inspection of Assumption 5(iii) reveals that if two models lead to the same $\bar{\mathbf{g}}_\infty(\boldsymbol{\theta}^*)$, then our procedure cannot distinguish between them. This is the case, for example, of the square root stochastic volatility models considered in Section 6, which cannot be told apart from GARCH stochastic volatility models.

The test statistic for the validity of moment restrictions is given by

$$S_{T,M} = T \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M})' \mathbf{W}_{T,M}^{-1} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M}). \quad (17)$$

The following theorem establishes the limiting behaviour of $S_{T,M}$ under both hypotheses.

6. The spectral decomposition of multivariate diffusion is analysed by Hansen and Scheinkman (1995) and by Chen, Hansen and Scheinkman (2000). The latter paper also addresses the issue of non-parametric estimation of the drift in multi-dimensional diffusion processes.

Theorem 1. Let Assumptions 1–5 hold. If as $T, M \rightarrow \infty$, $T/b_M^2 \rightarrow 0$, $p_T \rightarrow \infty$, and $p_T/T^{1/4} \rightarrow 0$, then, under H_0 ,

$$S_{T,M} \xrightarrow{d} \chi_1^2,$$

and, under H_A ,

$$\Pr(T^{-1}|S_{T,M}| > \varepsilon) \rightarrow 1, \quad \text{for some } \varepsilon > 0.$$

Notice that we require that T grows at a slower rate than b_M^2 . Thus, the slower the rate of growth of b_M , the stronger is this requirement. The rate of growth of b_M depends on the specific realized measure $RM_{t,M}$ used and will be specified explicitly in Section 5.

As usual, once the null hypothesis is rejected, inspection of the moment condition vector provides some insights on the nature of the violation.

Remark 1. The available central limit theorems for the different realized measures (Jacod, 1994; Jacod and Protter, 1998 for realized volatility; Barndorff-Nielsen *et al.*, 2006 for bipower variation; Zhang *et al.*, 2005 for modified subsampled realized volatility) apply to the case in which the discrete interval between successive observations approaches 0 and the time span remains fixed. In general, the cited papers show that $b_M^{1/2}(RM_{\bar{T}M} - IV_{\bar{T}})$ has a mixed normal limiting distribution, when $M \rightarrow \infty$ and \bar{T} is fixed. In this paper we deal with a double asymptotics in which both T and M go to infinity, and in order to have a valid limit theory we first need to show that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T (RM_{t,M} - E(RM_{t,M})) = \frac{1}{\sqrt{T}} \sum_{t=1}^T (IV_t - E(IV_t)) + o_p(1)$$

and then that $T^{-1/2} \sum_{t=1}^T (IV_t - E(IV_t))$ satisfies a central limit theorem. For this reason, we need that the memory of IV_t and of $N_{t,M}$ declines at a sufficiently fast rate. This holds for the class of stochastic eigenfunction volatility models; in fact, in this class integrated volatility has a memory decaying at a geometric rate and has an $ARMA(p, p)$ structure, when the number of eigenfunctions, p , is finite.

4. THE CASE WHERE THE MOMENTS ARE NOT KNOWN EXPLICITLY

The testing procedure suggested above requires the knowledge of the specific functional form of the eigenvalues and of the coefficients of the eigenfunctions, λ_i and $a_0, a_i, i = 1, \dots, p$, in terms of the parameters characterizing the volatility process under the null hypothesis.

For the case where this information is not available we can, nevertheless, construct a test based on the comparison between the sample moments of the observed volatility measure and the sample moments of simulated integrated volatility. If the null hypothesis is true, the two sets of moments approach the same limit, as T and M approach infinity, otherwise they will converge to two different sets of limiting moments.

As one can notice from (11), a test for the correct specification of mean, variance, and covariance structure of integrated volatility can be performed without knowledge of the leverage parameter ρ and/or the (return) drift parameter m . This is because we rule out the possibility of a feedback effect from the observable state variable to the unobservable volatility. Then, our objective is to approximate by simulation the first two moments and a given number of covariances (depending on the number of eigenfunctions of the model under the null hypothesis) of the daily volatility process.

This is somewhat different from the situation in which we simulate the path of the process describing (the log of) the price of the financial asset. We sample the simulated paths at the same

frequency as the data and then we match (functions of) the sample moments of the data and of the simulated data using only observations at discrete time $t = 1, \dots, T$. In fact, in the latter case it suffices to ensure that, for $t = 1, \dots, T$, the difference between the simulated skeleton and the simulated continuous trajectories is approaching 0, in a mean square sense, as the sampling interval approaches 0. Broadly speaking, in the latter case it suffices to have a good approximation of the continuous trajectory only at the same frequency of the data, that is, at $t = 1, \dots, T$.

On the other hand, in the current context, this does no longer suffice as we need to approximate all the path, given that daily volatility (say from $t - 1$ to t) is defined as the integral of (instantaneous) volatility over the interval $t - 1$ and t . Pardoux and Talay (1985) provide conditions for uniform, almost sure convergence of the discrete simulated path to the continuous path, for given initial conditions. However, such a result holds only on a finite time span. The intuitive reason is that the uniform, almost sure convergence follows from the modulus of continuity of a diffusion (and of the Brownian motion), which holds only over a finite time span.

Therefore, we shall proceed in the following manner. For any value θ in the parameter space Θ we simulate a path of length $k + 1$, where k is the highest ordered autocovariance that we want to include into the moment conditions, using a discrete time interval, which can be set to be arbitrarily small. As with a finite time span we cannot rely on the ergodic properties of the underlying diffusion; we need to draw the initial value from the invariant distribution of the volatility model under the null hypothesis. Such an invariant distribution is indeed known in most cases; for example, it is a gamma for the square root volatility and it is an inverse gamma for the GARCH-diffusion volatility. Also, at least in the univariate case, we always know the functional form of the invariant density. For each $\theta \in \Theta$, we simulate S paths of length $k + 1$, for S sufficiently large. We then construct the simulated sample moments by averaging the relevant quantities over S . More formally, we proceed as follows.

For any simulation $i = 1, \dots, S$, for $j = 1, \dots, N$ and for any $\theta \in \Theta$, we simulate the volatility paths of length $k + 1$ using a Milstein scheme, that is,

$$\begin{aligned} f_{i,j\xi}(\theta) = & f_{i,(j-1)\xi}(\theta) + \mu(f_{i,(j-1)\xi}(\theta), \theta)\xi - \frac{1}{2}\sigma'(f_{i,(j-1)\xi}(\theta), \theta)\sigma(f_{i,(j-1)\xi}(\theta), \theta)\xi \\ & + \sigma(f_{i,(j-1)\xi}(\theta), \theta)(W_{j\xi} - W_{(j-1)\xi}) \\ & + \frac{1}{2}\sigma'(f_{i,(j-1)\xi}(\theta), \theta)\sigma(f_{i,(j-1)\xi}(\theta), \theta)(W_{j\xi} - W_{(j-1)\xi})^2, \end{aligned} \quad (18)$$

where $\sigma'(\cdot)$ denotes the derivative of $\sigma(\cdot)$ with respect to its first argument, $\{W_{j\xi} - W_{(j-1)\xi}\}$ is i.i.d. $N(0, \xi)$, $f_{i,0}(\theta)$ is drawn from the invariant distribution of the volatility process under the null hypothesis, and finally $N\xi = k + 1$. For each i it is possible to compute the simulated integrated volatility as

$$IV_{i,\tau,N}(\theta) = \frac{1}{N/(k+1)} \sum_{j=1}^{N/(k+1)} \sigma_{i,\tau-1+j\xi}^2(\theta), \quad \tau = 1, \dots, k+1, \quad (19)$$

where $N/(k+1) = \xi^{-1}$, assumed to be an integer for the sake of simplicity, and

$$\sigma_{i,\tau-1+j\xi}^2(\theta) = \psi(f_{i,\tau-1+j\xi}(\theta)).$$

Also, averaging the quantity calculated in (19) over the number of simulations S and over the length of the path $k + 1$ yields

$$\overline{IV}_{S,\tau,N}(\theta) = \frac{1}{S} \sum_{i=1}^S IV_{i,\tau,N}(\theta),$$

and

$$\overline{IV}_{S,N}(\boldsymbol{\theta}) = \frac{1}{k+1} \sum_{\tau=1}^{k+1} \overline{IV}_{S,\tau,N}(\boldsymbol{\theta}),$$

respectively. We are now in a position to define the set of moment conditions as

$$\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1}^T \mathbf{g}_{t,M}^* - \frac{1}{S} \sum_{i=1}^S \mathbf{g}_{i,N}(\boldsymbol{\theta}), \quad (20)$$

where $\mathbf{g}_{t,M}^*$ is defined as in (14) and

$$\frac{1}{S} \sum_{i=1}^S \mathbf{g}_{i,N}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{S} \sum_{i=1}^S IV_{i,1,N}(\boldsymbol{\theta}) \\ \frac{1}{S} \sum_{i=1}^S (IV_{i,1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta}))^2 \\ \left(\frac{1}{S} \sum_{i=1}^S (IV_{i,1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta}))(IV_{i,2,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})) \right) \\ \vdots \\ \frac{1}{S} \sum_{i=1}^S (IV_{i,1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta}))(IV_{i,k+1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})) \end{pmatrix}. \quad (21)$$

Similarly to the case analysed in the previous section, it is possible to define the simulated method of moments estimator as the minimizer of the quadratic form

$$\hat{\boldsymbol{\theta}}_{T,S,M,N} = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} (\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}))' \mathbf{W}_{T,M}^{-1} (\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta})), \quad (22)$$

where $\mathbf{W}_{T,M}^{-1}$ is defined in (13). Also, define

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} (\bar{\mathbf{g}}_{\infty}^* - \bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta}))' \mathbf{W}_{\infty}^{-1} (\bar{\mathbf{g}}_{\infty}^* - \bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta})), \quad (23)$$

where $\bar{\mathbf{g}}_{\infty}^*$, $\bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta})$, and \mathbf{W}_{∞}^{-1} are the probability limits of $\bar{\mathbf{g}}_{T,M}^*$, $\bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta})$, and $\mathbf{W}_{T,M}^{-1}$, respectively, as T , S , M , and N go to ∞ .

Finally, the statistic for the validity of the moment restrictions is given by

$$Z_{T,S,M,N} = T (\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N}))' \mathbf{W}_{T,M}^{-1} (\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N})). \quad (24)$$

Analogously to the case in which the moment conditions were known, we consider the following hypothesis

$$H_0 : (\bar{\mathbf{g}}_{\infty}^* - \bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta}^*)) = \mathbf{0} \quad \text{versus} \quad H_A : (\bar{\mathbf{g}}_{\infty}^* - \bar{\mathbf{g}}_{\infty}(\boldsymbol{\theta}^*)) \neq \mathbf{0}.$$

Before moving on the study of the asymptotic properties of $Z_{T,S,M,N}$ we need some further assumptions.

Assumption 6. *The drift and variance functions $\mu(\cdot)$ and $\sigma(\cdot)$, as defined in (3), satisfy the following conditions:*

- (1a) $|\mu(f_r(\boldsymbol{\theta}_1), \boldsymbol{\theta}_1) - \mu(f_r(\boldsymbol{\theta}_2), \boldsymbol{\theta}_2)| \leq K_{1,r} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|,$
 $|\sigma(f_r(\boldsymbol{\theta}_1), \boldsymbol{\theta}_1) - \sigma(f_r(\boldsymbol{\theta}_2), \boldsymbol{\theta}_2)| \leq K_{2,r} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|,$
for $0 \leq r \leq k+1$, where $\|\cdot\|$ denotes the Euclidean norm, any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, with $K_{1,r}$, $K_{2,r}$ independent of $\boldsymbol{\theta}$ and $\sup_{r \leq k+1} K_{1,r} = O_p(1)$, $\sup_{r \leq k+1} K_{2,r} = O_p(1)$.

