Predicting Volatility Conditional Confidence Intervals via Realized Measures*

Valentina Corradi[†] Walter Distaso[‡] Norman R. Swanson[§]

Queen Mary, University of London University of Exeter Rutgers University

February 2005

preliminary and incomplete draft

Abstract

Keywords: Daily volatility, conditional confidence intervals, realized measures, jumps, microstructure noise

JEL classification: C22, C51.

*Corradi and Distaso gratefully acknowledge ESRC grant RES-000-23-0006, and Swanson acknowledges financial support from a Rutgers University Research Council grant.

†Queen Mary, University of London, Department of Economics, Mile End, London, E14NS, UK, email: v.corradi@qmul.ac.uk.

‡University of Exeter, Department of Economics, Streatham Court, Exeter EX4 4PU, UK, email: w.distaso@ex.ac.uk.

§Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, email: nswanson@econ.rutgers.edu.

1 Introduction

In a recent paper, Andersen, Bollerslev, Diebold and Labys (ABDL: 2003) have suggested a novel, model free, approach for forecasting daily volatility. More precisely, they advocate the use of simple, reduced form time series models for realized volatility, where the latter is constructed summing up intradaily squared returns. The predictive ability of a given model is measured via the R^2 from the autoregressive or ARMA models constructed using (the log of) realized volatility. Their findings suggest that these simple ARMA based forecasts for realized volatility outperform most of the volatility models commonly used by practitioners, such as different type GARCH models, for example. The rationale behind their approach is that, as the time interval shrinks, realized volatility converges to the "true" daily volatility, whenever the underlying asset price is a continuous semimartingale. Though tick by tick and ultra high frequency data are now available, they are often contamined by microstrustructure noise and so realized volatility is typically constructed using 5-minute interval returns or even lower frequency. Therefore, these reduced form time series forecasts for realized volatility imply a loss in efficiency relative to the infeasible optimal forecast for the daily volatility process, based on the entire volatility path. For the class of eigenfunction stochastic volatility models of Meddahi (2001), an analytical expression for such loss in efficiency is provided by Andersen, Bollerslev and Meddahi (2004). In particular, they show that the error associated with realized volatility induces a downward bias in the estimated degree of predictability obtained via the R^2 approach mentioned above. To overcome this issue, Andersen, Bollerslev and Meddahi (2005) develop a general model free, feasible procedure to compute adjusted R^2.

All the papers mentioned above are concerned with pointwise prediction of volatility via some ARMA models based on realized measures. On the other hand, there are situations in which we may be interested in predicting conditional confidence interval of daily volatility.

The main objective of this paper is to propose a feasible, model free estimator of the conditional confidence intervals of integrated volatility. From Meddahi (2003), we know that, within the context of eigenfunction stochastic volatility models, integrated volatility follows an $ARMA(p, p)$ structure, where p denotes the number of eigenfunctions. Though, we have a complete characterization of the autoregressive part only, and furthermore, we do not know the marginal distribution of the innovation. For these reasons, we cannot exploit the ARMA representation in order to construct consistent estimator of the conditional confidence intervals of integrated volatility. Thus, we need to follow a different route. We construct a kernel estimator of the conditional density of a given realized volatility measure, conditional on recent observed values of the realized measure itself. By integrating over the evaluation point, we obtain estimators of the conditional distribution function. We provide general conditions on the measurement error between realized measure and integrated volatility, in terms of its moment and covariance structure, under which we can define a sequence of bandwidth parameters under which the kernel estimator of the conditional density is uniformly consistent. We also provide a uniform rate of convergence, which depends on the bias and variance of the kernel estimator as well as on the measurement error. Also, we derive the relative rate, in terms of the number of days T, at which the first absolute moment of the measurement error and the bandwidth parameter have to approach zero, in order to ensure that all three components (bias, variance and contribution of measurement error) approach zero at the same speed.

Suppose that we knew the data generating process for the instantaneous volatility. Unfortunately, this information does not allow to recover the data generating process for the integrated volatility process. Nevertheless, in that case we can construct the kernel density estimator using the integrated volatility process simulated under the null model (and "evaluated" at the estimated parameters) instead of using a realized measure. Under mild regularity conditions, as the sample size and the number of simulations grow at an appropriate rate, the conditional confidence intervals based on kernel estimators of simulated volatility converge to the "true" conditional confidence interval of integrated volatility. A natural question is whether there is some advantage, in term of a faster rate of convergence, in using simulated volatility reather than realized measures. Basically, if the absolute moments of the measurement error approaches zero, as the number of intradaily observations grows, at a rate faster than $T^{-1/2}$, then there is no gain in using simulated integrated volatility, rather than realized measures. Needless to say, the rate at which the absolute moment of the measurement error approaches zero depends on the specific realized measure we use.

Thus, we show that three well known realized measures, that is realized volatility, bipower variation (Barndorff-Nielsen and Shephard 2004a,b) and the robust subsampled realized volatility of Zhang, Mykland and Aït-Sahalia (2004) satisfy the conditions on the measurement error required for the uniform consistency of the estimator based on realized measures. This means that we can provide a feasible model free estimator of the conditional confidence intervals of integrated volatility even in presence of jumps or microstructure noise. We also note that, in the case we knew the data generating process for the instantaneous volatility, there is no gain (in terms of faster rate of convergence) in using simulated integrated volatility instead of realized measures, whenever the number of intraday observations, M, grows at a rate faster than T, T^2 and T^3 for the case of realized volatility, bipower and robust subsampled realized volatility respectively. Though, it may seem quite unplausible to require M to grow faster than T^3, it should pointed out that in the case of robust subsampled realized volatility we do not have to be concerned about the presence of microstructure noise, and therefore we can use very high frequency data, such as at a few seconds interval.

In order to evaluate the goodness of the approximations of the conditional confidence interval estimators based on the three different realized measures, and on simulated integrated volatility, we compare them against an unfeasible, optimal estimator of the confidence interval of integrated volatility, conditional on all the path of past volatility. Finally, we report the findings of an empirical illustration, based on three very liquid stock in the Dow Jones Industrial Average.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 provides a uniform rate of convergence for the conditional confidence interval estimator based on a given realized measure. Section 4 provides a uniform rate of convergence for the conditional confidence interval estimator on simulated integrated volatility, for the case in which we knew the data generating process of the instantaneous volatility process. Section 5 provides conditions under which realized volatility, bipower variation and robust subsampled realized volatility satisfy the conditions on the mesurement error, required for the uniform consistency of the kernel estimator based on realized measures. Section 6 reports a small Monte Carlo study which evaluates the goodness of the approximations of the conditional confidence interval estimators based on the three different realized measures, and on simulated integrated volatility, for different relative rates of growth of T and M. An empirical illustration is given in Section 7. Finally, Section 8 contains some concluding remarks. All proofs are gathered in the Appendix.

2 The Model

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, is modelled as a jump diffusion process with constant drift term and variance term modelled as a measurable function of a latent factor, f_t, which is also generated by a diffusion process. Thus,

$$\mathrm{d}Y_t = m\mathrm{d}t + \mathrm{d}z_t + \sqrt{\sigma_t^2}\left(\sqrt{1-\rho^2}\mathrm{d}W_{1,t} + \rho\mathrm{d}W_{2,t}\right), \tag{1}$$

$W_{1,t}$ and $W_{2,t}$ refer to two independent Brownian motions and volatility is modelled according to the eigenfunction stochastic volatility model of Meddahi (2001), so that

$$\begin{aligned} \sigma_t^2 &= \psi(f_t) = \sum_{i=1}^{p} a_i \mathrm{P}_i(f_t) \\ \mathrm{d}f_t &= \mu(f_t, \boldsymbol{\theta})\mathrm{d}t + \sigma(f_t, \boldsymbol{\theta})\mathrm{d}W_{2,t}, \end{aligned} \tag{2}$$

for some $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\mathrm{P}_i(f_t)$ denotes the i-th eigenfunction of the infinitesimal generator $\mathcal{A}$ associated with the unobservable state variable f_t.[1] The pure jump process $\mathrm{d}z_t$ specified in (1) is such that

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

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

As customary in the literature on stochastic volatility models, the volatility process ia 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 modeled as a linear combination of the eigenfunctions of $\mathcal{A}$ associated to 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 $\mathrm{P}_0(f_t) = 1$ and that $\mathrm{var}(\mathrm{P}_i(f_t)) = 1$, for any $i \neq 0$. Also, when p is infinite, we also require $\sum_{i=1}^{\infty} a_i < \infty$. 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 (2001,2003). Finally, notice that we have assumed a constant drift term.[2]

[1] 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 $\mathrm{P}_i(f_t)$ and eigenvalues $-\lambda_i$ are given by $\mathcal{A}\mathrm{P}_i(f_t) = -\lambda_i \mathrm{P}_i(f_t)$.

[2] This is in line with Bollerslev abd 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 the empirical findings of Andersen and Bollerslev (1997). Indeed, the test statistics suggested below do not require the knowledge of the drift term. However, some of the proofs make use of the fact that the drift is constant.

