Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities*

Torben G. Andersen†, Tim Bollerslev‡, and Nour Meddahi§

First Draft: September 2002
This Draft: November 5, 2003

Abstract

We develop general model-free adjustment procedures for the calculation of unbiased volatility loss functions based on practically feasible realized volatility benchmarks. The procedures, which exploit the recent non-parametric asymptotic distributional results in Barndorff-Nielsen and Shephard (2002a, 2003a) along with new results explicitly allowing for leverage effects, are both easy-to-implement and highly accurate in empirically realistic situations. On properly accounting for the measurement errors in the volatility forecast evaluations reported in Andersen, Bollerslev, Diebold and Labys (2003), the adjustments result in markedly higher estimates for the true degree of return-volatility predictability.

*This work was supported by a grant from the National Science Foundation to the NBER (Andersen and Bollerslev), and from FQRSC, IFM2, MITACS, NSERC, SSHRC, and Jean-Marie Dufour's Econometrics Chair of Canada (Meddahi). Some of this material was circulated earlier as part of the paper "Analytical Evaluation of Volatility Forecasts." Detailed comments by Neil Shephard, Tom McCurdy, two anonymous referees, and the editor (Costas Meghir) have importantly improved the paper. We would also like to thank Francis X. Diebold for many discussions on closely related ideas, as well as seminar participants at the 2002 NBER/NSF Time Series Conference and the 2003 CIRANO-CIREQ Financial Econometrics Conference. The third author thanks the Bendheim Center for Finance at Princeton University for its hospitality during his visit where part of the research was done.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208, and NBER, USA, phone: 847-467-1285, e-mail: t-andersen@kellogg.northwestern.edu.

‡Department of Economics, Duke University, Durham, NC 27708, and NBER, USA, phone: 919-660-1846, e-mail: boller@econ.duke.edu.

§Département de sciences économiques, CIRANO, CIREQ, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), H3C 3J7, Canada, phone: 514-343-2399, e-mail: nour.meddahi@umontreal.ca; web-page: http://www.mapageweb.umontreal.ca/meddahin/.

"ARCH models have become indispensable tools not only for researchers, but also for analysts on financial markets, who use them in asset pricing and in evaluating portfolio risk." Press release, October 8, 2003: The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel.

1 Introduction

The burgeoning literature on time-varying financial market volatility is abound with empirical studies in which competing models are evaluated and compared on the basis of their forecast performance. Contrary to the typical setting for economic forecast evaluation, the variable of interest in that context - the volatility - is not directly observable but rather inherently latent. Consequently, any ex-post assessment of forecast precision must contend with a fundamental errors-in-variable problem associated with the measurement of the realization of the forecasted variable. Growing recognition of the importance of this issue has led a number of recent studies to advocate the use of so-called realized volatilities, constructed from the summation of finely sampled squared high-frequency returns, as a practical method for improving the ex-post volatility measures.

The use of realized volatility as the practical benchmark may be justified by standard continuous-time arguments. Assuming that the sampling frequency of the squared returns utilized in the realized volatility computations approaches zero, the realized volatility consistently estimates the true (latent) integrated volatility under quite general conditions, where importantly, the latter concept corresponds to the realization of the (cumulative) instantaneous variance process over the relevant horizon (see, e.g., Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold and Labys, 2001; Barndorff-Nielsen and Shephard, 2001, 2002a,b; Comte and Renault, 1998; along with the recent survey by Andersen, Bollerslev and Diebold, 2003). Unfortunately, market microstructure frictions distort the measurement of returns at the highest frequencies so that, e.g., tick-by-tick return processes blatantly violate the theoretical semi-martingale restrictions implied by the no-arbitrage assumptions in continuous-time asset pricing models. These same features also bias empirical realized volatility measures constructed directly from the ultra high-frequency returns, so in practice the measures are instead typically constructed from intraday returns sampled at an intermediate frequency.[1] As such, the integrated volatility is invariably measured with error (see, e.g., the numerical calculations in Andreou and Ghysels, 2002, and Bai, Russell, and Tiao, 2000). The exact form of the measurement error will, of course, depend on the assumed model structure (see, e.g., Meddahi, 2002, and Barndorff-Nielsen and Shephard, 2002a), but it will generally result in a downward bias in the estimated degree of predictability obtained