- (1b) $|\mu(f_{r,N}(\theta_1), \theta_1) - \mu(f_{r,N}(\theta_2), \theta_2)| \leq K_{1,r,N} \|\theta_1 - \theta_2\|$,
 $|\sigma(f_{r,N}(\theta_1), \theta_1) - \sigma(f_{r,N}(\theta_2), \theta_2)| \leq K_{2,r,N} \|\theta_1 - \theta_2\|$, where $f_{r,N}(\theta) = f_{\lfloor \frac{Nr}{k+1} \rfloor}(\theta)$ and
 for any $\theta_1, \theta_2 \in \Theta$, with $K_{1,r,N}, K_{2,r,N}$ independent of θ , and $\sup_{r \leq k+1} K_{1,r,N} = O_p(1)$,
 $\sup_{r \leq k+1} K_{2,r,N} = O_p(1)$, uniformly in N .
- (2) $|\mu(x, \theta) - \mu(y, \theta)| \leq C_1 \|x - y\|$, $|\sigma(x, \theta) - \sigma(y, \theta)| \leq C_2 \|x - y\|$,
 where C_1, C_2 are independent of θ .
- (3) $\sigma(\cdot)$ is three times continuously differentiable and $\psi(\cdot)$ is a Lipschitz-continuous function.

Assumption 7.

- (1) $\hat{\theta}_{T,S,M,N}$ and θ^* are in the interior of Θ .
- (2) $\bar{g}_S(\theta)$ is twice continuously differentiable in the interior of Θ , where

$$\mathbf{g}_S(\theta) = \frac{1}{S} \sum_{i=1}^S \mathbf{g}_i(\theta), \quad (25)$$

where

$$\bar{\mathbf{g}}_S(\theta) = \frac{1}{S} \sum_{i=1}^S \mathbf{g}_i(\theta) = \begin{pmatrix} \frac{1}{S} \sum_{i=1}^S \text{IV}_{i,1}(\theta) \\ \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\theta) - \bar{\text{IV}}_S(\theta))^2 \\ \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\theta) - \bar{\text{IV}}_S(\theta))(\text{IV}_{i,2}(\theta) - \bar{\text{IV}}_S(\theta)) \\ \vdots \\ \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\theta) - \bar{\text{IV}}_S(\theta))(\text{IV}_{i,k+1}(\theta) - \bar{\text{IV}}_S(\theta)) \end{pmatrix}, \quad (26)$$

and, for $\tau = 1, \dots, k+1$,

$$\text{IV}_{i,\tau}(\theta) = \int_{\tau-1}^{\tau} \sigma_{i,s}^2(\theta) ds, \quad \bar{\text{IV}}_S(\theta) = \frac{1}{k+1} \sum_{\tau=1}^{k+1} \frac{1}{S} \sum_{i=1}^S \int_{\tau-1}^{\tau} \sigma_{i,s}^2(\theta) ds.$$

- (3) $E(\partial \bar{\mathbf{g}}_1(\theta) / \partial \theta |_{\theta=\theta^*})$ exists and is of full rank.

Assumption 6, (2) and (3) correspond to assumption (ii)' in theorem 6 in Pardoux and Talay (1985), apart from the fact that we also require uniform Lipschitz continuity on the parameter space Θ . Uniform Lipschitz continuity on the real line is a rather strong requirement, which is violated by the most popular stochastic volatility models. However, most stochastic volatility models are locally uniform Lipschitz. For example, the square root volatility model, analysed in the empirical application, is uniform Lipschitz provided that f_t is bounded above from 0, a condition which is satisfied with unit probability. As for the Lipschitz continuity of $\psi(\cdot)$, it is satisfied over bounded sets. Now, note that, since we simulate the paths only over a finite time span, this is not too strong a requirement. In fact, as the diffusion is stationary and (geometrically) ergodic, then the probability that the process escapes from a (large enough) compact set is 0 over a finite time span.

Then, we can state the limiting distribution of $Z_{T,S,M,N}$ under H_0 and the properties of the associated specification test.

Theorem 2. *Let Assumptions 1–7 hold. Also, assume that as $T \rightarrow \infty$, $M \rightarrow \infty$, $S \rightarrow \infty$, $N \rightarrow \infty$, $T/N^{(1-\delta)} \rightarrow 0$, $\delta > 0$, $T/b_M^2 \rightarrow 0$, $p_T \rightarrow \infty$, $p_T/T^{1/4} \rightarrow 0$, and $T/S \rightarrow 0$. Then,*

under H_0 ,

$$Z_{T,S,M,N} \xrightarrow{d} \chi_1^2,$$

and, under H_A ,

$$\Pr(T^{-1}|Z_{T,S,M,N}| > \varepsilon) \rightarrow 1, \quad \text{for some } \varepsilon > 0.$$

Given that we require $T/S \rightarrow 0$, the simulation error is asymptotically negligible, and so it is not surprising that the standard J -test for overidentifying restrictions and the simulation based J -test are asymptotically equivalent. If $T/S \rightarrow \pi$, with $0 < \pi < \infty$, one may expect that $(1 + \pi)^{-1/2} Z_{T,S,M,N}$ still has a χ_1^2 limiting distribution. However, this is not the case. The intuitive reason is that we simulate S volatility paths of finite length $k + 1$, instead of a single path of length S . Therefore, the long-run variance of the simulated moment conditions does not coincide with the long-run variance of the realized volatility moment conditions.

Remark 2. Notice that, in Theorems 1 and 2, we have considered the case of mean, variance, and a given number of autocovariances of IV_t . In principle, there is no particular reason to confine our attention to the set of conditions based on the moments defined in (6). In fact, we could just consider a generic set of moment conditions $E(\phi(RM_{t,M}, \dots, RM_{t-k,M}))$, with the function $\phi: \mathbb{R}^{k+1} \rightarrow \mathbb{R}^{2p+r}$, $r \geq 1$, not necessarily known in closed form, satisfying Assumption 7 above. For any $i = 1, \dots, 2p + r$, we could use a Taylor expansion around integrated volatility, yielding

$$\begin{aligned} \phi_i(RM_{t,M}, \dots, RM_{t-k,M}) &= \phi_i(IV_{t,M}, \dots, IV_{t-k,M}) + \sum_{j=1}^2 \frac{\partial \phi_i}{\partial RM_{t-j,M}} \Big|_{IV_{t-j}} N_{t-j,M} \\ &\quad + \frac{1}{2} \sum_{j=1}^2 \sum_{h=1}^2 \frac{\partial^2 \phi_i}{\partial RM_{t-j,M} \partial RM_{t-h,M}} \Big|_{IV_{t-j}, IV_{t-h}} N_{t-j,M} N_{t-h,M} \\ &\quad + \sum_{j=1}^2 \sum_{h=1}^2 o_p(N_{t-j,M} N_{t-h,M}). \end{aligned}$$

Therefore, the asymptotic validity of a test based on $E(\phi(RM_{t,M}, \dots, RM_{t-k,M}))$ follows by the same argument used in the proof of Theorem 1 if $E(\phi(RM_{t,M}, \dots, RM_{t-k,M}))$ is known explicitly and of Theorem 2 otherwise.

Finally, in order to construct a simulated GMM test, we could also follow an alternative route. We can simulate the trajectories of both the volatility and log-price processes and then sample the latter at the same frequency of the data. Then we can compare the moments of the realized measure of volatility computed using actual and simulated data. If data are simulated from a model, which is correctly specified for both the observable asset and the volatility process, then the two set of moments converge to the same limit as $T \rightarrow \infty$ regardless of M . In the context of the applications analysed below, this is viable only in the case in which we use realized volatility as the chosen realized measure. In that case, if we properly model the leverage effect and if the constant drift specification is correct, then moments of realized volatility and simulated realized volatility approach the same limit as the time span goes to infinity, regardless of whether $M \rightarrow \infty$. However, this is not a viable solution when we use either normalized bipower variation or the modified subsampled realized volatility as the chosen realized measures. In fact, if we simulate the log-price process without jumps, then the moments of actual and simulated realized bipower variation measures do not converge to the same limit for $T \rightarrow \infty$, for fixed

M , unless also the actual log-price process does not exhibit jumps. Analogously, also the moments of actual and simulated subsampled realized volatility cannot converge to the same limit as $T \rightarrow \infty$, for fixed M , unless also the actual log-price process is observed without measurement error.

In the next section the testing procedure outlined above will be specialized to the three considered measures of integrated volatility, namely, realized volatility, bipower variation, and modified subsampled realized volatility.

5. APPLICATIONS TO SPECIFIC ESTIMATORS OF INTEGRATED VOLATILITY

Assumption 1 states some primitive conditions on the measurement error between integrated volatility and realized measure. Basically, it requires that the first, second, and fourth moments of the error approach 0 as $M \rightarrow \infty$, thus implying that the realized measure is a consistent estimator of integrated volatility and that the autocorrelations of $N_{t,M}$, $\text{corr}(N_{t,M}, N_{s,M})$, decline to 0 at a fast enough rate, as $M \rightarrow \infty$ and/or $|t - s| \rightarrow \infty$.

More precisely, if T grows at a slower rate than b_M^2 , then averages over the number of days (scaled by \sqrt{T}) of sample moments of the realized measure and of the integrated volatility process are asymptotically equivalent. It is immediate to see that the slower the rate at which b_M grows, the stronger is the requirement that $T/b_M^2 \rightarrow 0$. In this section, we provide exact rates of growth for b_M and necessary restrictions on the model in (1) and on the measurement error in (4), under which realized volatility, defined as $RV_{t,M}$ in (7), bipower variation, defined as $BV_{t,M}$ in (8) and modified subsampled realized volatility, defined as $\widehat{RV}_{t,M}^\mu$ in (9), satisfy Assumption 1 and then lead to asymptotically valid specification tests.

5.1. Realized volatility

Realized volatility has been suggested as an estimator of integrated volatility by Andersen *et al.* (2001, 2003) and Barndorff-Nielsen and Shephard (2002). When the (log)-price process is a continuous semi-martingale, then realized volatility is a consistent estimator of the increments of the quadratic variation (see, for example, Karatzas and Shreve, 1988, ch. 1). The relevant limit theory, under general conditions, also allowing for generic leverage effects, has been provided by Jacod (1994) and Jacod and Protter (1998), who have shown that

$$\sqrt{M} \left(RV_{\bar{T}M} - \int_0^{\bar{T}} \sigma_s^2 ds \right) \xrightarrow{d} MN \left(0, 2 \int_0^{\bar{T}} \sigma_s^4 ds \right), \quad (27)$$

for given \bar{T} , where the notation $RM_{\bar{T}M}$ in (27), (28), and (29) means that the realized measure has been constructed using intra-daily observations between 0 and \bar{T} .

The result stated above holds for a fixed time span and therefore the asymptotic theory is based on the interval between successive observations approaching 0.

The regularity conditions for the specification test obtained using realized volatility are contained in the following proposition.

Proposition 1. *Let $dz_t = 0$, a.s. and $v = 0$, where dz_t and v are defined in (1) and in (4), respectively. Then Assumption 1 holds with $RM_{t,M} = RV_{t,M}$ for $b_M = O(M)$.*

From the proposition above, we see that, when there are no jumps and no microstructure noise in the price process, then Assumption 1 is satisfied for $b_M = M$ and so Proposition 1 holds with $T/M^2 \rightarrow 0$.

5.2. Bipower variation

Bipower variation has been introduced by Barndorff-Nielsen and Shephard (2004b); the relevant central limit theorem, allowing for leverage effects, is given by Barndorff-Nielsen *et al.* (2006), who have shown that

$$\sqrt{M} \left(\mu_1^{-2} \text{BV}_{\bar{T}M} - \int_0^{\bar{T}} \sigma_s^2 ds \right) \xrightarrow{d} MN \left(0, 2.6090 \int_0^{\bar{T}} \sigma_s^4 ds \right). \quad (28)$$

Again, the provided limit theory holds over a finite time span. As one can immediately see by comparing (27) and (28), robustness to jumps is achieved at the expense of some loss in efficiency. The intuition behind the results by Barndorff-Nielsen *et al.* (2006) is very simple. Since only a finite number of jumps can occur over a finite time span, the probability of having a jump over two consecutive observations will be low, and then this will not induce a bias on the estimator. The fact that when there are no jumps, both $\text{RV}_{\bar{T},M}$ and $\mu_1^{-2} \text{BV}_{\bar{T},M}$ are consistent estimators for $\text{IV}_{\bar{T}}$, with the former being more efficient can be used to construct Hausman-type tests for the null hypothesis of no jumps. For example, Huang and Tauchen (2005) suggest different variants of Hausman tests based on the limit theory of Barndorff-Nielsen and Shephard (2006) and Andersen *et al.* (2005a) provide empirical findings about the relevance of jumps in predicting volatility.