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 intradaily observations for T days, for

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

where

$$\epsilon_{t+j/M} \sim \text{iid}(0, \nu) \text{ and } \mathrm{E}(\epsilon_{t+j/M} Y_{s+i/M}) = 0 \text{ for all } t, s, j, i. \tag{3}$$

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 independent of the underlying prices. This is consistent with the model considered by Ait-Sahalia, Mykland and Zhang (2003), Zhang, Mykland and Ait-Sahalia (ZMA:2004), Bandi and Russell (2003, 2004).[3] Needless to say, when $\nu = 0$, then $\epsilon_{t+j/M} = 0$ (almost surely), and therefore $X_{t+j/M} = Y_{t+j/M}$.

The daily integrated volatility process at day t is defined as

$$IV_t = \int_{t-1}^{t} \sigma_s^2 \mathrm{d}s. \tag{4}$$

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 IV_t. 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 to the realized measure $RM_{t,M}$. Note that, in the case where $\nu > 0$, any realized measure of integrated volatility is contaminated by two measurement errors, given that the realized measure is constructed using contaminated data.

In the sequel, we shall first construct functionals of kernel estimator of conditional densities based on realized measure.

Then, we will first provide primitive conditions on the measurement error $N_{t,M}$, in terms of its moments and memory structure, ensuring that the kernel conditional density estimators based on realized measures are consistent for the integrated volatility conditional density and we provide a

[3]Recently, Hansen and Lunde (2004) address the issue of time dependence in the microstructure noise, while Awartani, Corradi and Distaso (2004) allow for correlation between the underlying price.

uniform rate. Finally, 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

$$RV_{t,M} = \sum_{j=1}^{M-1} \left(X_{t+(j+1)/M} - X_{t+j/M}\right)^2, \tag{5}$$

(b) normalized bipower variation, given by

$$(\mu_1)^{-1} BV_{t,M} = (\mu_1)^{-1} \frac{M}{M-1} \sum_{j=2}^{M-1} \left|X_{t+(j+1)/M} - X_{t+j/M}\right| \left|X_{t+j/M} - X_{t+(j-1)/M}\right| \tag{6}$$

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

(c) a microstructure robust subsampled based realized volatility measure, $\widehat{RV}^{u}_{t,l,M}$, suggested by Zhang, Mykland and Ait-Sahalia (2004), defined as

$$\widehat{RV}^{u}_{t,l,M} = RV^{avg}_{t,l,M} - 2l\widehat{\nu}_{t,M}, \tag{7}$$

where

$$\widehat{\nu}_{t,M} = \frac{RV_{t,M}}{2M} = \frac{1}{2M} \sum_{j=1}^{M-1} \left(X_{t+\frac{j}{M}} - X_{t+\frac{j-1}{M}}\right)^2$$

and

$$RV^{avg}_{t,l} = \frac{1}{B} \sum_{b=0}^{B} RV^{b}_{t,l} = \frac{1}{B} \sum_{b=0}^{B-1} \sum_{j=b+1}^{M-(B-b-1)} \left(X_{t+\frac{jB}{M}} - X_{t+\frac{(j-1)B}{M}}\right)^2, \tag{8}$$

with $Bl \cong M$, where l denotes the subsample size and B the number of subsamples. The idea of ZMA is the following: first construct B realized volatility measures using l non overlapping subsamples, then take an average of this B realized volatility measures and correct this average by an estimator of the bias term due to market microstructure, where the bias estimator is constructed using a finer grid.[4] Hereafter, $\widehat{RV}^{u}_{t,l,M}$ will be termed modified subsampled realized volatility.

In particular, for each considered realized measure we will provide regularity conditions for the relative speed at which T, M, l go to infinity for the asymptotic validity of the associated specification test for integrated volatility.

[4] ZMA 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 grow.

3 Predicting Volatility Conditional Confidence Intervals via Realized Volatility Measures

Our objective is to predict confidence intervals for daily volatility, via realized volatility measures. We first construct a nonparametric estimator of the conditional density of integrated volatility, constructed using realized measures (and conditional on a given realized volatility measure actually observed at time T), and we integrate over the dependent variable, in order to obtain cumulative conditional distribution. For example, for the confidence interval $[0, u]$, as one-step ahead predictor, we use

$$\widehat{F}_{RM_{T+1}|RM_{T,M}}(u|RM_{T,M}) = \int_0^u \left(\frac{\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1} \mathbf{K}\left(\frac{RM_{t+1,M}-x}{\xi_{2,T}}, \frac{RM_{t,M}-RM_{T,M}}{\xi_{2,T}}\right)}{\frac{1}{T\xi_{1,T}}\sum_{t=1}^{T-1} K\left(\frac{RM_{t,M}-RM_{T,M}}{\xi_{1,T}}\right)} \right) \mathrm{d}x. \tag{9}$$

Thus, if we want to predict $\Pr\left(u_1 \leq IV_{T+1} \leq u_2 | IV_T\right)$ we use

$$\widehat{F}_{RM_{T+1,M}|RM_{T,M}}(u_2|RM_{T,M}) - \widehat{F}_{RM_{T+1,M}|RM_{T,M}}(u_1|RM_{T,M}).$$

Hereafter, let's define $f_{IV_{T+1}|IV_T}(\cdot|\cdot)$ and $f_{IV_T}(\cdot)$ the conditional density of IV_{T+1} given IV_T and the marginal density of IV_T respectively. Recall that, $IV_t = \int_{t-1}^t \sigma_s^2 \mathrm{d}s = \int_{t-1}^t \psi(g_s)\mathrm{d}s$, and note that, because of assumption A2 below, IV_t is a strictly stationary process, and so $f_{IV_{T+1}|IV_T}(\cdot|\cdot) = f_{IV_{t+1}|IV_t}(\cdot|\cdot)$ and $f_{IV_T}(\cdot) = f_{IV_t}(\cdot)$, for $t = 1, 2, \ldots, T$.

In the sequel, we will need the following assumptions.

Assumption A1: There is a sequence b_M, with $b_M \to \infty$ as $M \to \infty$, such that, uniformly in t,

(i) $\mathrm{E}\left(N_{t,M}\right) = O(b_M^{-1})$,

(ii) $\mathrm{E}\left(N_{t,M}^2\right) = O(b_M^{-1})$,

(iii) $\mathrm{E}\left(N_{t,M}^4\right) = O(b_M^{-3/2})$,

(iv) either

- (a) $N_{t,M}$ is strong mixing with size $-r$, where $r > 2$; or
- (b) $\mathrm{E}\left(N_{t,M}N_{s,M}\right) = O(b_M^{-2}) + \alpha_{t-s}O(b_M^{-1})$, where $\alpha_{t-s} = O(|t-s|^{-2})$,

(v) $\frac{1}{T}\sum_{t=1}^{T} |N_{t,M}| = O_P(b_M^{-1/2})$.

Assumption A2: f_t is a time reversible process.

Assumption A3: the spectrum of the infinitesimal generator operator $\mathcal{A}$ of f_t is discrete, and denoted by $0 < \lambda_1 < ... < \lambda_i < ... < \lambda_N$, where λ_i is the eigenvalue associated with the $i-th$ eigenfunction $\mathrm{E}_i(f_t)$.

Assumption A4:

(i) The kernel $\mathbf{K} : \mathbb{R}^d \to \mathbb{R}^+$, $d = 1, 2$ is a symmetric, nonnegative, twice continuously differentiable function, with bounded derivatives, and

$$\int \mathbf{K}(x)\mathrm{d}s = 1, \quad \lim_{\|x\|\to\infty} \|x\|^d \mathbf{K}(x) = 0 \text{ and } \int_{\mathbb{R}^d} \|x\|^d \mathbf{K}(x)\mathrm{d}x < \infty$$

and for $\mu = 0, 1$ $D^\mu \mathbf{K}(\mathbf{x})$ has Fourier transform $\Psi_\mu(r) = (2\pi)^d \int \exp(ir'\mathbf{x}) D^\mu \mathbf{K}(\mathbf{x})\mathrm{d}\mathbf{x}$ that satisfies $\int (1 + \|r\|) |\Psi_\mu(r)| \, dr < \infty$.

(ii) $f_{IV_T}(\cdot)$ and $f_{IV_{T+1},IV_T}(\cdot,\cdot)$ are absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^2$, are twice continuously differentiable on $\mathbb{R}$ and $\mathbb{R}^2$, are bounded and have bounded first derivatives.

Then, we can state the following.

Theorem 1. *Let assumptions A1(ii),(v) and A2-A(4) hold. Then, if* $\xi_{2,T}^3/\xi_{1,T}^2 \to 0$, *as* $T, M \to \infty$, *for any* $u \in U$, *with* U *possibly unbounded, and for any, arbitrarily small* $\eta > 0$,

$$\begin{aligned} & \left|\widehat{F}_{RM_{T+1,M}|RM_{T,M}}(u|RM_{T,M}) - F_{IV_{T+1}|IV_T}(u|RM_{T,M})\right| \\ = \; & O_P\left(b_M^{-1/2+\eta}\xi_{2,T}^{-3}\right) + O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right) + O\left(\xi_{2,T}^2\right). \end{aligned} \tag{10}$$

Given the condition $\xi_{2,T}^3/\xi_{1,T}^2 \to 0$, we can ignore the error associated with the estimated marginal density, because it converges to zero at a faster rate.

The estimated cumulative distribution function has three sources of error. The first error is due to the fact that we use a realized measure, instead of the true integrated volatility, in the construction of the estimator. Thus, the error depends on b_M, which is linked to the number of intradaily observations.