[1] For instance, the daily realized volatilities in Andersen, Bollerslev, Diebold and Labys (2003) (henceforth ABDL (2003)) discussed further below are based on the summation of squared half-hourly foreign exchange rate returns, but either 5-minute or 15-minute returns are other common choices in the literature.

through any forecast evaluation criterion that simply uses the realized volatility in place of the true (latent) integrated volatility. Although this bias may be large (Andersen and Bollerslev, 1998), it is almost always ignored in empirical applications.

This note addresses that issue by developing general model-free adjustment procedures that allow for the calculation of simple unbiased loss functions in realistic forecast situations. Moreover, the adjustments are simple to implement in practice. The derivation exploits the recent asymptotic (for increasing sampling frequency) distributional results in Barndorff-Nielsen and Shephard (2002a). While these results explicitly rules out so-called leverage effects, we show that the same approximate adjustment procedures apply in the context of the general eigenfunction stochastic volatility class of models pioneered by Meddahi (2001) explicitly allowing for non-zero contemporaneous correlations between the separate shocks in the return and volatility processes. Following Andersen and Bollerslev (1998) and ABDL (2003), we focus our forecast comparisons on the value of the coefficient of multiple correlation, or R^2, in the Mincer-Zarnowitz style regressions of the ex-post realized volatility on the corresponding model forecasts,[2] but our procedures are general and could be applied in the adjustment of other loss functions used in the evaluation of any arbitrary set of volatility forecasts. On applying the procedures in the context of ABDL (2003), we obtain markedly higher estimates for the true degree of return-volatility predictability, with the adjusted R^2's exceeding their unadjusted counterparts by up to forty-percent.

We proceed as follows. The first subsection below introduces the notions of integrated and realized volatility, along with the (feasible) asymptotic distribution theory due to Barndorff-Nielsen and Shephard (2002a). The development of the practical and easy-to-implement adjustment procedures based on this theory is then presented in the next subsection, followed by our new theoretical results explicitly allowing for leverage effects. Utilizing these results, the last section provides a reassessment of the empirical evidence in ABDL (2003) related to the fit of the Mincer-Zarnowitz style volatility regressions. The accuracy of the asymptotic approximations - which form the basis for our approach - is confirmed through Monte Carlo simulations for models calibrated to reflect empirically relevant and challenging specifications. The details of the simulations and all of the technical proofs are deferred to two Appendixes.

2 Theory

We focus on a single asset traded in a liquid financial market. Assuming that the sample-path of the logarithmic price process, $\{\log(S_t), 0 \leq t\}$, is continuous, the class of continuous-time stochastic volatility models employed in the finance literature is then conveniently expressed

[2] This particular loss function is directly inspired by the work of Mincer and Zarnowitz (1969), and we will refer to the corresponding regressions as such; see also the discussion in Chong and Hendry (1986).

in terms of the following generic stochastic differential equation (sde),

$$d\log(S_t) \;=\; \mu_t dt \;+\; \sigma_t dW_t \;, \tag{1}$$

where W_t denotes a standard Brownian motion, the drift term μ_t is (locally) predictable and of finite variation, and σ_t is a càd-làg process such that $\int_0^t \sigma_u^2 du < \infty$ a.s. for any $t > 0$. Consequently, $\sigma_t dW_t$ is a local martingale and $\log(S_t)$ a semi-martingale (see, for instance, Protter, 1995).