The following proposition states the regularity conditions on the relative rates of growth of T and M for the specification test constructed using bipower variation.

Proposition 2. *Let $v = 0$, where v is defined in (4). Then Assumption 1 holds with $\text{RM}_{t,M} = \text{BV}_{t,M}$ for $b_M = O(M^{1/2})$.*

In the case of large and occasional jumps, the measurement error associated with the bipower variation process satisfies Assumption 1 for $b_M = M^{1/2}$. Thus, in this case, Theorems 1 and 2 apply provided that $T/M \rightarrow 0$.

It may seem a little surprising that, although Barndorff-Nielsen *et al.* (2006) have shown that $\sqrt{M}N_{t,M}$ is asymptotically mixed normal, as in the case of realized volatility, Assumption 1 holds only with $b_M = O(M^{1/2})$. The key point is that their central limit theorem simply ensures that

$$E(N_{t,M}) = o(M^{-1/2}),$$

but, to the best of our knowledge, one cannot in general show that

$$E(N_{t,M}) = O(M^{-1})$$

holds.

5.3. Modified subsampled realized volatility

In order to provide an estimator of integrated volatility robust to microstructure errors, Zhang *et al.* (2005) have proposed a subsampling procedure. Under the specification for the microstructure error term detailed in (4), they show that, in the absence of jumps in the price process,

$$M^{1/6} \left(\widehat{\text{RV}}_{\bar{T}M}^u - \int_0^{\bar{T}} \sigma_s^2 ds \right) \xrightarrow{d} (s^2)^{1/2} N(0, 1), \quad (29)$$

for given \bar{T} , where the asymptotic spread s^2 depends on the variance of the microstructure noise, the length of the fixed time span, and on integrated quarticity. Inspection of the limiting result given in (29) reveals that the cost of achieving robustness to microstructure noise is paid in terms of a slower convergence rate. The logic underlying the subsampled robust realized volatility of Zhang *et al.* is the following. By constructing realized volatility over non-overlapping subsamples, using subsamples of size l , we reduce the bias due to the microstructure error; in fact, the effect of doing so is equivalent to using a lower intra-day frequency. By averaging over different non-overlapping subsamples, we reduce the variance of the estimator. Finally, the estimator of the bias term is constructed using all the M intra-daily observations, and so the error due to the fact that we correct the realized volatility measure using an estimator of the bias instead of the true bias, is asymptotically negligible.⁷ Thus, if there are no jumps, and if the subsample length l is of order $O(M^{1/3})$, and so the number of non-overlapping subsamples is of order $M^{2/3}$, Assumption 1 is satisfied with $RM_{t,M} = \bar{RV}_{t,M}^u$. The regularity conditions are stated precisely in the following proposition.

Proposition 3. *Let $dz_t = 0$ a.s., where dz_t is defined in (1). If $l = O(M^{1/3})$, then Assumption 1 holds with $RM_{t,M} = \bar{RV}_{t,M}^u$ for $b_M = M^{1/3}$.*

It is immediate to see that in this case, T has to grow at a rate slower than $M^{2/3}$. However, this is not too big a problem. In fact, one reason for not using the highest possible frequency is that prices are likely to be contaminated by microstructure error, and in general, the signal-to-noise ratio decreases as the sampling frequency increases. Nevertheless, if we employ a volatility measure which is robust to the effect of microstructure error, we can indeed employ the highest available frequency. In this sense, the requirement that $T/M^{2/3} \rightarrow 0$ is not as stringent as it may seem.

6. EMPIRICAL ILLUSTRATION

In this section, an empirical application of the testing procedure proposed in the previous section will be detailed. A stochastic volatility model very popular both in the theoretical and empirical literature is the square root model proposed by Heston (1993). The model takes its name from the fact that the variance process $\sigma_t^2(\theta)$ is square root, that is,

$$d\sigma_t^2(\theta) = \kappa(\mu - \sigma_t^2(\theta))dt + \eta\sigma_t(\theta)dW_{2,t}, \quad \kappa > 0.$$

Following Meddahi (2001) it is then possible to define α and the unobservable state variable f_t by

$$\alpha = \frac{2\kappa\mu}{\eta^2} - 1, \quad f_t(\theta) = \frac{2\kappa}{\eta^2}\sigma_t^2(\theta).$$

Then the equation describing the dynamic behaviour of f_t is given by

$$df_t(\theta) = \kappa(\alpha + 1 - f_t(\theta))dt + \sqrt{2\kappa}\sqrt{f_t(\theta)}dW_{2,t},$$

and it turns out that the variance process $\sigma_t^2(\theta)$ is explained completely by the first eigenfunction of the infinitesimal generator associated with $f_t(\theta)$ through the equation

$$\begin{aligned} \sigma_t^2(\theta) &= a_0 + a_1 P_1(f_t(\theta)) \\ &= \mu - \frac{\sqrt{\mu}\eta}{\sqrt{2\kappa}} \frac{2\kappa\mu/\eta^2 - f_t(\theta)}{\sqrt{2\kappa\mu/\eta^2}}. \end{aligned} \quad (30)$$

7. Zhang *et al.* (2005) consider a more general set-up in which the sampling interval can be irregular. Also note that, as subsamples cannot overlap, Bl is not exactly equal to M ; however, such an error is negligible as B and l tend to infinity with M .

The corresponding eigenvalue is given by κ . Moreover, in this case $\theta = (\mu, \sqrt{\mu}\eta/\sqrt{2\kappa}, \kappa)'$ and the marginal distribution of $\sigma_t^2(\theta)$ is given by a gamma γ ($\alpha + 1, \mu/(\alpha + 1)$).

Using (6), it is possible to obtain the relevant moments for this specific stochastic volatility model. In fact, by considering

$$\begin{aligned} E(\text{IV}_t(\theta)) &= a_0 = \mu \\ \text{var}(\text{IV}_t(\theta)) &= 2 \frac{a_1^2}{\lambda_1^2} (\exp(-\lambda_1) + \lambda_1 - 1) = \frac{\mu\eta^2}{\kappa^3} (\exp(-\kappa) + \kappa - 1) \\ \text{cov}(\text{IV}_t(\theta), \text{IV}_{t-1}(\theta)) &= a_1^2 \exp \frac{(1 - \exp(-\lambda_1))^2}{\lambda_1} = \frac{\mu\eta^2}{2\kappa} \frac{(1 - \exp(-\kappa))^2}{\kappa^2} \\ \text{cov}(\text{IV}_t(\theta), \text{IV}_{t-2}(\theta)) &= a_1^2 \exp(-\lambda_1) \frac{(1 - \exp(-\lambda_1))^2}{\lambda_1} = \frac{\mu\eta^2}{2\kappa} \exp(-\kappa) \frac{(1 - \exp(\kappa))^2}{\kappa^2} \end{aligned} \quad (31)$$

one obtains exactly one overidentifying restriction to test and the elements of the test statistic defined in (17) are given respectively by

$$\bar{\mathbf{g}}_{T,M}(\theta) = \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T \text{RM}_{t,M} - \mu \\ \frac{1}{T} \sum_{t=1}^T (\text{RM}_{t,M} - \overline{\text{RM}}_M)^2 - \frac{\mu\eta^2}{\kappa^3} (\exp(-\kappa) + \kappa - 1) \\ \frac{1}{T} \sum_{t=1}^T (\text{RM}_{t,M} - \overline{\text{RM}}_M)(\text{RM}_{t-1,M} - \overline{\text{RM}}_M) - \frac{\mu\eta^2}{2\kappa} \frac{(1 - \exp(-\kappa))^2}{\kappa^2} \\ \frac{1}{T} \sum_{t=1}^T (\text{RM}_{t,M} - \overline{\text{RM}}_M)(\text{RM}_{t-2,M} - \overline{\text{RM}}_M) - \frac{\mu\eta^2}{2\kappa} \exp(-\kappa) \frac{(1 - \exp(\kappa))^2}{\kappa^2} \end{pmatrix} \quad (32)$$

and by the results of the calculation required in (13).

The empirical analysis is based on data retrieved from the Trade and Quotation (TAQ) database at the New York Stock Exchange. The TAQ database contains intra-day trades and quotes for all securities listed on the New York Stock Exchange, the American Stock Exchange, and the Nasdaq National Market System. The data is published monthly since 1993. Our sample contains the three most liquid stocks included in the Dow Jones Industrial Average, namely, General Electric, Intel, and Microsoft, and extends from 1 January 1997 until 24 December 2002, for a total of 1509 trading days.⁸ Our choice for the stocks included in the sample is motivated by the need of sufficient liquidity in order to compute the subsampled robust realized volatility.

From the original data-set, which includes prices recorded for every trade, we extracted 10-second and 5-minute interval data, similarly to Andersen and Bollerslev (1997). The 5-minute frequency is generally accepted as the highest frequency at which the effect of microstructure biases are not too distorting (see Andersen *et al.*, 2001). Conversely, 10-second data have been extracted in order to compute the subsampled robust realized volatility.

The price figures for each 10-second and 5-minute intervals are determined as the interpolated average between the preceding and the immediately following transaction prices, weighted linearly by their inverse relative distance to the required point in time. For example, suppose that the price at 15:29:56 was 11.75 and the next quote at 15:30:02 was 11.80, then the interpolated price at 15:30:00 would be $\exp(1/3 \times \log(11.80) + 2/3 \times \log(11.75)) = 11.766$. From the 10-second and 5-minute price series we calculated 10-second and 5-minute intra-daily returns as the difference between successive log prices.

The New York Stock Exchange opens at 9:30 a.m. and closes at 4:00 p.m. Therefore, a full trading day consists of 2341 (respectively 79) intra-day returns calculated over an interval of 10 seconds (respectively 5 minutes). For some stocks, and in some days, the first transactions

8. Trading days are divided over the different years as follows: 253, 252, 252, 252, 248, and 252 from 1997 to 2002. Note that there are 5 days missing in 2001 due to September 11.

TABLE 1
Values of the test statistic $S_{T,M}$ for different realized measures—General Electric

Days		$RV_{t,M}$	$\mu_1^{-2}BV_{t,M}$	$\widehat{RV}_{t,l,M}^u$	
1–100	1–50	1.83	1.69	0.04	2.79
	51–100			1.40	
101–200	101–150	0.18	0.04	0.45	0.43
	151–200			0.43	
201–300	201–250	2.54	1.87	3.04	1.07
	251–300			0.74	
301–400	301–350	1.85	2.15	0.07	0.83
	351–400			1.71	
401–500	401–450	14.52	15.83	2.08	11.17
	451–500			0.96	
501–600	501–550	0.77	0.27	0.83	1.92
	551–600			1.95	
601–700	601–650	5.24	5.99	1.38	3.78
	651–700			4.49	
701–800	701–750	4.54	4.38	0.13	0.55
	751–800			3.20	
801–900	801–850	0.53	0.25	0.04	1.87
	851–900			1.23	
901–1000	901–950	2.71	2.66	0.57	1.13
	951–1000			0.66	
1001–1100	1001–1050	3.07	3.45	4.36	3.44
	1051–1100			2.39	
1101–1200	1101–1150	2.69	2.96	0.30	1.34
	1151–1200			5.50	
1201–1300	1201–1250	0.27	0.30	0.30	0.80
	1251–1300			0.13	
1301–1400	1301–1350	4.45	2.84	0.50	2.35
	1351–1400			3.75	
1401–1500	1401–1450	7.89	6.00	1.71	5.21
	1451–1500			0.12	

arrive some time after 9:30; in these cases we always set the first available trading price after 9:30 a.m. to be the price at 9:30 a.m. Highly liquid stocks may have more than one price at certain points in time (e.g. 5 or 10 quotations at the same time stamp is very common for Intel and Microsoft); when there exists more than one price at the required interval, we select the last provided quotation. For interpolating a price from a multiple-price neighbourhood, we select the closest provided price for the computation.