The last two terms are the classical variance and bias term present in the literature on kernel based estimators. As usual, there is a trade-off between minimizing the bias and minimizing the variance of the estimator. In fact, the faster the bandwidth approaches zero, the faster (the slower) the bias (the variance) approaches zero.

Recalling the definition of $\widehat{F}_{RM_{T+1}|RM_T}(u|RM_{T,M})$, given in (9), and recalling that U is possibly unbounded, the rates on the right hand side of (10) are driven by the rates at which

$$\sup_{x\in\mathbb{R}^+}\left|\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1}\mathbf{K}\left(\frac{RM_{t+1,M}-x}{\xi_{2,T}},\frac{RM_{t,M}-RM_{T,M}}{\xi_{2,T}}\right)-\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1}\mathbf{K}\left(\frac{IV_{t+1}-x}{\xi_{2,T}},\frac{IV_t-RM_{T,M}}{\xi_{2,T}}\right)\right|$$

$$\sup_{x\in\mathbb{R}^+}\left(\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1}\mathbf{K}\left(\frac{IV_{t+1}-x}{\xi_{2,T}},\frac{IV_t-RM_{T,M}}{\xi_{2,T}}\right)-E\left(\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1}\mathbf{K}\left(\frac{IV_{t+1}-x}{\xi_{2,T}},\frac{IV_t-RM_{T,M}}{\xi_{2,T}}\right)\right)\right)^2$$

$$\sup_{x\in\mathbb{R}^+}\left(E\left(\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1}\mathbf{K}\left(\frac{IV_{t+1}-x}{\xi_{2,T}},\frac{IV_t-RM_{T,M}}{\xi_{2,T}}\right)\right)-f_{IV_{T+1},IV_T}\left(x,RM_{T,M}\right)\right),$$

approach zero, respectively, uniformly in $x\in\mathbb{R}^+$. In other words, the three terms on the RHS of (10) reflect the uniform rate at which the contribution to measurement error in the estimation of the joint density, the variance and the bias terms of the joint density approach zero uniformly in $x\in\mathbb{R}^+$.

In particular, note that the error due to the variance component, i.e. the second term on the RHS of (10), is of a larger order of probability than the typical one occurring in the pointwise case(see, e.g., Bosq, Ch. 2, 1998). In fact, in the pointwise case we would have $O_P\left(T^{-1/2}\xi_{2,T}^{-1}\right)$ instead of $O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)$. The slower rate is due to the need of deriving a result which holds uniformly on $\mathbb{R}^+$; it comes from a proof based on the Fourier transform of the kernel, first introduced by Bierens (1982) for regression functions with strong mixing processes and then extended to the case of generic derivatives of density and/of regression functions for general near epoch dependent, possibly heterogeneous, processes by Andrews (1990,1995).

It is immediate to see that the necessary conditions to ensure uniform consistency are that

(i) The bandwidth $\xi_{2,T}$ has to approach zero at a slower rate than $T^{-1/4}$;

(ii) b_M has to approach infinity at a faster rate than $\xi_{2,T}^{-6/(1-2\eta)}$.

Needless to say, the uniform rate of convergence is driven by the slower component. Thus, we want to determine the relative rate of growth, in terms of the number of days T, of b_M and $\xi_{2,T}$ which ensure that all the terms on the RHS of (10) are of the same order of probability. After a few simple manipulations, we see that for

$$\xi_{2,T}=O\left(T^{-1/8}\right),\ \ b_M=O\left(T^{5/(4(1-2\eta))}\right)$$

all the terms on the RHS of (10) are of order $T^{-1/4}$.

4 Predicting Volatility Conditional Confidence Intervals via Simulated Daily Volatility

In this section we consider the case in which we know the model generating the instantaneous volatility process, though we do not know the closed form of the conditional density of the integrated volatility process. We proceed in the following way: for any value in the parameter space, we generate S (instantaneous) volatility paths of length k ($k \geq 2$), using as initial value a draw from the invariant distribution, and construct the associated daily volatility process. Parameter can be estimated by SGMM, as in Corradi and Distaso (2004, Theorem 2). Then, we simulate S paths of length 2 using the estimated parameters, and again drawing the initial value from the invariant distribution. More formally: for any simulation $i = 1, \ldots, S$, for $j = 1, \ldots, N$ and for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, we simulate the volatility paths of length $k+1$ using a Milstein scheme, i.e.

$$\begin{aligned} f_{i,jh}(\boldsymbol{\theta}) \quad = \quad & f_{i,(j-1)h}(\boldsymbol{\theta}) + \mu(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta})h - \frac{1}{2}\sigma'(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta})\sigma(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta})h \\ & + \sigma(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta}) \left(W_{i,jh} - W_{i,(j-1)h}\right) \\ & + \frac{1}{2}\sigma'(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta})\sigma(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta}) \left(W_{i,jh} - W_{i,(j-1)h}\right)^2, \end{aligned} \tag{11}$$

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

$$IV_{i,\tau,N}(\boldsymbol{\theta}) = \frac{1}{N/(k+1)} \sum_{j=1}^{N/(k+1)} \sigma^2_{i,\tau-1+jh}(\boldsymbol{\theta}), \quad \tau = 1, \ldots, k+1, \tag{12}$$

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

$$\sigma^2_{i,\tau-1+jh}(\boldsymbol{\theta}) = \psi(f_{i,\tau-1+jh}(\boldsymbol{\theta})).$$

Also, averaging the quantity calculated in (12) over the number of simulations S and over the length of the path $k+1$ yields respectively

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

and

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

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

$$\overline{\mathbf{g}}_{T,M}^{*} - \overline{\mathbf{g}}_{S,N}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_{t,M}^{*} - \frac{1}{S}\sum_{i=1}^{S}\mathbf{g}_{i,N}(\boldsymbol{\theta}), \tag{13}$$

where $\mathbf{g}_{t,M}^{*}$ is defined as

$$\mathbf{g}_{t,M}^{*} = \begin{pmatrix} RM_{t,M} \\ \left(RM_{t,M} - \overline{RM}_{T,M}\right)^2 \\ \left(RM_{t,M} - \overline{RM}_{T,M}\right)\left(RM_{t-1,M} - \overline{RM}_{T,M}\right) \\ \vdots \\ \left(RM_{t,M} - \overline{RM}_{T,M}\right)\left(RM_{t-k,M} - \overline{RM}_{T,M}\right) \end{pmatrix}, \tag{14}$$

$RM_{t,M}$ denotes the particular realized measure used, and $\overline{RM}_{T,M} = \sum_{t=1}^{T} RM_{t,M}$. Also

$$\frac{1}{S}\sum_{i=1}^{S}\mathbf{g}_{i,N}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{S}\sum_{i=1}^{S} IV_{i,1,N}(\boldsymbol{\theta}) \\ \frac{1}{S}\sum_{i=1}^{S}\left(IV_{i,1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})\right)^2 \\ \frac{1}{S}\sum_{i=1}^{S}\left(IV_{i,1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})\right)\left(IV_{i,2,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})\right) \\ \vdots \\ \frac{1}{S}\sum_{i=1}^{S}\left(IV_{i,1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})\right)\left(IV_{i,k+1,N}(\boldsymbol{\theta}) - \overline{IV}_{S,N}(\boldsymbol{\theta})\right) \end{pmatrix}. \tag{15}$$

We can define the Simulated GMM estimator as the minimizer of the quadratic form

$$\widehat{\boldsymbol{\theta}}_{T,S,M,N} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}(\overline{\mathbf{g}}_{T,M}^{*} - \overline{\mathbf{g}}_{S,N}(\boldsymbol{\theta}))'\mathbf{W}_{T,M}^{-1}(\overline{\mathbf{g}}_{T,M}^{*} - \overline{\mathbf{g}}_{S,N}(\boldsymbol{\theta})), \tag{16}$$

where $\mathbf{W}_{T,M}$ is defined as

$$\begin{aligned}\mathbf{W}_{T,M} =& \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{g}_{t,M}^{*} - \overline{\mathbf{g}}_{T,M}^{*}\right)\left(\mathbf{g}_{t,M}^{*} - \overline{\mathbf{g}}_{T,M}^{*}\right)' \\ &+ \frac{2}{T}\sum_{v=1}^{p_T} w_v \sum_{t=v+1}^{T}\left(\mathbf{g}_{t,M}^{*} - \overline{\mathbf{g}}_{T,M}^{*}\right)\left(\mathbf{g}_{t-v,M}^{*} - \overline{\mathbf{g}}_{T,M}^{*}\right)'.\end{aligned} \tag{17}$$

Also, define

$$\boldsymbol{\theta}^{*} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}(\overline{\mathbf{g}}_{\infty}^{*} - \overline{\mathbf{g}}_{\infty}(\boldsymbol{\theta}))'\mathbf{W}_{\infty}^{-1}(\overline{\mathbf{g}}_{\infty}^{*} - \overline{\mathbf{g}}_{\infty}(\boldsymbol{\theta})), \tag{18}$$

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

We can now construct kernel conditional density estimators based on the integrated volatility simulated under the estimated parameter. For $i = 1, \ldots, S$, define:

$$\widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}(u|RM_{T,M})$$

$$= \int_0^u \left(\frac{\frac{1}{T\zeta_{2,T}^2} \sum_{i=1}^S \mathbf{K}\left(\frac{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}}, \frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{2,T}} \right)}{\frac{1}{T\zeta_{1,T}} \sum_{i=1}^S K\left(\frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{1,T}} \right)} \right) \mathrm{d}x, \tag{19}$$

where the bandwidths $\zeta_{1,T}, \zeta_{2,T}$ need not be equal to $\xi_{1,T}, \xi_{2,T}$ used in the previous section. Thus, if we want to predict $\Pr\left(u_1 \leq IV_{T+1} \leq u_2 | IV_T\right)$ we use,

$$\widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}(u_2|RM_{T,M}) - \widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}(u_1|RM_{T,M}).$$

We also need the following further assumptions.