2.1 Integrated and Realized Volatility

Although the sde in equation (1) is very convenient from a theoretical arbitrage-free pricing perspective, practical return calculations and volatility measurements are invariably restricted to discrete time intervals. In particular, focusing on the unit time interval, the one-period continuously compounded return for the price process in equation (1) is formally given by,[3]

$$r_t \;\equiv\; \log(S_t) - \log(S_{t-1}) \;=\; \int_{t-1}^{t} \mu_u du \;+\; \int_{t-1}^{t} \sigma_u dW_u \;, \tag{2}$$

with the corresponding *integrated volatility*,

$$IV_t \;\equiv\; \int_{t-1}^{t} \sigma_u^2 du \;, \tag{3}$$

affording a natural measure of the inherent, or notional, return variability (see, e.g., Andersen, Bollerslev and Diebold, 2003, for further discussion of the integrated and notional volatility concepts).[4]

Of course, integrated volatility is not directly observable. However, by the theory of quadratic variation (see, e.g., Protter, 1995, for a general discussion), the corresponding *realized volatility* defined by the summation of the $1/h$ intra-period squared returns, $r_t^{(h)} \equiv \log(S_t) - \log(S_{t-h})$,

$$RV_t(h) \;\equiv\; \sum_{i=1}^{1/h} r_{t-1+ih}^{(h)2}, \tag{4}$$

where $1/h$ is assumed to be an integer, converges uniformly in probability to IV_t as $h \to 0$.

The consistency of the realized volatility rely on the (conceptual) idea of an ever increasing number of finer sampled high-frequency returns, or $h \to 0$. However, as previously noted, the requisite semi-martingale property of returns invariably breaks down at ultra-high

[3]For notational simplicity, we focus our discussion on one-period return and volatility measures, but the general results and associated measurement error adjustment extend in a straightforward manner to the multi-period case.

[4]The integrated volatility also plays a crucial role in the pricing of options; see, e.g., Garcia, Ghysels, and Renault (2003).

frequencies, so that in actual applications market microstructure frictions in effect put a limit on the number of return observations per unit time interval that may be used productively in the computation of the realized volatility measures. As such, the realized volatility will necessarily be subject to a finite-sample (non-zero h) measurement error vis-a-vis the true (latent) integrated volatility, say

$$U_t(h) \equiv RV_t(h) - IV_t \ . \tag{5}$$

This observation was the original motivation for the development of the Barndorff-Nielsen and Shephard (2002a) asymptotic theory, which gives us a tool with which to study the errors for finite h.

Specifically, assuming that the mean, $\{\mu_u, u \geq 0\}$, and volatility, $\{\sigma_u, u \geq 0\}$, processes are jointly independent of $\{W_u, u \geq 0\}$, it follows from Barndorff-Nielsen and Shephard (2002a, 2003b) that,[5]

$$z_t \equiv \sqrt{h^{-1}} \ \frac{U_t(h)}{\sqrt{2IQ_t}} \overset{law}{\rightarrow} \mathcal{N}(0,1), \tag{6}$$

where the *integrated quarticity*, IQ_t, is defined by

$$IQ_t \ \equiv \ \int_{t-1}^{t} \sigma_u^4 du \ . \tag{7}$$

Moreover, under these same assumptions the integrated quarticity may be consistently estimated by the (standardized) *realized quarticity*

$$RQ_t(h) \ \equiv \ \frac{1}{h}\,\frac{1}{3}\sum_{i=1}^{1/h} r_{t-1+ih}^{(h)4} \ . \tag{8}$$

This remarkable set of asymptotic results allows for general model-free approximations to the distribution of the realized volatility error. In particular no parametric or finite moment assumptions are made on the $\{\mu_u, u \geq 0\}$ and $\{\sigma_u, u \geq 0\}$ processes to derive this result.