The square root model for the volatility component has been tested considering all the realized measures considered in Section 5. In particular, the test has been conducted for

- realized volatility, using a time span of 100 days ($T = 100$) and an intra-daily frequency of 5 minutes ($M = 78$);
- normalized bipower variation, using two different daily time spans ($T = 50, 100$), with $M = 78$;
- subsampled robust realized volatility, using $T = 100$, $M = 2340$, the size of the blocks $l = 30$ and finally the number of blocks $B = 78$.

The results are summarized in Tables 1–3.⁹ The tables reveal some interesting findings. First, it seems that the square root model is a good candidate to describe the dynamic behaviour of

9. Since the analysis using bipower variation has been conducted using two different daily time spans, the corresponding results are split into two columns: the first reports entries for $T = 100$, while the other one refers to the first and last 50 days in the subsample.

TABLE 2
Values of the test statistic $S_{T,M}$ for different realized measures—Intel

Days		$RV_{t,M}$	$\mu_1^{-2}BV_{t,M}$	$\widehat{RV}_{t,l,M}^\mu$
1–100	1–50	1.64	0.13	0.01
	51–100			0.31
101–200	101–150	0.68	0.23	0.09
	151–200			0.71
201–300	201–250	1.11	1.09	0.16
	251–300			3.88
301–400	301–350	1.23	1.56	1.27
	351–400			0.06
401–500	401–450	5.00	3.04	0.64
	451–500			11.67
501–600	501–550	0.01	0.01	0.17
	551–600			0.81
601–700	601–650	6.07	3.18	1.12
	651–700			2.32
701–800	701–750	6.32	3.81	0.38
	751–800			1.97
801–900	801–850	18.33	8.22	1.48
	851–900			2.00
901–1000	901–950	6.39	9.56	2.95
	951–1000			3.40
1001–1100	1001–1050	0.74	2.93	2.21
	1051–1100			1.36
1101–1200	1101–1150	5.29	1.96	1.38
	1151–1200			1.90
1201–1300	1201–1250	6.45	6.12	6.39
	1251–1300			0.16
1301–1400	1301–1350	10.09	0.15	2.42
	1351–1400			3.40
1401–1500	1401–1450	5.41	4.03	0.82
	1451–1500			3.03

the volatility, at least in the chosen sample. In fact, especially for General Electric and Microsoft, the model is rejected only for a relatively small fraction of times, irrespective of the realized measure used.

Second, realized volatility is the measure, which leads to more frequent rejections. This is not surprising, since realized volatility is not robust to either jumps or microstructure noise in the price process. There are cases where the only measure, which does not reject the model, is normalized bipower variation; this could be an indication that in that period, jumps have occurred in the log-price process. For example, this happens for Intel for the periods going from days 701 to 800 and from days 1101 to 1200; for Microsoft, for the periods going from days 901 to 1000. Of course, this hypothesis could be verified more rigorously, for example, by comparing the parameter estimates under the different realized measures and construct Hausman-type tests.

Conversely, there are cases where the only measure, which does not reject the model, is modified subsampled realized volatility. This happens for General Electric for the periods going from days 601 to 700 and from days 701 to 800; for Microsoft, for the periods going from days 1101 to 1200 and from days 1201 to 1300. These findings could be interpreted as a signal that in that period prices are strongly contaminated by microstructure effects.

The test using normalized bipower variation and conducted with $T = 50$, to conform to the regularity conditions in Proposition 2, generally confirms the findings of the test with $T = 100$.

TABLE 3
*Values of the test statistic $S_{T,M}$ for different realized
measures—Microsoft*

Days		$RV_{t,M}$	$\mu_1^{-2}BV_{t,M}$	$\widehat{RV}_{t,l,M}^u$
1–100	1–50	2.50	1.08	3.74
	51–100			1.17
101–200	101–150	1.08	1.25	1.23
	151–200			2.23
201–300	201–250	3.70	1.91	3.89
	251–300			2.26
301–400	301–350	0.76	0.07	0.01
	351–400			1.52
401–500	401–450	1.78	1.20	0.59
	451–500			1.42
501–600	501–550	0.26	0.67	0.83
	551–600			0.96
601–700	601–650	1.69	0.39	0.42
	651–700			0.79
701–800	701–750	1.55	1.93	0.36
	751–800			1.32
801–900	801–850	1.65	0.81	0.20
	851–900			1.06
901–1000	901–950	8.39	2.62	1.21
	951–1000			1.67
1001–1100	1001–1050	1.08	3.30	1.69
	1051–1100			2.02
1101–1200	1101–1150	10.08	9.71	1.42
	1151–1200			2.66
1201–1300	1201–1250	4.20	4.32	2.53
	1251–1300			2.58
1301–1400	1301–1350	14.45	2.79	5.76
	1351–1400			2.21
1401–1500	1401–1450	6.67	7.13	6.69
	1451–1500			3.08

Of course, in this case the power of the test may be particularly low, due to the smaller number of observations used.

Finally, it is worth mentioning the relative stability of the estimated parameters over different stocks and over different time spans. Specifically, μ ranges from 0.0003 to 0.0004, η from -0.02 to 0.05, and κ from 1 to 3.

7. CONCLUSIONS AND EXTENSIONS

In this paper, a testing procedure for the hypothesis of correct specification of the integrated volatility process is proposed.

The procedure is derived by employing the flexible eigenfunction stochastic volatility model of Meddahi (2001), which embeds most of the stochastic volatility models employed in the empirical literature. The proposed tests rely on some recent results of Barndorff-Nielsen and Shephard (2001, 2002), Andersen *et al.* (2004), and Meddahi (2003), establishing the moments and the autocorrelation structure of integrated volatility.

In the paper, we have focused on the case of one asset and one latent factor. Extensions to the case of two or more factors driving the volatility process are straightforward. In fact, following Meddahi (2001) and considering, without loss of generality, the case of two independent factors

$f_{1,t}$ and $f_{2,t}$, it is possible to expand the instantaneous volatility as

$$\sigma_t^2 = \psi(f_{1,t}, f_{2,t}) = \sum_{i=0}^{p_1} \sum_{j=0}^{p_2} a_{i,j} P_{1,i}(f_{1,t}) P_{2,j}(f_{2,t}) \quad \text{with} \quad \sum_{i=0}^{p_1} \sum_{j=0}^{p_2} a_{i,j}^2 < \infty.$$

Then defining

$$P_{i,j}(\mathbf{f}_t) = P_{1,i}(f_{1,t}) P_{2,j}(f_{2,t}) \quad \text{with} \quad \mathbf{f}_t = (f_{1,t}, f_{2,t})'$$

it is possible to use all the results given in the previous sections. Of course, in the multi-factor case, the reversibility assumption is not necessarily satisfied (a test for the reversibility hypothesis has been provided by Darolles, Florens and Gouriéroux, 2004).

Another assumption that can be easily relaxed involves the constancy of the drift term. The suggested test statistics do not require the knowledge of the drift component, but the analytical expressions for the moments of integrated volatility were derived by Andersen *et al.* (2004) under the maintained hypothesis of a constant drift. Therefore, if one wants to remove that assumption, he or she has to resort to the simulated version of the test, which continues to hold.

APPENDIX

In the sequel, let IV_t and $IV_t(\theta^*)$ denote respectively the “true” underlying daily volatility and the daily volatility implied by the null model, respectively.

Also, let \tilde{X}_t denote the log-price process after the jump component has been removed. Correspondingly, $\tilde{BV}_{t,M}$ denotes the bipower variation process in the case of no jumps. For notational simplicity, hereafter we omit the scaling factor $M/(M-1)$, used in the definition of (8).

The proofs of the theorems and propositions require the following lemmas.

A.1. Technical lemmas

Lemma A1. Given Assumptions 1–5, if as $T, M \rightarrow \infty$, $T/b_M^2 \rightarrow 0$, then, under H_0 ,

$$\sqrt{T} \bar{\mathbf{g}}_{T,M}(\theta^*) \xrightarrow{d} N(\mathbf{0}, \mathbf{W}_\infty),$$

where $\mathbf{W}_\infty = \lim_{T \rightarrow \infty} \text{var}(\sqrt{T} \bar{\mathbf{g}}_T(\theta^*))$, and

$$\begin{aligned} & \bar{\mathbf{g}}_T(\theta^*) \\ &= \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T IV_t - E(IV_1(\theta^*)) \\ \frac{1}{T} \sum_{t=1}^T (IV_t - \bar{IV})^2 - \text{var}(IV_1(\theta^*)) \\ \frac{1}{T} \sum_{t=1}^T (IV_t - \bar{IV})(IV_{t-1} - \bar{IV}) - \text{cov}(IV_2(\theta^*), IV_1(\theta^*)) \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T (IV_t - \bar{IV})(IV_{t-k} - \bar{IV}) - \text{cov}(IV_{k+1}(\theta^*), IV_1(\theta^*)) \end{pmatrix}, \end{aligned} \quad (33)$$

with the moments of $IV_t(\theta^*)$ given as in (6) but evaluated at θ^* .

Lemma A2. Given Assumptions 1–5,

(i) if as $M, T \rightarrow \infty$, $T/b_M^2 \rightarrow 0$, $p_T \rightarrow \infty$, and $p_T/T^{1/4} \rightarrow 0$, then

$$\mathbf{W}_{T,M}^{-1} \xrightarrow{p} \mathbf{W}_\infty^{-1},$$

where, under H_0 , $\mathbf{W}_\infty = \lim_{T, M \rightarrow \infty} \text{var}(\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{g}_{t,M}(\theta^*))$.

(ii) if as $M, T \rightarrow \infty$, $T/b_M^2 \rightarrow 0$, $p_T \rightarrow \infty$, and $p_T/T^{1/4} \rightarrow 0$, then

$$\hat{\boldsymbol{\theta}}_{T,M} \xrightarrow{p} \theta^*.$$

Lemma A3. Given Assumptions 1–7, if as $T, S, M, N \rightarrow \infty$, $T/b_M^2 \rightarrow 0$, $T/N^{(1-\delta)} \rightarrow 0$, $\delta > 0$, $T/S \rightarrow 0$, then, under H_0 ,

$$\sqrt{T}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}^*)) \xrightarrow{d} N(\mathbf{0}, \mathbf{W}_\infty),$$

where \mathbf{W}_∞ is the probability of limit of $\mathbf{W}_{T,M}$ as defined in (13).

Lemma A4. Given Assumptions 1–7, if as $M, T, S, N \rightarrow \infty$, $T/b_M^2 \rightarrow 0$, $T/N^{(1-\delta)} \rightarrow 0$, $\delta > 0$, $T/S \rightarrow 0$, $p_T \rightarrow \infty$, and $p_T/T^{1/4} \rightarrow 0$, then

$$\hat{\boldsymbol{\theta}}_{T,S,M,N} \xrightarrow{p} \boldsymbol{\theta}^*.$$

Lemma A5. Given Assumptions 2 and 3, as $T, M \rightarrow \infty$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \text{BV}_{t,M} = \frac{1}{\sqrt{T}} \sum_{t=1}^T \widetilde{\text{BV}}_{t,M} + O_p(1), \quad (34)$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \text{BV}_{t,M}^2 = \frac{1}{\sqrt{T}} \sum_{t=1}^T \widetilde{\text{BV}}_{t,M}^2 + O_p(1). \quad (35)$$

A.2. Proofs of theorems and propositions

Proof of Theorem 1.

We begin by showing the limiting distribution of the test statistic under the null hypothesis. Via a mean value expansion around $\boldsymbol{\theta}^*$, we have that

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_{T,M} - \boldsymbol{\theta}^*) = (-\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M})' \mathbf{W}_{T,M}^{-1} \nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\bar{\boldsymbol{\theta}}_{T,M}))^{-1} \nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M})' \mathbf{W}_{T,M}^{-1} \sqrt{T} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}^*),$$

where $\bar{\boldsymbol{\theta}}_{T,M} \in (\hat{\boldsymbol{\theta}}_{T,M}, \boldsymbol{\theta}^*)$.