Assumption A5: The drift and variance functions $\mu(\cdot)$ and $\sigma(\cdot)$, as defined in (2), 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) $|\mu(f_{r,N}(\boldsymbol{\theta}_1), \boldsymbol{\theta}_1) - \mu(f_{r,N}(\boldsymbol{\theta}_2), \boldsymbol{\theta}_2)| \leq K_{1,r,N} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|$,

$|\sigma(f_{r,N}(\boldsymbol{\theta}_1), \boldsymbol{\theta}_1) - \sigma(f_{r,N}(\boldsymbol{\theta}_2), \boldsymbol{\theta}_2)| \leq K_{2,r,N} \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|$, where $f_{r,N}(\boldsymbol{\theta}) = f_{\left[\frac{Nrh}{k+1}\right]}(\boldsymbol{\theta})$ and for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, with $K_{1,r,N}$, $K_{2,r,N}$ independent of $\boldsymbol{\theta}$, 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) $|\mu(x, \boldsymbol{\theta}) - \mu(y, \boldsymbol{\theta})| \leq C_1 \|x - y\|$, $\quad |\sigma(x, \boldsymbol{\theta}) - \sigma(y, \boldsymbol{\theta})| \leq C_2 \|x - y\|$,

where C_1, C_2 are independent of $\boldsymbol{\theta}$.

(3) $\sigma(\cdot)$ is three times continuously differentiable and $\psi(\cdot)$ is a Lipschitz-continuous function.

Assumption A6: $(\overline{\mathbf{g}}_\infty^* - \overline{\mathbf{g}}_\infty(\boldsymbol{\theta}^*))'\mathbf{W}_\infty^{-1}(\overline{\mathbf{g}}_\infty^* - \overline{\mathbf{g}}_\infty(\boldsymbol{\theta}^*)) < (\overline{\mathbf{g}}_\infty^* - \overline{\mathbf{g}}_\infty(\boldsymbol{\theta}))'\mathbf{W}_\infty^{-1}(\overline{\mathbf{g}}_\infty^* - \overline{\mathbf{g}}_\infty(\boldsymbol{\theta}))$, for any $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$.

Assumption A7:

(1) $\widehat{\boldsymbol{\theta}}_{T,S,M,N}$ and $\boldsymbol{\theta}^*$ are in the interior of $\boldsymbol{\Theta}$.

(2) $\overline{\mathbf{g}}_S(\boldsymbol{\theta})$ is twice continuously differentiable in the interior of $\boldsymbol{\Theta}$, where

$$\mathbf{g}_S(\boldsymbol{\theta}) = \frac{1}{S}\sum_{i=1}^S \mathbf{g}_i(\boldsymbol{\theta}), \tag{20}$$

where

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

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

$$IV_{i,\tau}(\boldsymbol{\theta}) = \int_{\tau-1}^{\tau}\sigma_{i,s}^2(\boldsymbol{\theta})\,\mathrm{d}s, \qquad \overline{IV}_S(\boldsymbol{\theta}) = \frac{1}{k+1}\sum_{\tau=1}^{k+1}\frac{1}{S}\sum_{i=1}^{S}\int_{\tau-1}^{\tau}\sigma_{i,s}^2(\boldsymbol{\theta})\,\mathrm{d}s.$$

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

Theorem 2. *Let A1-A7, if as* $M, T, S, N \to \infty$, $\zeta_{2,T}^3/\zeta_{1,T}^2 \to 0$, $T/b_M^2 \to 0$, $T/N^{(1-\delta)} \to 0$, $T/S \to 0$, $p_T \to \infty$ *and* $p_T/T^{1/4} \to 0$, *then for any* $u \in U$, *with* U *possibly unbounded,*

$$\begin{aligned} & \left|\widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}(u|RM_{T,M}) - F_{IV_{T+1}|IV_T}(u|RM_{T,M})\right| \\ = \; & O_P\left(T^{-1/2}\zeta_{2,T}^{-3}\right) + O\left(\zeta_{2,T}^2\right) \end{aligned} \quad (22)$$

Theorem 2 reports the uniform rate of convergence for the case where we construct kernel density estimators based on integrated volatility, simulated using a $\sqrt{T}$-consistent estimator for the parameters.

The variance term in this case is different from the one in which we observe the true volatility process, i.e. the second term of the right hand side of (10); in fact in this case the exponent of the bandwidth is lower, and the difference is due to the fact that estimated parameters are used.

Broadly speaking, the simulation error is negligible, as N, the reciprocal of the discrete interval in the path simulation, and S, the number of simulations, grow at a rate faster than T; in fact N and S can be set arbitrarily large. On the other hand, the fact that we simulate the volatility paths using $\sqrt{T}$ estimated parameters increases the probability order of the variance component, which is $T^{-1/2}\zeta_{2,T}^{-3}$ instead of $T^{-1/2}\zeta_{2,T}^{-2}$. This result may seem a little bit surprising; in fact, generally kernel estimators constructed using estimated residuals are asymptotically equivalent to those constructed using true errors. The key condition for this result is that the derivative of the kernel with respect to the parameter has mean zero. In this case, this is in general not true, hence the extra $\zeta_{2,T}^{-1}$ in the right hand side of (22).

It is immediate to see that the necessary condition for uniform consistency is that the bandwidth parameter $\zeta_{2,T}$ converges to zero at a slower rate than $T^{-1/6}$. It is also immediate to see that the rate of growth for the bandwidth, which ensures that the two sources of error are of the same order of probability, is $\zeta_{2,T} = T^{-1/10}$, which implies a uniform convergence rate of $T^{-1/5}$. First, from Theorem 1, we note that for $\xi_{2,T} = O(T^{-1/8})$ and $b_M = O\left(T^{5/(4(1-2\eta))}\right)$, the estimator constructed using realized measure conveges at $T^{-1/4}$ and so at a faster rate than $T^{-1/5}$. This is not overly surprising, as estimated parameters converge at $\sqrt{T}$ while the contribution of measurement error approaches zero at rate $\sqrt{b_M}$ with b_M growing at a rate faster than $T^{5/4}$. As the estimator based on realized measures is model free and easy to compute, it is natural to ask under which conditions on the relative rate of growth of b_M and $\xi_{2,T}$, in terms of T, there is no gain in using an estimator based on simulated integrated volatility. Basically, we need to see under which conditions the RHS of (10) approaches zero at a rate equal to or faster than $T^{-1/5}$.

Thus we set $\xi_{2,T} = T^{-\phi}, \phi > 0$, and find b_M such that $b_M^{-1/2+\eta} T^{3\phi}$ goes to zero at a equal or faster rate than $T^{-1/5}$. This occurs when

$$\frac{1}{10} \leq \phi \leq \frac{3}{20}, \text{ and } b_M \text{ grows at a rate equal or faster than } T^{\frac{2+30\phi}{5}\frac{1}{1-2\eta}}.$$

Therefore it is enough that b_M grows at a faster rate than T, to ensure that there is no advantage in using simulated integrated volatility instead of a given realized measure. As we shall see below, the rate at which b_M grow to infinity with the number of intradaily observations, depends on the specific realized measure we use.

5 Applications to Specific Volatility Realized Measures

Assumption A1 states some primitive conditions on the measurement error between integrated volatility and realized measure. Basically, A1(ii) and A1(v), stating conditions on the order of magnitude of the absolute first moment and second moment of the measurement error, are required for Theorem 1, which establishes a uniform rate of convergence for kernel estimators based on realized measures. On the other hand, A1(i) and A1(iii)(iv), stating conditions on the mean, fourth moment and covariance structure of the measurement error, are required for the $\sqrt{T}-$consistency of the SGMM, which is used in the proof of Theorem 2.

Realized volatility has been suggested as an estimator of integrated volatility by Barndorff-Nielsen and Shephard (2002) and Andersen, Bollerslev, Diebold and Labys (2001, 2003). When the (log) price process is a continuous semimartingale, then realized volatility is a consistent estimator of the increments of the quadratic variation (see e.g. Karatzas and Shreve, 1991, Ch.1). The relevant limit theory, under general conditions, also allowing for generic leverage effects, has been provided by Barndorff-Nielsen and Shephard (2004c), who have shown that

$$\sqrt{M}\left(RV_{\overline{T},M}-\int_0^{\overline{T}}\sigma_s^2\mathrm{d}s\right)\xrightarrow{d}\mathrm{MN}\left(0,2\int_0^{\overline{T}}\sigma_s^4\mathrm{d}s\right),$$

for given $\overline{T}$.