Importantly, however, the volatility process is assumed to be independent of the $\{W_u, u \geq 0\}$ process. This assumption formally rules out so-called leverage effects corresponding to a negative contemporaneous correlation between the volatility and the Brownian innovation process. While that is a reasonable assumption for the exchange rate data in ABDL (2003) discussed below, equity index returns, in particular, are often found to be negatively correlated with future volatility (e.g., Black, 1976; and Nelson, 1991). Nonetheless, the simulations reported in Barndorff-Nielsen and Shephard (2003b) suggests that the same

[5]The proof in Barndorff-Nielsen and Shephard (2003b) also requires that the volatility process, $\{\sigma_u, u \geq 0\}$, is (pathwise) locally bounded away from 0, with the property that, $\lim_{h\to 0} h^{1/2}\sum_{i=1}^{1/h} \mid \sigma_{\eta_i}^r - \sigma_{\xi_i}^r \mid = 0$, for some $r > 0$ and any two sequences $\eta_i \equiv \eta_i(h)$ and $\xi_i \equiv \xi_i(h)$ satisfying $t-1 \leq \eta_1 \leq \xi_1 \leq t-1+h \leq \eta_2 \leq \xi_2 \leq t-1+2h \leq ... \leq \eta_{1/h} \leq \xi_{1/h} \leq t$. Moreover, the mean process $\{\mu_u, u \geq 0\}$ must satisfy (pathwise) the condition that $\limsup_{h\to 0} \max_{1\leq i\leq 1/h} h^{-1} \mid \mu_{t-1+ih} - \mu_{t-1+(i-1)h} \mid < \infty$.

(approximate) arguments underlying the Central Limit Theorem in (6) carry over to the case of a non-zero correlation. We further support this contention through new theoretical results along with additional simulation based evidence pertaining directly to the measurement error adjustment procedures developed in the next section.

2.2 Practical Measurement Error Adjustments

The results discussed in the previous section implies that the time $t+1$ realized volatility error is (approximately) serially uncorrelated and orthogonal to any variables (volatility forecasts) in the time t information set. This justifies the common use of realized volatility as a convenient simple and unbiased, albeit potentially noisy, benchmark in ex-post volatility forecast evaluations and model comparisons.

Specifically, consider the Mincer-Zarnowitz style regressions of the realized volatility on a set of predetermined regressors (volatility forecasts) employed in ABDL (2003) among others.[6] Assuming that the underlying continuous time process satisfies a weak uniform integrability condition so that the consistency of $RQ_t(h)$ for IQ_t also guarantees convergence in mean (see, e.g., Billingsley, 1994, and Hoffmann-Jørgensen, 1994), it follows directly from equation (6), that for small values of h,

$$Var[IV_t] \;=\; Var[RV_t(h)] \;-\; 2hE[RQ_t(h)] \;+\; o(h). \tag{9}$$

Thus, any MSE type forecast evaluation criteria based on a comparison of the volatility forecasts with the ex-post $RV_t(h)$ in place of IV_t will on average overstate the true variability of the forecast errors by $2hE[RQ_t(h)]$. In particular, ignoring the $o(h)$ term, it follows that the (feasible) R^2 from the commonly employed Mincer-Zarnowitz regression will underestimate the true predictability as measured by the (infeasible) R^2 from the regression of the future (latent) integrated volatility on the same set of predetermined regressors (volatility forecasts) by the multiplicative factor: $Var[RV_t(h)]/\{Var[RV_t(h)] - 2hE[RQ_t(h)]\}$.[7]

Meanwhile, the predictive regressions and related loss functions reported in the extant volatility literature are often formulated in terms of the realized standard deviation, $RV_t(h)^{1/2}$, or the logarithmic standard deviation, $\log RV_t(h)^{1/2}$. To properly gauge the true predictability in those situations the sample variances of the transformed realized

[6] Although this is not required for the Barndorff-Nielsen and Shephard (2002a, 2003b) asymptotic theory discussed in the previous section, the Mincer-Zarnowitz regression implicitly assumes that the variable of interest, i.e., the integrated and realized volatility processes, have finite second order moments. This in turn requires that the fourth moment of σ_t is finite. This holds for any affine and log-normal diffusion, and is also satisfied for the GARCH diffusions considered in the Monte Carlo experiment discussed in the Appendix.