By the uniform law of large numbers for strong mixing processes, as $M, T \rightarrow \infty$ and $T/b_M^2 \rightarrow 0$,

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} |\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}) - \mathbb{E}(\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}))| \xrightarrow{p} 0,$$

where $\bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta})$ is defined as in (33), but evaluated at a generic $\boldsymbol{\theta}$. Given Lemma 2, part (ii), it follows that

$$\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M}) \xrightarrow{p} \mathbf{d}_0 = \mathbb{E}(\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}^*)).$$

Analogously, the same convergence result can be established for $\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_{T,M}(\bar{\boldsymbol{\theta}}_{T,M})$. Now, given Lemma 2,

$$\sqrt{T} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M}) = (\mathbf{I} - \mathbf{d}_\infty (\mathbf{d}'_\infty \mathbf{W}_\infty \mathbf{d}_\infty)^{-1} \mathbf{d}'_\infty \mathbf{W}_\infty^{-1}) \sqrt{T} \bar{\mathbf{g}}_{T,M}(\boldsymbol{\theta}^*) + o_p(1),$$

and therefore, given Lemma 1,

$$\sqrt{T} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M}) \xrightarrow{d} N(\mathbf{0}, (\mathbf{I} - \mathbf{d}_\infty (\mathbf{d}'_\infty \mathbf{W}_\infty \mathbf{d}_\infty)^{-1} \mathbf{d}'_\infty \mathbf{W}_\infty^{-1}) \mathbf{W}_\infty (\mathbf{I} - \mathbf{d}_\infty (\mathbf{d}'_\infty \mathbf{W}_\infty \mathbf{d}_\infty)^{-1} \mathbf{d}'_\infty \mathbf{W}_\infty^{-1})).$$

Finally, given Lemma 2, part (ii), and by noting that $(\mathbf{I} - \mathbf{W}_\infty^{-1/2} \mathbf{d}_\infty (\mathbf{d}'_\infty \mathbf{W}_\infty \mathbf{d}_\infty)^{-1} \mathbf{d}'_\infty \mathbf{W}_\infty^{-1/2})$ is idempotent, then

$$\sqrt{T} \mathbf{W}_{T,M}^{-1/2} \bar{\mathbf{g}}_{T,M}(\hat{\boldsymbol{\theta}}_{T,M}) \xrightarrow{d} N(\mathbf{0}, (\mathbf{I} - \mathbf{W}_\infty^{-1/2} \mathbf{d}_\infty (\mathbf{d}'_\infty \mathbf{W}_\infty \mathbf{d}_\infty)^{-1} \mathbf{d}'_\infty \mathbf{W}_\infty^{-1/2})).$$

The limiting distribution under H_0 then follows straightforwardly from Lemma 4.2 in Hansen (1982).

The rate of divergence under the alternative comes straightforwardly from the fact that $\bar{\mathbf{g}}_\infty(\boldsymbol{\theta}^*) \neq \mathbf{0}$. \parallel

Proof of Theorem 2.

We begin by analysing the behaviour of the statistic under the null hypothesis. By a similar argument as in the proof of Theorem 1, we have that

$$\begin{aligned} & \sqrt{T}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N})) \\ &= [\mathbf{I} - \nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\bar{\boldsymbol{\theta}}_{T,S,M,N}))(\nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N}))' \mathbf{W}_{T,M}^{-1} \\ & \quad \times \nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\bar{\boldsymbol{\theta}}_{T,S,M,N})))]^{-1} \nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\bar{\boldsymbol{\theta}}_{T,S,M,N}))' \mathbf{W}_{T,M}^{-1} \sqrt{T}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}^*)). \end{aligned}$$

Now by Lemma 3,

$$\sqrt{T}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}^*)) \xrightarrow{d} N(\mathbf{0}, \mathbf{W}_\infty),$$

and by Lemma 2, part (i) and Lemma 4, $\mathbf{W}_{T,M}^{-1} \xrightarrow{p} \mathbf{W}_\infty^{-1}$, $\hat{\boldsymbol{\theta}}_{T,S,M,N} \xrightarrow{p} \boldsymbol{\theta}^*$ and $\bar{\boldsymbol{\theta}}_{T,S,M,N} \xrightarrow{p} \boldsymbol{\theta}^*$.

We now need to show that

$$\nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N})) \xrightarrow{p} \mathbf{d}_\infty, \quad \text{and} \quad \nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N})) \xrightarrow{p} \mathbf{d}_\infty,$$

where $\mathbf{d}_\infty = E(\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_T(\boldsymbol{\theta}^*))$, and $\bar{\mathbf{g}}_T(\boldsymbol{\theta}^*)$ is defined in (33).

Given Assumptions 6 and 7,

$$\frac{1}{S} \sum_{i=1}^S \nabla_{\boldsymbol{\theta}} \mathbf{g}_{i,N}(\boldsymbol{\theta}) = \frac{1}{S} \sum_{i=1}^S \nabla_{\boldsymbol{\theta}} \mathbf{g}_i(\boldsymbol{\theta}) + o_p(1),$$

and

$$\frac{1}{S} \sum_{i=1}^S |\nabla_{\boldsymbol{\theta}} \mathbf{g}_i(\boldsymbol{\theta}) - E(\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_1(\boldsymbol{\theta}))| = o_p(1),$$

uniformly in $\boldsymbol{\theta}$, by the uniform law of large numbers. Now, $E(\nabla_{\boldsymbol{\theta}} \bar{\mathbf{g}}_1(\boldsymbol{\theta}))$ is equal to the vector of derivatives of mean, variance and autocovariances of daily volatility under the null hypothesis. Therefore, by noting that $\bar{\mathbf{g}}_{T,M}^*$ does not depend on $\boldsymbol{\theta}$,

$$\nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\hat{\boldsymbol{\theta}}_{T,S,M,N})) = \nabla_{\boldsymbol{\theta}}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}^*)) + o_p(1),$$

and the first term on the R.H.S. above converges in probability to \mathbf{d}_0 .

Finally, divergence under the alternative can be shown along the same lines as in the proof of Theorem 1. \parallel

Proof of Proposition 1.

Let $N_{t,M} = \sum_{j=1}^M N_{t-1+j/M}$, where, from proposition 2.1 in Meddahi (2002a),

$$\begin{aligned} N_{t-1+j/M} &= (m/M)^2 + 2m \int_{t-1+(j-1)/M}^{t-1+j/M} \sigma_s(\boldsymbol{\theta}) dW_s \\ &\quad + 2 \int_{t-1+(j-1)/M}^{t-1+j/M} \left(\int_{t-1+(j-1)/M}^u \sigma_s dW_s \right) \sigma_u dW_u, \end{aligned}$$

and $W_s = \sqrt{1-\rho^2}W_{1,s} + \rho W_{2,s}$. Therefore, $E(N_{t,M}) = \sum_{j=1}^M E(N_{t-1+j/M}) = m^2/M$. This satisfies Assumption 1 part (i).

Also, given that

$$\int_{t-1+(j-1)/M}^{t-1+j/M} \sigma_s(\boldsymbol{\theta}) dW_s \quad \text{and} \quad \int_{t-1+(j-1)/M}^{t-1+j/M} \left(\int_{t-1+(j-1)/M}^u \sigma_s dW_s \right) \sigma_u dW_u$$

are martingale difference series, then $N_{t-1+j/M}$ is uncorrelated with its past.

From proposition 4.2 in Meddahi (2002a), $\text{var}(N_{t,M}) = M^{-1} \sum_{i=1}^p a_i^2 + o_p(1/M)$ uniformly in t . This proves that Assumption 1(ii) is satisfied.

Within the class of eigenfunction stochastic volatility models, both IV_t and σ_t^2 are geometrically mixing processes, and then it follows that $N_{t,M}$ is geometrically mixing. Therefore, Assumption 1(iii-a) is satisfied. \parallel

Proof of Proposition 2.

Given Lemma 5, it suffices to show that Assumption 1 is satisfied in the case of no jumps. Now, by theorem 2.3 in Barndorff-Nielsen *et al.* (2006),

$$\begin{aligned} \sqrt{M}N_{t,M} &= \sqrt{M} \left(\sum_{j=1}^{M-1} \mu_1^{-2} |\tilde{X}_{t-1+(j+1)/M} - \tilde{X}_{t-1+j/M}| |\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}| - \int_{t-1}^t \sigma_s^2 ds \right) \\ &\xrightarrow{d} \left(2.6090 \int_{t-1}^t \sigma_s^4 ds \right)^{1/2} Z, \end{aligned} \quad (36)$$

where Z is a standard normal variate independent of $(\int_{t-1}^t \sigma_s^4 ds)$. It then follows that

$$E(N_{t,M}) = o(M^{-1/2}) \quad \text{and} \quad E(N_{t,M}^2) = O(M^{-1}).$$

Therefore, Assumption 1(i) holds with $b_M = M^{1/2}$ and Assumption 1(ii) holds with $b_M = M$. Finally,

$$|E(N_{t,M} N_{t,s})| \leq E(N_{t,M}^2)^{1/2} E(N_{s,M}^2)^{1/2} = O(M^{-1}) \quad \text{and} \quad E(N_{t,M}^4) = O(b_M^{-2}) \quad \text{for} \quad b_M = M^{1/2},$$

because of Lemma 1(ii) in Corradi, Distaso and Swanson (2006), provided that $E((\sigma_t^2)^9) < \infty$. Thus, Assumption 1(iii-b) holds for $b_M = M^{1/2}$. \parallel

Proof of Proposition 3.

The error term can be rearranged as

$$N_{t,M} = \left(\widehat{RV}_{t,l,M}^u - RV_{t,l,M}^u \right) + \left(RV_{t,l,M}^u - RV_{t,l,M}^{*,\text{avg}} \right) + \left(RV_{t,l,M}^{*,\text{avg}} - IV_t \right),$$

where

$$RV_{t,l,M}^u = RV_{t,l,M}^{\text{avg}} - 2lv,$$

with $RV_{t,l,M}^{\text{avg}}$ defined as in (10), and

$$RV_{t,l,M}^{*,\text{avg}} = \frac{1}{B} \sum_{b=0}^{B-1} RV_{t,l,M}^{*,b} = \frac{1}{B} \sum_{b=1}^B \sum_{j=1}^{l-1} (Y_{t+(jB+b)/M} - Y_{t+((j-1)B+b)/M})^2.$$

Now

$$\begin{aligned} E\left(\widehat{RV}_{t,l,M}^u - RV_{t,l,M}^u\right) &= 2E(l(\widehat{v}_{t,M} - v)) = 2lE\left(\frac{1}{2M} \sum_{j=1}^{M-1} (X_{t+(j+1)/M} - X_{t+j/M})^2 - v\right) \\ &= 2lE\left(\frac{1}{2M} \sum_{j=1}^{M-1} (Y_{t+(j+1)/M} - Y_{t+j/M})^2 + \frac{1}{2M} \sum_{j=1}^{M-1} ((\epsilon_{t+(j+1)/M} - \epsilon_{t+j/M})^2 - 2v)\right. \\ &\quad \left.+ \frac{1}{M} \sum_{j=1}^{M-1} (Y_{t+(j+1)/M} - Y_{t+j/M})(\epsilon_{t+(j+1)/M} - \epsilon_{t+j/M})\right) \\ &= 2lE\left(\frac{1}{2M} \sum_{j=1}^{M-1} (Y_{t+(j+1)/M} - Y_{t+j/M})^2\right) = O(lM^{-1}). \end{aligned}$$

Also,

$$\begin{aligned} E\left(RV_{t,l,M}^u - RV_{t,l,M}^{*,\text{avg}}\right) &= \frac{1}{B} \sum_{b=1}^B \sum_{j=1}^{l-1} E\left((\epsilon_{t+(jB+b)/M} - \epsilon_{t+((j-1)B+b)/M})^2 - 2v\right) \\ &\quad + \frac{1}{B} \sum_{b=1}^B \sum_{j=1}^{l-1} E\left(\epsilon_{t+(jB+b)/M} - \epsilon_{t+((j-1)B+b)/M}\right) \\ &\quad \times (Y_{t+(jB+b)/M} - Y_{t+((j-1)B+b)/M}) = 0. \end{aligned} \quad (37)$$

Let $RV_{t,M}^y = \sum_{j=1}^{M-1} (Y_{t+(j+1)/M} - Y_{t+j/M})^2$, and note that

$$RV_{t,l}^{*,\text{avg}} - IV_t = \left(RV_{t,l}^{*,\text{avg}} - RV_{t,M}^y \right) + \left(RV_{t,M}^y - IV_t \right).$$

The second term on the right hand side above satisfies Assumption 1, by Proposition 1. As for the first term, from the proof of Theorem 2 in Zhang *et al.* (2005),

$$E\left(RV_{t,l}^{*,\text{avg}} - RV_{t,M}^y\right) = O(BM^{-1}).$$

Thus, when $l = O(M^{1/3})$, Assumption 1(i) holds with $b_M = M^{1/3}$.