Proposition 1. *Let* $\mathrm{d}z_t=0$*, a.s. and* $\nu=0$*, where* $\mathrm{d}z_t$ *and* ν *are defined in (1) and in (3), respectively. Then Assumption A1 holds with* $RM_{t,M}=RV_{t,M}$ *for* $b_M=O(M)$*.*

Thus, under the same assumptions, the statement in Theorem 1 can be restated as:

$$\begin{aligned}&\left|\widehat{F}_{RV_{T+1,M}|RV_{T,M}}(u|RV_{T,M})-F_{IV_{T+1}|IV_T}(u|RV_{T,M})\right|\\=\;&O_P\left(M^{-1/2+\eta}\xi_{2,T}^{-3}\right)+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O\left(\xi_{2,T}^2\right).\end{aligned}\tag{23}$$

So, a necessary condition for uniform convergence is that M has to approach infinity at a faster rate than $\xi_{2,T}^{-6/(1-2\eta)}$. Also if,

$$\xi_{2,T}=O\left(T^{-1/8}\right),\quad M=O\left(T^{5/(4(1-2\eta))}\right),$$

then all the terms on the RHS of (23) are of order $T^{-1/4}$.

Finally, note that in the case of realized volatility, there is no gain in using an estimator based on integrated simulated volatility whenever $\xi_{2,T}=T^{-\phi}$, $\frac{1}{10}\leq\phi\leq\frac{3}{20}$ and M grows at a rate equal or faster than $T^{\frac{2+30\phi}{5}\frac{1}{1-2\eta}}$.

Bipower variation has been introduced by Barndorff-Nielsen and Shephard (2004b), who have shown that, when the (log) price process contains a finite number of jumps, and when there is no leverage effect, then

$$\sqrt{M}\left(\mu_1^{-2}BV_{\overline{T},M}-\int_0^{\overline{T}}\sigma_s^2\mathrm{d}s\right)\xrightarrow{d}\mathrm{MN}\left(0,2.6090\int_0^{\overline{T}}\sigma_s^4\mathrm{d}s\right).$$

Proposition 2. *Let* $\rho=0$ *and* $\nu=0$*, where* ρ *and* ν *are defined in (1) and in (3), respectively. Then Assumption A1 holds with* $RM_{t,M}=BV_{t,M}$ *for* $b_M=O(M^{1/2})$*.*

Thus, under the same assumptions, the statement in Theorem 1 can be restated as:

$$\begin{aligned}&\left|\widehat{F}_{BV_{T+1,M}|BV_{T,M}}(u|BV_{T,M})-F_{IV_{T+1}|IV_T}(u|BV_{T,M})\right|\\=\;&O_P\left(M^{-1/4+\eta}\xi_{2,T}^{-3}\right)+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O\left(\xi_{2,T}^2\right).\end{aligned}\tag{24}$$

So, a necessary condition for uniform convergence is that M has to approach infinity at a faster rate than $\xi_{2,T}^{-12/(1-2\eta)}$. Also if,

$$\xi_{2,T} = O\left(T^{-1/8}\right), \quad M = O\left(T^{10/(4(1-2\eta))}\right),$$

then all the terms on the RHS of (24) are of order $T^{-1/4}$.

Finally, note that in the case of bipower variation, there is no gain in using an estimator based on integrated simulated volatility whenever $\xi_{2,T} = T^{-\phi}$, $\frac{1}{10} \leq \phi \leq \frac{3}{20}$ and M grows at a rate equal or faster than $T^{\frac{4+60\phi}{5}\frac{1}{1-2\eta}}$.

In order to provide an estimator of integrated volatility robust to microstructure errors, ZMA have proposed a subsampling procedure. Under the specification for the microstructure error term detailed in (3), they show that, in the absence of jumps in the price process,

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

for given $\overline{T}$, 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 (3) 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 ZMA is the following. By constructing realized volatility over non overlapping subsamples, using susbamples 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 intraday frequency. By averaging over different non overlapping subsamples, we reduce the variance of the estimator. Finally, the bias estimator is constructed using all the M intradaily 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. 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 $RM_{t,M} = \widehat{RV}^{u}_{t,M}$. The regularity conditions are stated precisely in the following Proposition.

Proposition 3. *Let* $\mathrm{d}z_t = 0$ *a.s., where* $\mathrm{d}z_t$ *is defined in (1). If* $l = O(M^{1/3})$*, then Assumption A1 holds with* $RM_{t,M} = \widehat{RV}^{u}_{t,l,M}$*, for* $b_M = M^{1/3}$.

Thus, under the same assumptions, the statement in Theorem 1 can be restated as:

$$\left|\widehat{F}_{BV_{T+1,M}|BV_{T,M}}(u|BV_{T,M}) - F_{IV_{T+1}|IV_T}(u|BV_{T,M})\right|$$

$$= O_P\left(M^{-1/4+\eta}\xi_{2,T}^{-3}\right) + O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right) + O\left(\xi_{2,T}^{2}\right). \tag{25}$$

So, a necessary condition for uniform convergence is that M has to approach infinity at a faster rate than $\xi_{2,T}^{-18/(1-2\eta)}$. Also, if

$$\xi_{2,T} = O\left(T^{-1/8}\right), \quad M = O\left(T^{15/(4(1-2\eta))}\right),$$

then, all the terms on the RHS of (25) are of order $T^{-1/4}$.

Finally, note that in the case of realized volatility, there is no gain in using an estimator based on integrated simulated volatility whenever $\xi_{2,T} = T^{-\phi}$, $\frac{1}{10} \leq \phi \leq \frac{3}{20}$ and
M grows at a rate equal or faster than $T^{\frac{6+90\phi}{5}\frac{1}{1-2\eta}}$.

As noted below Theorem 2, we have no gain in using (model based) simulated integrated volatily wheneve b_M grows at a faster rate of T. This implies that the number of intradaily observations should growth at a rate faster than T, T^2 and T^3, for the cases of realized volatility, bipower variation and modified subsampled realized volatility, respectively. Also, unless we let the bandwidth go to zero very slowly, even the necessary condition ensuring that contribution of measurement error approaches zero is quite stringent, and gets more and more stringent passing from realized volatility, bipower variation and modified subsampled realized volatility. Clearly, we also want to choose a rather large value for T, as the estimator for the conditional confidence interval is consistent (at a nonparametric rate) only for T going to infinity. Thus, we need a very large number of intradaily observations, specially for the case of modified subsampled realized volatility. However, the latter measure is robust to microstructure noise, at least for the type outlined in Section 2, and therefore we can use all the available information and sample at very high frequency, such as 1-5 seconds.

6 A Simulation Exercise

We now want to evaluate the accuracy of the volatility confidence interval prediction based on realized measures or on simulated integrated volatility with the best, "unfeasible" predictor based on the "true" confidence interval conditional on the past history of the volatility path. We simulate S paths of f_t as in (11), with

$$\mu(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \kappa\left(\alpha + 1 - f_{i,(j-1)h}\right)$$

$$\sigma(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \sqrt{2\kappa} f_{i,(j-1)h}$$

$$\sigma'(f_{i,(j-1)h}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \frac{1}{2}\sqrt{2\kappa} f_{i,(j-1)h}^{-1/2},$$

so that we are assuming that volatility follows a square root process (see Meddahi (2001) for the one to one representation of square root stochastic volatility and eigenfunction stochastic volatility models). Also, we set $k = 1$ and $\boldsymbol{\theta} = \boldsymbol{\theta}^{\dagger}$, and keep $\left(W_{i,jh} - W_{i,(j-1)h}\right)$ fixed across simulations for $j = 1, \dots, N/2$, while we let it independent across i, for $j = N/2 + 1, \dots, N$. Thus, $f_{i,jh}$ is fixed across simulations and consequently does not depend on i, for $j = 1, \dots, N/2$. Define,

$$IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right) = \frac{1}{N/2}\sum_{j=1}^{N/2} \sigma_{jh}^2(\boldsymbol{\theta}^{\dagger}) \tag{26}$$

$$IV_{i,2,N}\left(\boldsymbol{\theta}^{\dagger}\right) = \frac{1}{N/2}\sum_{j=N/2+1}^{N} \sigma_{i,jh}^2(\boldsymbol{\theta}^{\dagger})$$

and construct estimators of integrated confidence intervals, conditional on $\sigma_{jh}^2(\boldsymbol{\theta}^{\dagger})$, $j = 1, \dots, N/2$ as

$$\widehat{F}_{N,S}(u_2) - \widehat{F}_{N,S}(u_1) = \frac{1}{S}\sum_{i=1}^{S} 1_{\left\{u_1 \leq IV_{i,2,N}\left(\boldsymbol{\theta}^{\dagger}\right) \leq u_2\right\}}$$

and note that, as show in the Appendix,

$$\begin{aligned}
&\left(\widehat{F}_{N,S}(u_2) - \widehat{F}_{N,S}(u_1)\right) - \left(F\left(u_1|\sigma_\tau^2, \tau \in [0,1]\right) - F\left(u_2|\sigma_\tau^2, \tau \in [0,1]\right)\right) \\
= \;& O_P\left(\frac{1}{S^{1/2}}\right) + O_P\left(\frac{1}{N^{1/2-\delta/2}}\right), \text{ for any } \delta > 0,
\end{aligned} \tag{27}$$

where the error on the RHS can be made arbitrarily small by choosing S and N sufficiently large. We then simulate a path of lenght T for X_t, using constant drift and the same specification for instantaneous volatility, i.e. we specify