[7] As previously noticed by Meddahi (2002), the approximation in (9) also allows for the construction of more efficient (in the sense of MSE) model-free integrated volatility estimates, by downweighting the realized volatility by the multiplicative factor $\{Var[RV_t(h)] - 2hE[RQ_t(h)]\}/Var[RV_t(h)]$ and adding the constant $\{E[RV_t(h)]2hE[RQ_t(h)]\}/Var[RV_t(h)]$.

volatilities may be similarly replaced by (feasible) expressions for the true (latent) variances, $Var[IV_t^{1/2}]$ and $Var[\log IV_t^{1/2}]$, respectively.[8] To this end, it follows from equation (6) and a second-order Taylor series expansion of the square-root function of $RV_t(h)$ around IV_t, that conditional on the sample-path realization of the (latent) point-in-time volatility process (see the Appendix),

$$RV_t(h)^{1/2} \;\approx\; IV_t^{1/2} \;+\; 2^{-1/2}h^{1/2}IV_t^{-1/2}IQ_t^{1/2}z_t \;-\; \frac{1}{4}hIV_t^{-3/2}IQ_tz_t^2 \;. \tag{10}$$

Thus, subject to the necessary uniform integrability conditions on the underlying continuous-time process ensuring convergence in mean of the relevant quantities (see also Barndorff-Nielsen and Shephard, 2003b),

$$Var[IV_t^{1/2}] \;=\; E[RV_t(h)] - \left\{E[RV_t(h)^{1/2}] \;+\; \frac{h}{4}E[RV_t(h)^{-3/2}RQ_t(h)]\right\}^2 \;+\; o(h). \tag{11}$$

The variance of the square-root of the realized volatility, as used in a number of previous empirical studies, obviously exceeds the expression in (11) by the absence of the second (positive) term in the last squared bracket. This in turn will result in a downward bias in the R^2's from the (feasible) Mincer-Zarnowitz predictive regressions formulated in terms of $RV_t(h)^{1/2}$ in place of $IV_t^{1/2}$.

By similar arguments (see the Appendix),[9]

$$\log RV_t(h) \;\approx\; \log IV_t \;+\; 2^{1/2}h^{1/2}IV_t^{-1}IQ_t^{1/2}z_t \;-\; hIV_t^{-2}IQ_tz_t^2 \;, \tag{12}$$

and,

$$\begin{aligned}[\log RV_t(h)]^2 \;\approx\; & [\log IV_t]^2 \;+\; 2^{3/2}h^{1/2}IV_t^{-1}[\log IV_t]IQ_t^{1/2}z_t \;+ \\ & - 2hIV_t^{-2}(1-\log IV_t)IQ_tz_t^2 \;,\end{aligned} \tag{13}$$

so that again subject to the necessary integrability conditions,

$$\begin{aligned}Var[\log IV_t] \;=\; & E[[\log RV_t(h)]^2] \;-\; 2hE[RV_t(h)^{-2}(1-\log RV_t(h))RQ_t(h)] \\ & - \left\{E[\log RV_t(h)] \;+\; hE[RV_t(h)^{-2}RQ_t(h)]\right\}^2 \;+\; o(h).\end{aligned} \tag{14}$$

The accuracy of the distributional assumption and second-order Taylor series expansions underlying the (feasible) expressions for the latent variances in equations (9), (11), and (14)

[8] Any transformed unbiased forecast for IV_{t+1} will generally not be unbiased for $IV_{t+1}^{1/2}$ or $\log IV_{t+1}^{1/2}$. However, allowing for a non-zero intercept and a slope coefficient different from unity in the Mincer-Zarnowitz regression of the future transformed realized volatilities on the transformed forecasts explicitly corrects this (unconditional) bias in the forecasts; see also the discussion in Andersen, Bollerslev and Meddahi (2003).