As for the variance of the error term,

$$\begin{aligned}\text{var}(N_{t,M}) &= \text{var}\left(\widehat{RV}_{t,l,M}^u - RV_{t,l,M}^u\right) + \text{var}\left(RV_{t,l,M}^u - RV_{t,l,M}^{*,\text{avg}}\right) + \text{var}\left(RV_{t,l,M}^{*,\text{avg}} - IV_t\right) \\ &\quad + 2\text{cov}\left(\left(\widehat{RV}_{t,l,M}^u - RV_{t,l,M}^u\right)\left(RV_{t,l,M}^u - RV_{t,l,M}^{*,\text{avg}}\right)\right) \\ &\quad + 2\text{cov}\left(\left(\widehat{RV}_{t,l,M}^u - RV_{t,l,M}^u\right)\left(RV_{t,l,M}^{*,\text{avg}} - IV_t\right)\right) \\ &\quad + 2\text{cov}\left(\left(RV_{t,l,M}^u - RV_{t,l,M}^{*,\text{avg}}\right)\left(RV_{t,l,M}^{*,\text{avg}} - IV_t\right)\right).\end{aligned}\quad (38)$$

The first term of the right hand side of (38) can be rearranged as

$$\begin{aligned}\text{var}\left(\left(\widehat{RV}_{t,l,M}^u - RV_{t,l,M}^u\right)\right) &= \text{var}\left(2l\left(\widehat{v}_{t,M} - v\right)\right) = 4l^2\text{E}\left(\left(\frac{1}{2M}\sum_{j=1}^{M-1}\left(X_{t+(j+1)/M} - X_{t+j/M}\right)^2 - v\right)^2\right) \\ &= \frac{l^2}{M^2}\sum_{j=1}^{M-1}\text{E}\left(\left(\epsilon_{t+(j+1)/M} - \epsilon_{t+j/M}\right)^2 - 2v\right)^2 \\ &\quad + \frac{2l^2}{M^2}\sum_{j=1}^{M-1}\text{E}\left(\left(\left(\epsilon_{t+(j+1)/M} - \epsilon_{t+j/M}\right)^2 - 2v\right)\left(Y_{t+(j+1)/M} - Y_{t+j/M}\right)^2\right) \\ &= O\left(l^2M^{-1}\right).\end{aligned}$$

Similarly, the second term

$$\begin{aligned}\text{var}\left(RV_{t,l,M}^u - RV_{t,l,M}^{*,\text{avg}}\right) &= \text{var}\left(\frac{1}{B}\sum_{b=1}^B\sum_{j=1}^{l-1}\left(\left(\epsilon_{t+(jB+b)/M} - \epsilon_{t+((j-1)B+b)/M}\right)^2 - 2v\right)\right. \\ &\quad \left.+ \frac{1}{B}\sum_{b=1}^B\sum_{j=1}^{l-1}\left(\epsilon_{t+(jB+b)/M} - \epsilon_{t+((j-1)B+b)/M}\right)\right. \\ &\quad \left.\times\left(Y_{t+(jB+b)/M} - Y_{t+((j-1)B+b)/M}\right)\right)\end{aligned}\quad (39)$$

and

$$\text{var}\left(\frac{1}{B}\sum_{b=1}^B\sum_{j=1}^{l-1}\left(\left(\epsilon_{t+(jB+b)/M} - \epsilon_{t+((j-1)B+b)/M}\right)^2 - 2v\right)\right) = O\left(lB^{-1}\right).\quad (40)$$

Also note that

$$\text{var}\left(\frac{1}{B}\sum_{b=1}^B\sum_{j=1}^{l-1}\left(\epsilon_{t+(jB+b)/M} - \epsilon_{t+((j-1)B+b)/M}\right)\left(Y_{t+(jB+b)/M} - Y_{t+((j-1)B+b)/M}\right)\right)$$

and the covariance term obtained expanding (39) are of a smaller order than the variance term in (40).

We are left with the third variance term on the R.H.S. of (38). It can be decomposed into

$$\begin{aligned}\text{var}\left(RV_{t,l,M}^{*,\text{avg}} - IV_t\right) &= \text{var}\left(RV_{t,l,M}^{*,\text{avg}} - RV_{t,M}^y\right) + \text{var}\left(RV_{t,M}^y - IV_t\right) \\ &\quad + 2\text{cov}\left(\left(RV_{t,l,M}^{*,\text{avg}} - RV_{t,M}^y\right)\left(RV_{t,M}^y - IV_t\right)\right).\end{aligned}$$

From the proof of Theorem 2 in Zhang *et al.* (2005), it follows that

$$\text{var}\left(RV_{t,l,M}^{*,\text{avg}} - RV_{t,M}^y\right) = O(BM^{-1}).$$

From Proposition 1, we know that $\text{var}(RV_{t,M}^y - IV_t) = O(M^{-1})$.

Notice also that $\text{cov}((RV_{t,l,M}^{*,\text{avg}} - RV_{t,M}^y)(RV_{t,M}^y - IV_t))$ and the covariance terms in (38) are $o(BM^{-1})$. Thus, Assumption 1(ii) is satisfied for $b_M = M^{1/3}$, given $l = M^{1/3}$ and $B = M^{2/3}$.

Finally, as the microstructure noise is i.i.d. and is independent of the price process,

$$RV_{t,l}^{*,\text{avg}} - RV_{t,M}^y = 2 \sum_{j=1}^{M-1} (Y_{t+(j+1)/M} - Y_{t+j/M}) \sum_{i=1}^{B \wedge j} (1 - i/B) (Y_{t+(j+1-i)/M} - Y_{t+(j-i)/M})$$

is also geometrically mixing. Therefore, Assumption 1(iii-a) holds. \parallel

A.3. Proofs of lemmas

Proof of Lemma 1.

We first need to show that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T RM_{t,M} = \frac{1}{\sqrt{T}} \sum_{t=1}^T IV_t + o_p(1), \quad (41)$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \left(RM_{t,M} - \frac{1}{T} \sum_{t=1}^T RM_{t,M} \right)^2 = \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(IV_t - \frac{1}{T} \sum_{t=1}^T IV_t \right)^2 + o_p(1) \quad (42)$$

and

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(RM_{t,M} - \frac{1}{T} \sum_{t=1}^T RM_{t,M} \right) \left(RM_{t-k,M} - \frac{1}{T} \sum_{t=1}^T RM_{t,M} \right) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(IV_t - \frac{1}{T} \sum_{t=1}^T IV_t \right) \left(IV_{t-k} - \frac{1}{T} \sum_{t=1}^T IV_t \right) + o_p(1). \end{aligned} \quad (43)$$

To show (41), it suffices to show that $\text{var}(\frac{1}{\sqrt{T}} \sum_{t=1}^T N_{t,M}) = o(1)$.

Given Assumption 1(i), and given that $T/b_M^2 \rightarrow 0$, then $\frac{1}{\sqrt{T}} \sum_{t=1}^T E(N_{t,M}) = o(1)$. Also, let $\bar{N}_{t,M} = N_{t,M} - E(N_{t,M})$ and note that

$$\begin{aligned} \text{var} \left(\frac{1}{T^{1/2}} \sum_{t=1}^T (N_{t,M} - E(N_{t,M})) \right) &= \frac{1}{T} \sum_{t=1}^T E(\bar{N}_{t,M})^2 + \frac{1}{T} \sum_{t=1}^T \sum_{s < t} E(\bar{N}_{t,M} \bar{N}_{s,M}) \\ &\quad + \frac{1}{T} \sum_{t=1}^T \sum_{s > t} E(\bar{N}_{t,M} \bar{N}_{s,M}), \end{aligned} \quad (44)$$

where $E(\bar{N}_{t,M}^2) = O(b_M^{-1})$, given A1(ii). Now, if Assumption 1(iii-a) holds, then for the information set defined as $\mathcal{F}_s = \sigma(X_u, \sigma_u, u \leq s)$

$$\begin{aligned} \left| \frac{1}{T} \sum_{t=1}^T \sum_{s < t} E(\bar{N}_{t,M} \bar{N}_{s,M}) \right| &= \left| \frac{1}{T} \sum_{t=1}^T \sum_{s < t} E(\bar{N}_{s,M} E(\bar{N}_{t,M}) | \mathcal{F}_s) \right| \\ &\leq C \frac{1}{T} \sum_{t=1}^T (E(N_{t,M}^2))^{1/2} (E(|N_{t,M}|^r))^{1/r} \sum_{s=0}^{\infty} \alpha_s^{1/2-1/r} \rightarrow 0, \text{ as } M \rightarrow \infty, \end{aligned}$$

because of mixing inequality (see, for example, Davidson, 1994, theorem 14.2). If instead Assumption 1(iii-b) holds, then

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{s < t} E(\bar{N}_{t,M} \bar{N}_{s,M}) = T O(b_M^{-2}) \rightarrow 0 \text{ as } T b_M^{-2} \rightarrow 0.$$

This completes the proof for (41). As for (42),

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(\text{RM}_{t,M} - \frac{1}{T} \sum_{t=1}^T \text{RM}_{t,M} \right)^2 - \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(\text{IV}_t - \frac{1}{T} \sum_{t=1}^T \text{IV}_t \right)^2 \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(\bar{N}_{t,M} - \frac{1}{T} \sum_{t=1}^T \bar{N}_{t,M} \right)^2 + \frac{2}{\sqrt{T}} \sum_{t=1}^T \left(\bar{N}_{t,M} - \frac{1}{T} \sum_{t=1}^T \bar{N}_{t,M} \right) \left(\text{IV}_t - \frac{1}{T} \sum_{t=1}^T \text{IV}_t \right) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \bar{N}_{t,M}^{2,c} + \frac{1}{\sqrt{T}} \sum_{t=1}^T \text{var}(N_{t,M}) + \frac{2}{\sqrt{T}} \sum_{t=1}^T \left(\bar{N}_{t,M} - \frac{1}{T} \sum_{t=1}^T \bar{N}_{t,M} \right) \left(\text{IV}_t - \frac{1}{T} \sum_{t=1}^T \text{IV}_t \right), \quad (45) \end{aligned}$$

where

$$\bar{N}_{t,M}^{2,c} = (N_{t,M} - E(N_{t,M}))^2 - \text{var}(N_{t,M}).$$

Now,

$$\text{var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \bar{N}_{t,M}^{2,c} \right) = \frac{1}{T} \sum_{t=1}^T \bar{N}_{t,M}^{4,c} + \frac{1}{T} \sum_{t=1}^T \sum_{s \neq t}^T \bar{N}_{t,M}^{2,c} \bar{N}_{s,M}^{2,c} \rightarrow 0,$$

if either Assumption 1(iii-a) or 1(iii-b) holds, by the same argument as the one used to prove (41). Also, the second term on the R.H.S. of (45) approaches zero as a direct consequence of Cauchy-Schwartz inequality. This completes the proof of (42).

Finally, (43) follows by a similar argument as that used to show (42). The statement in the lemma then follows from the central limit theorem for mixing processes. \parallel

Proof of Lemma 2.

- (i) Given Lemma 1, by the same argument used in the proof of Lemma 1 and by theorem 2 in Newey and West (1987).
- (ii) Given Assumption 5, it suffices to show that

$$\sup_{\theta \in \Theta} |\bar{\mathbf{g}}_{T,M}(\theta)' \mathbf{W}_{T,M}^{-1} \bar{\mathbf{g}}_{T,M}(\theta) - \bar{\mathbf{g}}_{\infty}(\theta)' \mathbf{W}_{\infty}^{-1} \bar{\mathbf{g}}_{\infty}(\theta)| \xrightarrow{p} 0. \quad (46)$$

The desired result then follows by, for example, Gallant and White (1988, ch. 3). From part (i), we know that $\mathbf{W}_{T,M}^{-1} \xrightarrow{p} \mathbf{W}_{\infty}^{-1}$. First note that, by the same argument used in the proof of Lemma 1

$$\bar{\mathbf{g}}_{T,M}(\theta) = \bar{\mathbf{g}}_T(\theta) + o_p(1)$$

where the remainder term does not depend on θ and $\bar{\mathbf{g}}_T(\theta)$ is defined as in (33), but is evaluated at a generic θ . As IV_t follows an ARMA process, (46) follows from the uniform law of large numbers for α -mixing processes. \parallel

Proof of Lemma 3.