$$dX_t = mh + \sqrt{\frac{\eta^2}{2\kappa} f_t} dW_{1,t}$$

where f_t follows the same square root process define above, and generate paths for X_t via a Milstein scheme. Sample the simulated process for the X_t at frequency $1/M$ and form the volatility realized measures based on M intradaily observations. Finally, along the same lines as in Section 3, construct

$$\widehat{F}_{RM_{T+1}|RM_{T,M}}\left(u|IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)\right) = \int_0^u \left(\frac{\frac{1}{T\xi_{2,T}^2}\sum_{t=1}^{T-1}\mathbf{K}\left(\frac{RM_{t+1,M}-x}{\xi_{2,T}}, \frac{RM_{t,M}-IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)}{\xi_{2,T}}\right)}{\frac{1}{T\xi_{1,T}}\sum_{t=1}^{T-1} K\left(\frac{RM_{t,M}-IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)}{\xi_{1,T}}\right)} \right) \mathrm{d}x, \tag{28}$$

where $IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)$ is the quantity computed in (26). Then generate simulated paths for the integrated volatility as in Section 4, i.e. simulating paths using estimated parameters, etc. and construct

$$\widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}\left(u|IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)\right)$$

$$= \int_0^u \left(\frac{\frac{1}{T\zeta_{2,T}^2} \sum_{i=1}^S \mathbf{K}\left(\frac{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}}, \frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)}{\zeta_{2,T}} \right)}{\frac{1}{T\zeta_{1,T}} \sum_{i=1}^S K\left(\frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)}{\zeta_{1,T}} \right)} \right) \mathrm{d}x,$$

We can then measure the degree of accuracy (for various range of M and T, and various combination of u_1 and u_2) of the conditional confidence interval estimators based on realized measures by comparing $\widehat{F}_{RM_{T+1}|RM_{T,M}}\left(u_2|IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)\right) - \widehat{F}_{RM_{T+1}|RM_{T,M}}\left((u_1|IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)\right)$ with $\widehat{F}_{N,S}(u_2) - \widehat{F}_{N,S}(u_1)$. Analogously, we can then measure the degree of accuracy (for various range of M and T, and various combination of u_1 and u_2) of the conditional confidence interval estimators based on simulated integrated volatility, by comparing $\widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}\left(u_2|IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)\right) - \widehat{F}_{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)|IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)}\left(u_1|IV_{1,N}\left(\boldsymbol{\theta}^{\dagger}\right)\right)$ with $\widehat{F}_{N,S}(u_2) - \widehat{F}_{N,S}(u_1)$.

to be complete.

7 Empirical Illustration

To be Done

8 Appendix

The proof of Theorem 1 requires the following two Lemmas.

Lemma 1. *Let assumptions A1 and A4 hold. Then, as $T, M \to \infty$, for any $\eta > 0$ arbitrarily small,*

(i)

$$\begin{aligned} &\left| \frac{1}{T} \sum_{t=1}^{T-1} \left(K\left(\frac{RM_{t,M} - RM_{T,M}}{\xi_{1,T}} \right) / \xi_{1,T} - K\left(\frac{IV_t - RM_{T,M}}{\xi_{1,T}} \right) / \xi_{1,T} \right) \right| \\ = \;& O_P(b_M^{-1/2+\eta} \xi_{1,T}^{-2}) \end{aligned} \tag{29}$$

(ii)

$$\begin{aligned} \sup_{x \in \mathbb{R}^+} \quad & \left| \frac{1}{T} \sum_{t=1}^{T-1} \left(\mathbf{K}\left(\frac{RM_{t+1,M} - x}{\xi_{2,T}}, \frac{RM_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^2 \right.\right. \\ & \left.\left. - \mathbf{K}\left(\frac{IV_{t+1} - x}{\xi_{2,T}}, \frac{IV_t - RM_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^2 \right) \right| = O_P(b_M^{-1/2+\eta} \xi_{2,T}^{-3}) \end{aligned} \tag{30}$$

8.1 Proof of Lemma 1

We prove (ii), as (i) follows by the same argument. By mean value expansion around IV_t,

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^{T-1} \left(\mathbf{K}\left(\frac{RM_{t+1,M} - x}{\xi_{2,T}}, \frac{RM_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^2 - \mathbf{K}\left(\frac{IV_{t+1} - x}{\xi_{2,T}}, \frac{IV_t - RV_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^2 \right) \\ = \;& \frac{1}{T} \sum_{t=1}^{T-1} \left(\mathbf{K}_1'\left(\frac{\widetilde{RM}_{t+1,M} - x}{\xi_{2,T}}, \frac{\widetilde{RM}_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^3 \right) N_{t+1,M} \\ & + \frac{1}{T} \sum_{t=1}^{T-1} \left(\mathbf{K}_2'\left(\frac{\widetilde{RM}_{t+1,M} - x}{\xi_{2,T}}, \frac{\widetilde{RM}_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^2 \right) N_{t,M} \end{aligned} \tag{31}$$

where $\widetilde{RM}_{t+1,M} \in (RM_{t+1,M}, IV_{t+1})$, $\mathbf{K}_j'$ denotes the first derivative with respect to the $j-$th argument of $\mathbf{K}$. We begin by showing that the first term on the RHS of (31) is $o_P(1)$ uniformly in x. Now,

$$\begin{aligned} & \sup_{x \in \mathbb{R}^+} \left| \frac{1}{T} \sum_{t=1}^{T-1} \left(\mathbf{K}_1'\left(\frac{\widetilde{RM}_{t+1,M} - x}{\xi_{2,T}}, \frac{\widetilde{RM}_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) / \xi_{2,T}^3 \right) N_{t+1,M} \right| \\ \leq \;& \sup_{x \in \mathbb{R}^+} \frac{1}{T} \sum_{t=1}^{T-1} \left| \mathbf{K}_1'\left(\frac{\widetilde{RM}_{t+1,M} - x}{\xi_{2,T}}, \frac{\widetilde{RM}_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) \right| \frac{1}{\xi_{2,T}^3} |N_{t+1,M}| \\ \leq \;& \sup_{x \in \mathbb{R}^+} \left| \mathbf{K}_1'\left(\frac{\widetilde{RM}_{t+1,M} - x}{\xi_{2,T}}, \frac{\widetilde{RM}_{t,M} - RM_{T,M}}{\xi_{2,T}} \right) \right| \frac{1}{T \xi_{2,T}^3} \sum_{t=1}^{T-1} |N_{t+1,M}| \end{aligned}$$

$$= \quad O_P(1) O_P\left(b_M^{-1/2}\xi_{2,T}^{-3}\right).$$

In fact $\sup_{x\in\mathbb{R}^+}\left|\mathbf{K}_1'\left(\frac{\widetilde{RM}_{t+1,M}-x}{\xi_{2,T}},\frac{\widetilde{RM}_{t,M}-RM_{T,M}}{\xi_{2,T}}\right)\right| = O_P(1)$, given A4(i),

$$\begin{aligned}\frac{1}{T\xi_{2,T}^3}\sum_{t=1}^{T-1}|N_{t+1,M}| &= \frac{1}{T\xi_{2,T}^3}\sum_{t=1}^{T-1}\mathrm{E}\,|N_{t+1,M}| + \frac{1}{T\xi_{2,T}^3}\sum_{t=1}^{T-1}\left(|N_{t+1,M}|-\mathrm{E}\,|N_{t+1,M}|\right)\\ &= O_p\left(b_M^{-1/2}\right)+\frac{1}{T\xi_{2,T}^3}\sum_{t=1}^{T-1}\left(|N_{t+1,M}|-\mathrm{E}\,|N_{t+1,M}|\right).\end{aligned}$$

Now,

$$\begin{aligned}&\mathrm{var}\left(\frac{b_M^{1/2-\eta}}{T}\sum_{t=1}^{T-1}\left(|N_{t+1,M}|-\mathrm{E}\,|N_{t+1,M}|\right)\right)\\ =\ &\frac{b_M^{1-2\eta}}{T^2}\sum_{t=1}^{T-1}\sum_{s=1}^{T-1}\mathrm{E}\left(\left(|N_{t+1,M}|-\mathrm{E}\,|N_{t+1,M}|\right)\left(|N_{s+1,M}|-\mathrm{E}\,|N_{t-s+1,M}|\right)\right)\\ \leq\ & b_M^{1-2\eta}\left(\mathrm{E}\left(\left(|N_{t+1,M}|-\mathrm{E}\,|N_{t+1,M}|\right)\right)^2\right)^{1/2}\left(\mathrm{E}\left(\left(|N_{t+1,M}|-\mathrm{E}\,|N_{t+1,M}|\right)\right)^2\right)^{1/2} = O(b_M^{-2\eta}),\end{aligned}$$

given A1(ii) and A1(v).

(i) can be shown in an analogous way, simply replacing $\xi_{2,T}^{-2}$ with $\xi_{1,T}^{-1}$. ■

Lemma 2. *Let assumption A(2)-A(4). Then, if* $\xi_{2,T}^3/\xi_{1,T}^2 \to 0$, *as* $T\to\infty$,

(i)

$$\sup_{x\in\mathbb{R}^+}\left|\frac{1}{T}\sum_{t=1}^{T-1}\left(K\left(\frac{IV_{t+1}-x}{\xi_{1,T}}\right)/\xi_{1,T} - f_{IV_{t+1}}(x)\right)\right| = O_P\left(T^{-1/2}\xi_{1,T}^{-1}\right)+O_P\left(\xi_{1,T}^2\right)$$

(ii)

$$\begin{aligned}&\sup_{x\in\mathbb{R}^+}\left|\frac{1}{T}\sum_{t=1}^{T-1}\left(\mathbf{K}\left(\frac{IV_{t+1}-x}{\xi_{1,T}},\frac{IV_{t+1}-RM_{T,M}}{\xi_{1,T}}\right)/\xi_{1,T} - f_{IV_{t+1},IV_t}(x,RM_{T,M})\right)\right|\\ =\ & O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O_P\left(\xi_{2,T}\right).\end{aligned}$$

8.2 Proof of Lemma 2

It follows from Theorem 1 in Andrews (1995), setting, in his notation, $\omega = 2$, $\lambda = 0$, $\eta = \infty$, $\sigma_{1T} = \sigma_{2T}$. In fact, given A2-A3, IV_t has an ARMA structure, and so is geometrically strong mixing, thus NP1 in Andrews holds with $\eta = \infty$, and $a(s)$ decaying at a geometric rate. Also, A4 implies that NP2 and NP4 in Andrews are satisfied. ■

8.3 Proof of Theorem 1

Define:

$$\widehat{f}_{1,T,M}(RM_{T,M}) = \frac{1}{T\xi_{1,T}} \sum_{t=1}^{T-1} K\left(\frac{RM_{t,M} - RM_{T,M}}{\xi_{1,T}}\right)$$

$$\widehat{f}_{2,T,M}(x, RM_{T,M}) = \frac{1}{T\xi_{2,T}^2} \sum_{t=1}^{T-1} \mathbf{K}\left(\frac{RM_{t+1,M} - x}{\xi_{2,T}}, \frac{IV_t - RM_{T,M}}{\xi_{2,T}}\right)$$

$$\widetilde{f}_{1,T}(RM_{T,M}) = \frac{1}{T\xi_{1,T}} \sum_{t=1}^{T-1} K\left(\frac{IV_t - RM_{T,M}}{\xi_{1,T}}\right)$$

$$\widetilde{f}_{2,T}(x, RM_{T,M}) = \frac{1}{T\xi_{2,T}^2} \sum_{t=1}^{T-1} \mathbf{K}\left(\frac{IV_{t+1} - x}{\xi_{2,T}}, \frac{RM_{t,M} - RM_{T,M}}{\xi_{2,T}}\right)$$

$$f_{IV_{t+1}|IV_t}(x|RM_{T,M}) = \frac{f_{IV_{t+1},IV_t}(x, RM_{T,M})}{f_{IV_t}(RM_{T,M})}$$

Now,

$$\begin{aligned}
& \frac{\widehat{f}_{2,T,M}(x, RM_{T,M})}{\widehat{f}_{1,T,M}(RM_{T,M})} - \frac{f_{IV_{t+1},IV_t}(x, RM_{T,M})}{f_{IV_t}(RM_{T,M})} \\
= \quad & \left(\frac{\widehat{f}_{2,T,M}(x, RM_{T,M}) - f_{IV_{t+1},IV_t}(x, RM_{T,M})}{f_{IV_t}(RM_{T,M})}\right) \\
& + f_{IV_{t+1},IV_t}(x, RM_{T,M}) \left(\frac{\widehat{f}_{1,T,M}(RM_{T,M}) - f_{IV_t}(RM_{T,M})}{\widehat{f}_{1,T,M}(RM_{T,M}) f_{IV_t}(RM_{T,M})}\right) \\
& + \frac{\left(\widehat{f}_{2,T,M}(x, RM_{T,M}) - f_{IV_{t+1},IV_t}(x, RM_{T,M})\right)\left(\widehat{f}_{1,T,M}(RM_{T,M}) - f_{IV_t}(RM_{T,M})\right)}{\widehat{f}_{1,T,M}(RM_{T,M}) f_{IV_t}(RM_{T,M})}
\end{aligned}$$

Thus,

$$\begin{aligned}
& \sup_{x\in\mathbb{R}^+} \left|\frac{\widehat{f}_{2,T,M}(x, RM_{T,M})}{\widehat{f}_{1,T,M}(RM_{T,M})} - \frac{f_{IV_{t+1},IV_t}(x, RM_{T,M})}{f_{IV_t}(RM_{T,M})}\right| \\
\leq \quad & \sup_{x\in\mathbb{R}^+} \left|\frac{\widehat{f}_{2,T,M}(x, RM_{T,M}) - f_{IV_{t+1},IV_t}(x, RM_{T,M})}{f_{IV_t}(RM_{T,M})}\right| \\
& + \sup_{x\in\mathbb{R}^+} \left|f_{IV_{t+1},IV_t}(x, RM_{T,M}) \left(\frac{\widehat{f}_{1,T,M}(RM_{T,M}) - f_{IV_t}(RM_{T,M})}{\widehat{f}_{1,T,M}(RM_{T,M}) f_{IV_t}(RM_{T,M})}\right)\right| \\
& + \frac{\left|\widehat{f}_{1,T,M}(RM_{T,M}) - f_{IV_t}(RM_{T,M})\right|}{\widehat{f}_{1,T,M}(RM_{T,M}) f_{IV_t}(RM_{T,M})} \sup_{x\in\mathbb{R}^+} \left|\widehat{f}_{2,T,M}(x, RM_{T,M}) - f_{IV_{t+1},IV_t}(x, RM_{T,M})\right| \quad (32)
\end{aligned}$$

Given that $f_{IV_t}(RM_{T,M})$ is bounded away from zero, as for the first term on the RHS of (32), note that

$$\sup_{x\in\mathbb{R}^+} \left|\widehat{f}_{2,T,M}(x, RM_{T,M}) - f_{IV_{t+1},IV_t}(x, RM_{T,M})\right|$$

$$
\begin{aligned}
&\le \sup_{x\in\mathbb{R}^+}\left|\widehat{f}_{2,T,M}(x,RM_{T,M})-\widetilde{f}_{2,T}(x,RM_{T,M})\right|+\sup_{x\in\mathbb{R}^+}\left|\widetilde{f}_{2,T}(x,RM_{T,M})-f_{IV_{t+1},IV_t}\left(x,RM_{T,M}\right)\right| \\
&= O_P(b_M^{-1/2+\eta}\xi_{2,T}^{-3})+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O_P\left(\xi_{2,T}\right),
\end{aligned}
$$

by Lemma 1(ii) and 2(ii). As for the second on the RHS of (32), given A4(ii)

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}^+}\left|f_{IV_{t+1},IV_t}\left(x,RM_{T,M}\right)\left(\frac{\widehat{f}_{1,T,M}(RM_{T,M})-f_{IV_t}\left(RM_{T,M}\right)}{\widehat{f}_{1,T,M}(RM_{T,M})f_{IV_t}\left(RM_{T,M}\right)}\right)\right| \\
&\le C\left(\left|\widehat{f}_{1,T,M}(RM_{T,M})-\widetilde{f}_{1,T}(RM_{T,M})\right|+\left|\widetilde{f}_{1,T}(RM_{T,M})-f_{IV_t}\left(RM_{T,M}\right)\right|\right) \\
&= O_P(b_M^{-1/2}\xi_{1,T}^{-2})+O_P\left(T^{-1/2}\xi_{1,T}^{-1}\right)+O_P\left(\xi_{1,T}^{2}\right),
\end{aligned}
$$

given Lemma 1(i) and Lemma 2(i). Finally, it is immediate to see that the last term on the RHS of (32) is of a smaller order than the previous two. Thus, for $\xi_{2,T}^3/\xi_{1,T}^2\to 0$,

$$
\sup_{x\in\mathbb{R}^+}\left|\frac{\widehat{f}_{2,T,M}(x,RM_{T,M})}{\widehat{f}_{1,T,M}(RM_{T,M})}-\frac{f_{IV_{t+1},IV_t}\left(x,RM_{T,M}\right)}{f_{IV_t}\left(RM_{T,M}\right)}\right|=O_P(b_M^{-1/2+\eta}\xi_{2,T}^{-3})+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O_P\left(\xi_{2,T}\right)
$$

Thus,

$$
\left|\widehat{F}_{RM_{T+1}|RM_T}(u|RM_{T,M})-F_{IV_{T+1}|IV_T}(u|RM_{T,M})\right|=O_P\left(b_M^{-1/2+\eta}\xi_{2,T}^{-3}\right)+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O_P\left(\xi_{2T}\right)
$$

Finally,

$$
\begin{aligned}
&\left|\widehat{F}_{RM_{T+1}|RM_T}(u|RM_{T,M})-F_{IV_{T+1}|IV_T}(u|IV_T)\right| \\
&\le \left|\widehat{F}_{RM_{T+1}|RM_T}(u|RM_{T,M})-F_{IV_{T+1}|IV_T}(u|RM_{T,M})\right|+\left|F_{IV_{T+1}|IV_T}(u|IV_T)-F_{IV_{T+1}|IV_T}(u|RM_{T,M})\right| \\
&= O_P\left(b_M^{-1/2+\eta}\xi_{2,T}^{-3}\right)+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O_P\left(\xi_{2T}\right)+O_P(b_M^{-1/2}) \\
&= O_P\left(b_M^{-1/2+\eta}\xi_{2,T}^{-3}\right)+O_P\left(T^{-1/2}\xi_{2,T}^{-2}\right)+O_P\left(\xi_{2T}\right),
\end{aligned}
$$

noting

$$
\left|F_{IV_{T+1}|IV_T}(u|IV_T)-F_{IV_{T+1}|IV_T}(u|RM_{T,M})\right|=\left|f_{IV_{t+1}|IV_t}(u|\widetilde{RM}_{t,M})\left(IV_T-RM_{T,M}\right)\right|=O_P(b_M^{-1/2+\eta}).
$$

The statement in the theorem then follows. ■

The proof of Theorem 2 requires the following Lemma.