$T^{-1} \sum_{t=1}^T \mathbf{g}_{t,M}^*$ can be treated as in the proof of Lemma 1, so that (41), (42), and (43) hold. Similar to the proof of Lemma 1, we first need to show that

$$\frac{1}{S} \sum_{i=1}^S \text{IV}_{i,1,N}(\theta^*) = \frac{1}{S} \sum_{i=1}^S \text{IV}_{i,1}(\theta^*) + o_p(T^{-1/2}), \quad (47)$$

$$\frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1,N}(\theta^*) - \bar{\text{IV}}_N(\theta^*))^2 = \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\theta^*) - \bar{\text{IV}}(\theta^*))^2 + o_p(T^{-1/2}) \quad (48)$$

and

$$\begin{aligned} & \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1,N}(\theta^*) - \bar{\text{IV}}_N(\theta^*)) (\text{IV}_{i,k+1,N}(\theta^*) - \bar{\text{IV}}_N(\theta^*)) \\ &= \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\theta^*) - \bar{\text{IV}}(\theta^*)) (\text{IV}_{i,k+1}(\theta^*) - \bar{\text{IV}}(\theta^*)) + o_p(T^{-1/2}). \end{aligned} \quad (49)$$

As for (47),

$$\frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1,N}(\boldsymbol{\theta}^*) - \text{IV}_{i,1}(\boldsymbol{\theta}^*)) = \frac{1}{S} \sum_{i=1}^S \left(\frac{1}{N/(k+1)} \sum_{j=1}^{N/(k+1)} \sigma_{i,j\zeta}^2(\boldsymbol{\theta}^*) - \int_0^1 \sigma_{i,u}^2(\boldsymbol{\theta}^*) du \right),$$

where $\zeta^{-1} = N/(k+1)$. Now, let $\sigma_{i,r,N}^2(\boldsymbol{\theta}^*) = \sigma_{i, \lfloor Nr\zeta/(k+1) \rfloor}^2(\boldsymbol{\theta}^*)$, for $0 \leq r \leq k+1$. Given Assumption 6, by corollary 1.8 in Pardoux and Talay (1985), then

$$\sup_{r \leq (k+1)} N^{1/2(1-\delta)} \left| \sigma_{i, \lfloor Nr\zeta/(k+1) \rfloor}^2(\boldsymbol{\theta}^*) - \sigma_{i,r}^2(\boldsymbol{\theta}^*) \right| \xrightarrow{a.s.} 0, \quad \text{as } N \rightarrow \infty \quad (\zeta \rightarrow 0).$$

Thus, it follows that

$$\left| \frac{1}{N/(k+1)} \sum_{j=1}^{N/(k+1)} \sigma_{i,j\zeta}^2(\boldsymbol{\theta}^*) - \int_0^1 \sigma_{i,u}^2(\boldsymbol{\theta}^*) du \right| = O_{a.s.}(N^{-1/2(1-\delta)}) = o_{a.s.}(T^{-1/2}),$$

as $T/N^{1/(1-\delta)} \rightarrow 0$. By a similar argument, (48) and (49) follow too.

Now, let

$$\frac{1}{T} \sum_{t=1}^T \mathbf{g}_t^* = \begin{pmatrix} \frac{1}{T} \sum_{t=1}^T \text{IV}_t \\ \frac{1}{T} \sum_{t=1}^T (\text{IV}_t - \overline{\text{IV}})^2 \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T (\text{IV}_t - \overline{\text{IV}})(\text{IV}_{t-k} - \overline{\text{IV}}) \end{pmatrix},$$

and

$$\frac{1}{S} \sum_{i=1}^S \mathbf{g}_i(\boldsymbol{\theta}^*) = \begin{pmatrix} \frac{1}{S} \sum_{i=1}^S \text{IV}_{i,1}(\boldsymbol{\theta}^*) \\ \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\boldsymbol{\theta}^*) - \overline{\text{IV}}(\boldsymbol{\theta}^*))^2 \\ \vdots \\ \frac{1}{S} \sum_{i=1}^S (\text{IV}_{i,1}(\boldsymbol{\theta}^*) - \overline{\text{IV}}(\boldsymbol{\theta}^*))(\text{IV}_{i,k+1}(\boldsymbol{\theta}^*) - \overline{\text{IV}}(\boldsymbol{\theta}^*)) \end{pmatrix}.$$

Then, given Lemma 1, and given (47), (48), and (49),

$$\begin{aligned} \sqrt{T}(\bar{\mathbf{g}}_{T,M}^* - \bar{\mathbf{g}}_{S,N}(\boldsymbol{\theta}^*)) &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{g}_t^* - \frac{\sqrt{T}}{\sqrt{S}} \frac{1}{\sqrt{S}} \sum_{i=1}^S \mathbf{g}_i(\boldsymbol{\theta}^*) + o_p(1) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T (\mathbf{g}_t^* - \mathbb{E}(\mathbf{g}_1^*)) - \frac{\sqrt{T}}{\sqrt{S}} \frac{1}{\sqrt{S}} \sum_{i=1}^S (\mathbf{g}_i(\boldsymbol{\theta}^*) - \mathbb{E}(\bar{\mathbf{g}}_1(\boldsymbol{\theta}^*))) \\ &\quad + \sqrt{T}(\mathbb{E}(\mathbf{g}_1^*) - \mathbb{E}(\bar{\mathbf{g}}_1(\boldsymbol{\theta}^*))) + o_p(1). \end{aligned} \quad (50)$$

The second last term of the R.H.S. of (50) is 0 under the null hypothesis. As any simulation draw is independent of the others, by the central limit theorem for i.i.d. random variables,

$$\frac{1}{\sqrt{S}} \sum_{i=1}^S (\mathbf{g}_i(\boldsymbol{\theta}^*) - \mathbb{E}(\bar{\mathbf{g}}_1(\boldsymbol{\theta}^*))) = O_p(1)$$

and, as $T/S \rightarrow 0$, the second term of the R.H.S. of (50) is $o_p(1)$. The statement follows by the same argument as the one used in Lemma 1. \parallel

Proof of Lemma 4.

Given Assumption 7 (unique identifiability), it suffices to show that

$$\left| \frac{1}{N/(k+1)} \sum_{j=1}^{N/(k+1)} \sigma_{i,j\zeta}^2(\boldsymbol{\theta}) - \int_0^1 \sigma_{i,u}^2(\boldsymbol{\theta}^*) du \right| = o_p(1) \quad (51)$$

uniformly in θ . The statement then follows from the uniform law of large numbers and unique identifiability, as in the proof of Lemma 2, part (i). Now,

$$\left| \frac{1}{N/(k+1)} \sum_{j=1}^{N/(k+1)} \sigma_{i,j,\xi}^2(\theta) - \int_0^1 \sigma_{i,u}^2(\theta) du \right| \leq \sup_{r \leq k+1} |\sigma_{i,r,N}^2(\theta) - \sigma_{i,r}^2(\theta)| = O_p(N^{-1/2(1-\delta)}) = o_p(1)$$

pointwise in θ , by corollary 1.8 in Pardoux and Talay (1985), given Assumption 6(2) and (3). We now show that Assumption 6, parts (1a) and (1b), also ensure that $\sup_{r \leq k+1} |\sigma_{i,r,N}^2(\theta) - \sigma_{i,r}^2(\theta)|$ is stochastic equicontinuous over Θ .

In fact, for $\theta' \in \Theta$ and $S(\epsilon) = \{\theta : \|\theta - \theta'\| \leq \epsilon\}$,

$$\begin{aligned} & \sup_{r \leq k+1} \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |(\sigma_{i,r,N}^2(\theta) - \sigma_{i,r}^2(\theta)) - (\sigma_{i,r,N}^2(\theta') - \sigma_{i,r}^2(\theta'))| \\ & \leq \sup_{r \leq k+1} \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |\sigma_{i,r}^2(\theta) - \sigma_{i,r}^2(\theta')| + \sup_{r \leq k+1} \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |\sigma_{i,r,N}^2(\theta) - \sigma_{i,r,N}^2(\theta')|. \end{aligned} \quad (52)$$

We begin by showing that the first term of the R.H.S. of (52) is $o_p(1)$. In fact, given the Lipschitz continuity of $\psi(\cdot)$,

$$\begin{aligned} \sup_{r \leq k+1} \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |\sigma_{i,r}^2(\theta) - \sigma_{i,r}^2(\theta')| & \leq C \sup_{r \leq k+1} \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |f_{i,r}(\theta) - f_{i,r}(\theta')| \\ & \leq (k+1) \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |\mu(f_r(\theta), \theta) - \mu(f_r(\theta'), \theta')| \\ & \quad + \sup_{r \leq k+1} |W_r| \sup_{\theta} \sup_{\theta' \in S(\epsilon)} |\sigma(f_r(\theta), \theta) - \sigma(f_r(\theta'), \theta')|, \end{aligned} \quad (53)$$

where $\sup_{r \leq k+1} |W_r| \stackrel{d}{=} |W_{k+1}|$ is bounded in probability (see, for example, Karatzas and Shreve, 1988, ch. 8), and $\stackrel{d}{=}$ denotes equality in distribution. Then, given Assumption 6, part (1a), the R.H.S. of (53) approaches 0 in probability, as $\epsilon \rightarrow 0$.

The second term on the R.H.S. of (52), is $o_p(1)$, uniformly in θ , by an analogous argument. By the same argument as in, for example, Davidson (1994, ch. 21.3), pointwise convergence plus stochastic equicontinuity implies uniform convergence. \parallel

Proof of Lemma 5.

We start from proving (34). We can expand

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} |X_{t-1+(j+1)/M} - X_{t-1+j/M}| |X_{t-1+j/M} - X_{t-1+(j-1)/M}| \\ & = \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} (\tilde{X}_{t-1+(j+1)/M} - \tilde{X}_{t-1+j/M} + c_{j+1,t} \delta_{j+1,t}) \text{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \\ & \quad \times (\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M} + c_{j,t} \delta_{j,t}) \text{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}) \\ & = \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} ((\tilde{X}_{t-1+(j+1)/M} - \tilde{X}_{t-1+j/M}) \text{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \\ & \quad \times (\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}) \text{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M})) \\ & \quad + \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} c_{j,t} \delta_{j,t} (\tilde{X}_{t-1+(j+1)/M} - \tilde{X}_{t-1+j/M}) \text{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \\ & \quad \times \text{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}) \\ & \quad + \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} c_{j+1,t} \delta_{j+1,t} (\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}) \\ & \quad \times \text{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \text{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}) \\ & \quad + \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} c_{j+1,t} \delta_{j+1,t} c_{j,t} \delta_{j,t} \text{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \text{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}), \end{aligned} \quad (54)$$

where $\delta_{j,t} = 1$ if there is at least a jump in the interval $[t-1+(j-1)/M, t-1+j/M]$ and 0 otherwise.

We have to show that

(a)

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} c_{j+1,t} c_{j,t} \delta_{j+1,t} \delta_{j,t} \operatorname{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \\ \times \operatorname{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}) = o_p(1),$$

(b)

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} c_{j+1,t} \delta_{j+1,t} (\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}) \operatorname{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \\ \times \operatorname{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}) = o_p(1).$$

Now, as for (a),

$$\left| \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} c_{j+1,t} c_{j,t} \delta_{j+1,t} \delta_{j,t} \operatorname{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \operatorname{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M}) \right| \\ \leq \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} |c_{j+1,t}| |c_{j,t}| \delta_{j+1,t} \delta_{j,t} = o_p(1).$$

This follows from

$$\mathbb{E} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} |c_{j+1,t}| |c_{j,t}| \delta_{j+1,t} \delta_{j,t} \right)^2 \\ = \frac{1}{T} \sum_{t=1}^T \sum_{j=1}^{M-1} \mathbb{E}(|c_{j+1,t}|^2) \mathbb{E}(|c_{j,t}|^2) \mathbb{E}(\delta_{j+1,t}) \mathbb{E}(\delta_{j,t}) = O(M^{-1}), \quad (55)$$

given that in any unit span of time we have only a finite number of jumps and $\mathbb{E}(\delta_{j,t}) = O(1/M)$.