Lemma 3. *Let A1-A7, if as* $M,T,S,N\to\infty$, $T/b_M^2\to 0$, $T/N^{(1-\delta)}\to 0$, $T/S\to 0$, $p_T\to\infty$ *and* $p_T/T^{1/4}\to 0$, *then for any* $u\in U$,

(i)

$$
\begin{aligned}
&\left|\frac{1}{S}\sum_{i=1}^{S}\left(K\left(\frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{1,T}}\right)/\zeta_{1,T}-K\left(\frac{IV_{i,1}(\boldsymbol{\theta}^{\dagger})-RM_{T,M}}{\zeta_{1,T}}\right)/\zeta_{1,T}\right)\right| \\
&= O_P(T^{-1/2}\zeta_{1,T}^{-2}),
\end{aligned}
$$

(ii)

$$\sup_{x\in\mathbb{R}^+}\frac{1}{S}\sum_{i=1}^{S-1}\left|\left(\mathbf{K}\left(\frac{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right.\right.$$
$$\left.\left.-\mathbf{K}\left(\frac{IV_{i,2}(\boldsymbol{\theta}^\dagger)-x}{\zeta_{2,T}},\frac{IV_{i,1}(\boldsymbol{\theta}^\dagger)-RV_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right)\right|=O_P(T^{-1/2}\zeta_{2,T}^{-3})$$

8.4 Proof of Lemma 3

We prove (ii), as (i) would follow by the same argument.

$$\begin{aligned}
&\frac{1}{S}\sum_{i=1}^{S}\left(\mathbf{K}\left(\frac{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right.\\
&\left.-\,\mathbf{K}\left(\frac{IV_{i,2}(\boldsymbol{\theta}^\dagger)-x}{\zeta_{2,T}},\frac{IV_{i,1}(\boldsymbol{\theta}^\dagger)-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right)\\
=\;&\frac{1}{S}\sum_{i=1}^{S}\left(\mathbf{K}\left(\frac{IV_{i,2,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right.\\
&\left.-\mathbf{K}\left(\frac{IV_{i,2,N}\left(\boldsymbol{\theta}^\dagger\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\boldsymbol{\theta}^\dagger\right)-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right)\\
&+\frac{1}{S}\sum_{i=1}^{S}\left(\mathbf{K}\left(\frac{IV_{i,2,N}\left(\boldsymbol{\theta}^\dagger\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\boldsymbol{\theta}^\dagger\right)-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right.\\
&\left.-\mathbf{K}\left(\frac{IV_{i,2}-x}{\zeta_{2,T}},\frac{IV_{i,1}-RM_{T,M}}{\zeta_{2,T}}\right)/\zeta_{2,T}^2\right) \qquad (33)
\end{aligned}$$

Via a mean value expansion, and taking the supremum over x, the first term on the RHS of (33) writes as:

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^+}\left|\frac{1}{S\zeta_{2,T}^3}\sum_{i=1}^{S}\mathbf{K}'\left(\frac{IV_{i,2,N}\left(\overline{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\overline{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{2,T}}\right)\frac{\partial IV_{i,2,N}\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}}\Big|_{\overline{\boldsymbol{\theta}}_{T,S,M,N}}\right.\\
&\left.\times\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}-\boldsymbol{\theta}^\dagger\right)\right|\\
\leq\;&\sup_{x\in\mathbb{R}^+}\frac{1}{S}\sum_{i=1}^{S}\left|\mathbf{K}'\left(\frac{IV_{i,2,N}\left(\overline{\boldsymbol{\theta}}_{T,S,M,N}\right)-x}{\zeta_{2,T}},\frac{IV_{i,1,N}\left(\overline{\boldsymbol{\theta}}_{T,S,M,N}\right)-RM_{T,M}}{\zeta_{2,T}}\right)\frac{\partial IV_{i,2,N}\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}}\Big|_{\overline{\boldsymbol{\theta}}_{T,S,M,N}}\right|\\
&\times\frac{1}{\zeta_{2,T}^3}\left|\widehat{\boldsymbol{\theta}}_{T,S,M,N}-\boldsymbol{\theta}^\dagger\right|\\
=\;&O_P(1)O_P\left(\zeta_{2,T}^3T^{-1/2}\right),
\end{aligned}$$

given that, from Corradi and Distaso (piece of proof of Thm 2), $\left(\widehat{\boldsymbol{\theta}}_{T,S,M,N}-\boldsymbol{\theta}^\dagger\right)=O_P\left(T^{-1/2}\right)$. Finally, the last term on the RHS of (33) is of a smaller order, as N grows faster than T. (DETAILS MISSING) ■

8.5 Proof of Theorem 2

Given Lemma 3, by the same argument used in the proof of Theorem 1. ■

8.6 Proof of Propositions 1,2 and 3

A1(i)-A1(iv) follow straightforwardly from the proofs of Propositions 1,2 and 3 in Corradi and Distaso (2004). It remains to show A1(v), which follows by noting that $|N_{t,M}| = O_P(b_M^{-1/2})$, (to be completed) ■

8.7 Proof of equation (27)

To be done

References

Aït-Sahalia, Y., P.A. Mykland and L. Zhang (2003). How Often to Sample a Continuous Time Process in the Presence of Market Microstructure Noise. *Review of Financial Studies*, forthcoming.

Andersen, T.G., T. Bollerslev, (1997), Intraday Periodicity and Volatility Persistence in Financial Markets, *Journal of Empirical Finance,* 4, 115-158

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, (2001), The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association, 96, 42-55.*

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, (2003), Modelling and Forecasting Realized Volatility, Econometrica, 71, 579-626

Andersen, T.G., T. Bollerslev and N. Meddahi, (2004), Analytic Evaluation of Volatility Forecasts. *International Economic Review*, 45, 1079-1110.

Andersen, T.G., T. Bollerslev and N. Meddahi, (2004b), Correcting the Errors: Volatility Forecast Evaluation Using High Frequency Data and Realized Volatilities, *Econometrica*, forthcoming.

Andrews, D.W.K, (1990), Asymptotics for Semiparametric Econometric Models: II. Stochastic Equicontinuity and Nonparametric Kernel Estimation, Cowles Foundation Discussion Paper 909R.

Andrews, D.W.K., (1995), Nonparametric Kernel Estimation for Semiparametric Models, *Econometric Theory,* 560-596.

Awartani, B., V. Corradi and W. Distaso (2004). Testing and Modelling Market Microstructure Effects with an Application to the Dow Jones Industrial Average, FERC Working Paper, WP04-06

http://www2.warwick.ac.uk/fac/soc/wbs/research/wfri/ferc/wrkingpaprseries/

Barndorff-Nielsen, O.E., and N. Shephard, (2004a), Power and Bipower Variation with Stochastic Volatility and Jumps, *Journal of Financial Econometrics.*

Barndorff-Nielsen, O.E., and N. Shephard, (2004b), Econometrics of Testing for Jumps in Financial Economics using Bipower Variation, University of Oxford.

Barndorff-Nielsen, O.E., and N. Shephard, (2004c), A Feasible Central Limit Theory for Realized Volatility under Leverage, University of Oxford.

Barndorff-Nielsen, O.E., and N. Shephard, (2002), Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility, *Journal of the Royal Statistical Society, Ser. B.,* 64, 243-280.

Bierens, H.J., (1983), Uniform Consistency of Kernel Estimators of a Regression Function under Generalized Conditions, *Journal of the American Statistical Association,* 77, 699-707.

Bollerslev, T. and H. Zhou, (2002), Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility, *Journal of Econometrics,* 109, 33-65.

Bosq, D. (1998), *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction,* Lecture Notes in Statistics #110, Springer and Verlag, New York.

Corradi, V. and W. Distaso (2004). Estimating and Testing Stochastic Volatility Models via Realized Measures, FERC Working Paper, WP04-12, http://www2.warwick.ac.uk/fac/soc/wbs/research/wfri

Hansen, P.R., and A. Lunde, (2004), An Unbiased Version of Realized Variance, Working Paper, Stanford University.

Karatzas, I. and Shreve, S.E., (1988), *Brownian Motion and Stochastic Calculus,* Springer and Verlag, New York.

Meddahi, N. (2001). An Eigenfunction Approach for Volatility Modeling, Manuscript, University of Montreal.

Meddahi, N. (2002). A Theoretical Comparison between Integrated and Realized Volatility. *Journal of Applied Econometrics*, 17, 475-508.

Meddahi, N. (2003). ARMA Representation of Integrated and Realized Variances. *Econometrics Journal*, 6, 334–355.

Zhang, L., P.A. Mykland and Y. Aït-Sahalia, (2004), A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High Frequency Data, Carnegie Mellon.