In order to prove (b), it suffices to show that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} |c_{j+1,t}| \delta_{j+1,t} |\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}| = o_p(1).$$

Since the jumps are independent of the randomness driving the volatility, and as we have assumed independent jumping times,

$$\mathbb{E}(\delta_{j+1,t} |\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}|) = \mathbb{E}(\delta_{i+1,s}) \mathbb{E}(|\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}|)$$

for all t, s, i, j . Also, since we have assumed independent jumping times,

$$\mathbb{E} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} |c_{j+1,t}| \delta_{j+1,t} |\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}| \right)^2 \\ = \frac{1}{T} \sum_{t=1}^T \sum_{j=1}^{M-1} \mathbb{E}(\delta_{i+1,s}) \mathbb{E}(|c_{j+1,t}|^2) \mathbb{E}(|\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}|^2) = O(M^{-1}).$$

Finally, given that the number of jumps per day is finite,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} [(\tilde{X}_{t-1+(j+1)/M} - \tilde{X}_{t-1+j/M}) \operatorname{sgn}(X_{t-1+(j+1)/M} - X_{t-1+j/M}) \\ \times (\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M} + c_{j,t} \delta_{j,t}) \operatorname{sgn}(X_{t-1+j/M} - X_{t-1+(j-1)/M})] \\ = \frac{1}{\sqrt{T}} \sum_{t=1}^T \sum_{j=1}^{M-1} |\tilde{X}_{t-1+(j+1)/M} - \tilde{X}_{t-1+j/M}| |\tilde{X}_{t-1+j/M} - \tilde{X}_{t-1+(j-1)/M}| + O_p(T^{1/2} M^{-1}).$$

This concludes the proof of (34); (35) follows by the same argument. \parallel

Acknowledgements. We wish to thank the editor, Bernard Salanié, two anonymous referees, and Karim Abadir, Andrew Chesher, Atsushi Inoue, Paul Labys, Oliver Linton, Enrique Sentana, Ron Smith, as well as the seminar participants at the 2003 Winter meeting of Econometric Society in Washington, DC, the 2003 Forecasting Financial Markets Conference in Paris, the 2003 Money, Macro and Finance Conference at London Metropolitan University, London School of Economics—Financial Market Group, the 2003 CIREQ-CIRANO Realized Volatility Conference at University of Montreal, IFS-UCL, Rutgers University, Università di Bari, and University of Glasgow for helpful comments. We are particularly indebted to Nour Meddahi and Peter Phillips for helpful suggestions on a previous version of the paper. The authors gratefully acknowledge financial support from the Economic and Social Research Council, grant code R000230006.

REFERENCES

- AÏT-SAHALIA, Y. (2004), "Disentangling Diffusion from Jumps", *Journal of Financial Economics*, **74**, 487–528.
- AÏT-SAHALIA, Y., HANSEN, L. P. and SCHEINKMAN, J. A. (2004), "Operator Methods for Continuous-Time Markov Processes", in Y. Aït-Sahalia and L. P. Hansen (eds.) *Handbook of Financial Econometrics* (Amsterdam: North-Holland) (forthcoming).
- AÏT-SAHALIA, Y., MYKLAND, P. A. and ZHANG, L. (2005a), "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., MYKLAND, P. A. and ZHANG, L. (2005b), "Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise" (NBER Working Paper No. 11380, National Bureau of Economic Research).
- ALTISSIMO, F. and MELE, A. (2005), "Simulated Nonparametric Estimation of Continuous Time Models of Asset Prices and Returns" (Manuscript, London School of Economics).
- ANDERSEN, T. G., BENZONI, L. and LUND, J. (2002), "An Empirical Investigation of Continuous-Time Equity Return Models", *The Journal of Finance*, **57**, 1239–1284.
- ANDERSEN, T. G. and BOLLERSLEV, T. (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets", *Journal of Empirical Finance*, **4**, 115–158.
- ANDERSEN, T. G., BOLLERSLEV, T. and DIEBOLD, F. X. (2005a), "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility" (Manuscript, University of Pennsylvania).
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. and LABYS, P. (2001), "The Distribution of Realized Exchange Rate Volatility", *Journal of the American Statistical Association*, **96**, 42–55.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. and LABYS, P. (2003), "Modelling and Forecasting Realized Volatility", *Econometrica*, **71**, 579–626.
- ANDERSEN, T. G., BOLLERSLEV, T. and MEDDAHI, N. (2004), "Analytic Evaluation of Volatility Forecasts", *International Economic Review*, **45**, 1079–1110.
- ANDERSEN, T. G., BOLLERSLEV, T. and MEDDAHI, N. (2005b), "Correcting the Errors: Volatility Forecast Evaluation Using High Frequency Data and Realized Volatilities", *Econometrica*, **73**, 279–296.
- AWARTANI, B., CORRADI, W. and DISTASO, W. (2005), "Testing Market Microstructure Effects with an Application to the Dow Jones Industrial Average Stocks" (Manuscript, Queen Mary, University of London).
- BANDI, F. M. and RUSSELL, J. R. (2003), "Microstructure Noise, Realized Volatility, and Optimal Sampling" (Manuscript, University of Chicago).
- BANDI, F. M. and RUSSELL, J. R. (2006), "Separating Microstructure Noise from Volatility", *Journal of Financial Economics*, **79**, 655–692.
- BARNDORFF-NIELSEN, O. E., GRAVERSEN, S. E., JACOD, J., PODOLSKII, M. and SHEPHARD, N. (2006), "A Central Limit Theorem for Realized Power and Bipower Variations of Continuous Semimartingales", in Y. Kabanov, R. Lipster and J. Stoyanov (eds.) *From Stochastic Analysis to Mathematical Finance. Festschrift for Albert Shiryaev* (New York: Springer-Verlag) 33–68.
- BARNDORFF-NIELSEN, O. E., GRAVERSEN, S. E., JACOD, J. and SHEPHARD, N. (2005a), "Limit Theorems for Bipower Variation in Financial Econometrics", *Econometric Theory* (forthcoming).
- BARNDORFF-NIELSEN, O. E., HANSEN, P. R., LUNDE, A. and SHEPHARD, N. (2005b), "Designing Realised Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise" (Manuscript, Oxford University).
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2001), "Non-Gaussian OU Based Models and Some of their Uses in Financial Economics", *Journal of the Royal Statistical Society, B*, **63**, 167–241.
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2002), "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models", *Journal of the Royal Statistical Society, B*, **64**, 253–280.
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2004a), "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression and Correlation in Financial Economics", *Econometrica*, **72**, 885–925.
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2004b), "Power and Bipower Variation with Stochastic Volatility and Jumps", *Journal of Financial Econometrics*, **2**, 1–48.
- BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2006), "Econometrics of Testing for Jumps in Financial Economics using Bipower Variation", *Journal of Financial Econometrics*, **4**, 1–30.
- BOLLERSLEV, T., KRETSCHMER, U., PIGORSCH, C. and TAUCHEN, G. (2005), "A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects" (Manuscript, Duke University).
- BOLLERSLEV, T. and ZHOU, H. (2002), "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility", *Journal of Econometrics*, **109**, 33–65.
- CHEN, X., HANSEN, L. P. and SCHEINKMAN, J. A. (2000), "Principal Components and the Long Run" (Manuscript, New York University).

- CHERNOV, M., GALLANT, A. R., GHYSELS, E. and TAUCHEN, G. (2003), "Alternative Models for Stock Price Dynamics", *Journal of Econometrics*, **116**, 225–257.
- CORRADI, V., DISTASO, W. and SWANSON, N. R. (2006), "Predictive Inference for Integrated Volatility", (Manuscript, Queen Mary, University of London).
- CORRADI, V. and WHITE, H. (1999), "Specification Tests for the Variance of a Diffusion", *Journal of Time Series Analysis*, **20**, 253–270.
- DAROLLES, S., FLORENS, J. P. and GOURIÉROUX, C. (2004), "Kernel Based Nonlinear Canonical Analysis and Time Reversibility", *Journal of Econometrics*, **119**, 323–353.
- DAVIDSON, J. (1994) *Stochastic Limit Theory* (Oxford: Oxford University Press).
- DETTE, H., PODOLSKIJ, M. and VETTER, M. (2004), "Estimation of Integrated Volatility in Continuous Time Financial Models with Application to Goodness-of-Fit Testing" (Manuscript, Ruhr-University).
- DETTE, H. and VON LIERES UND WILKAU, C. (2003), "On a Test for a Parametric Form of Volatility in Continuous Time Financial Models", *Finance and Stochastics*, **7**, 363–384.
- DUFFIE, D. and SINGLETON, K. J. (1993), "Simulated Method of Moment Estimation of Markov Models of Asset Prices", *Econometrica*, **61**, 929–952.
- FERMANIAN, J. D. and SALANIÉ, B. (2004), "A Nonparametric Simulated Maximum Likelihood Estimation Method", *Econometric Theory*, **20**, 701–734.
- GALLANT, A., HSIEH, D. and TAUCHEN, G. (1997), "Estimation of Stochastic Volatility Models with Diagnostics", *Journal of Econometrics*, **81**, 159–192.
- GALLANT, A. and TAUCHEN, G. (1996), "Which Moments to Match?", *Econometric Theory*, **12**, 657–681.
- GALLANT, A. and WHITE, H. (1988) *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models* (Oxford: Basil Blackwell).
- GOURIÉROUX, C., MONFORT, A. and RENAULT, E. (1993), "Indirect Inference", *Journal of Applied Econometrics*, **8**, S85–S118.
- HANSEN, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, **50**, 1029–1054.
- HANSEN, L. P. and SCHEINKMAN, J. A. (1995), "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes", *Econometrica*, **63**, 767–804.
- HANSEN, L. P., SCHEINKMAN, J. A. and TOUZI, N. (1998), "Spectral Methods for Identifying Scalar Diffusions", *Journal of Econometrics*, **86**, 1–32.
- HANSEN, P. R. and LUNDE, A. (2005), "Realized Variance and Market Microstructure Noise", *Journal of Business & Economic Statistics* (forthcoming).
- HESTON, S. L. (1993), "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies*, **6**, 327–344.
- HUANG, X. and TAUCHEN, G. (2005), "The Relative Contribution of Jumps to Total Price Variance", *Journal of Financial Econometrics*, **3**, 456–499.
- JACOD, J. (1994), "Limit of Random Measures Associated with the Increments of a Brownian Semimartingale" (Preprint number 120, Laboratoire de Probabilités, Université Pierre et Marie Curie, Paris).
- JACOD, J. and PROTTER, P. (1998), "Asymptotic Error Distribution for the Euler Method for Stochastic Differential Equations", *Annals of Probability*, **26**, 267–307.
- KARATZAS, I. and SHREVE, S. E. (1988) *Brownian Motion and Stochastic Calculus* (New York: Springer-Verlag).
- MEDDAHI, N. (2001), "An Eigenfunction Approach for Volatility Modeling" (Manuscript, University of Montreal).
- MEDDAHI, N. (2002a), "A Theoretical Comparison between Integrated and Realized Volatilities", *Journal of Applied Econometrics*, **17**, 475–508.
- MEDDAHI, N. (2002b), "Moments of Continuous Time Stochastic Volatility Models" (Manuscript, University of Montreal).
- MEDDAHI, N. (2003), "ARMA Representation of Integrated and Realized Variances", *Econometrics Journal*, **6**, 334–355.
- NEWBY, W. and WEST, K. (1987), "A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, **55**, 703–708.
- PARDOUX, E. and TALAY, D. (1985), "Discretization and Simulation of Stochastic Differential Equations", *Acta Applicandae Mathematicae*, **3**, 23–47.
- TAUCHEN, G. and ZHOU, H. (2005), "Identifying Realized Jumps on Financial Markets" (Manuscript, Duke University).
- ZHANG, L. (2004), "Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach" (Manuscript, Carnegie-Mellon University).
- ZHANG, L., MYKLAND, P. A. and AÏT-SAHALIA, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High Frequency Data", *Journal of the American Statistical Association*, **100**, 1394–1411.