[9] Interestingly, the Monte Carlo evidence in Barndorff-Nielsen and Shephard (2003b) also suggests that the asymptotic approximation obtained by equating z_t^2 to one in (12), i.e., $[\log RV_t(h) - \log IV_t + hIV_t^{-2}IQ_t]/[2^{1/2}h^{1/2}IV_t^{-1}IQ_t^{1/2}]$, is closer to a standard normal than the approximation obtained by applying the Δ-method directly to (6), i.e., $[\log RV_t(h) - \log IV_t]/[2^{1/2}h^{1/2}IV_t^{-1}IQ_t^{1/2}]$.

are underscored by the simulation results for the baseline models reported in Table A1 of the Appendix.[10] It is evident that the simulated medians and ninety-percent confidence intervals for the asymptotic approximations to $Var[IV_t]$, $Var[IV_t^{1/2}]$ and $Var[\log(IV_t^{1/2})]$ are extremely close to the simulated sampling distributions for the true variances (labelled $h = 1/\infty$) as long as the frequency of the returns used in the calculation of the realized volatility and quarticity measures, $RV_t(h)$ and $RQ_t(h)$, respectively, exceeds half-an-hour, or $h \leq 1/48$.

Similar arguments could, of course, be applied for any other twice continuously differentiable function of integrated volatility in order to obtain an approximate value for $Var[f(IV_t)]$, in turn allowing for simple model-free approximations to the true (infeasible) R^2's that would obtain in the hypothetical regressions of $f(IV_t)$ on any forecasts by scaling the (feasible) R^2's from the corresponding regressions based on $f(RV_t(h))$ by the multiplicative adjustment factor, $Var[f(RV_t(h))]/\{Var[f(IV_t(h))]\}$.

2.3 Leverage Effects

The assumptions underlying the distributional results and adjustment procedures discussed in the previous section formally rule out leverage effects. In order to assess the potential biases in the various volatility measurements caused by a violation of that assumption, this section presents analytical results pertaining to the eigenfunction stochastic volatility (ESV) class of models introduced by Meddahi (2001), explicitly allowing for leverage effects.[11] The ESV class of models is very general, encompassing all of the models most commonly employed in the literature, including the GARCH diffusion model of Nelson (1990), the log-normal diffusion model popularized by Hull and White (1987) and Wiggins (1987), and the square-root diffusion model of Heston (1993), along with multi-factor extensions of all these models.

For expositional purposes, we concentrate on the two-factor ESV model. This remains the empirically most relevant case. However, all of the analytic results could be extended to the case of three or more factors without too many technical difficulties, at the cost of considerable notational complications. The two-factor ESV model is formally defined by the

[10] The accuracy of (6) and the corresponding CLT for $RV_t(h)^{1/2}$ and $\log(RV_t(h))$ based on the Δ-method has also previously been investigated by Barndorff-Nielsen and Shephard (2003b).

[11] The use of eigenfunctions in modelling Markovian time series was pioneered by Chen, Hansen and Scheinkman (2000).

following three equations,

$$d\log(S_t) = \mu dt + \sigma_t\, dW_t = \mu dt + \sigma_t \left[\rho_1\, dW_{1,t} + \rho_2\, dW_{2,t} + \sqrt{1-\rho_1^2-\rho_2^2}\, dW_{3,t}\right], \quad (15)$$

$$\sigma_t^2 = \sum_{0\leq i,j\leq p} a_{i,j}\bar{P}_{1,i}(f_{1,t})\bar{P}_{2,j}(f_{2,t})\ , \quad (16)$$

$$df_{n,t} = m_n(f_{n,t})dt + \sqrt{v_n(f_{n,t})}\, dW_{n,t}\ , \quad n = 1,2, \quad (17)$$

where $W_{1,t}$, $W_{2,t}$ and $W_{3,t}$ denote three independent standard Brownian motions; p denotes an integer (possible infinite); the $a_{i,j}$'s are real numbers; and the $\bar{P}_{n,i}(f_{n,t})$'s denote the eigenfunctions of the infinitesimal generator, $\mathcal{A}_n$, associated with the $n = 1,2$ stationary and time reversible (latent) volatility factors, $f_{n,t}$, where for notational simplicity and without loss of generality we impose the normalization that $\bar{P}_{n,0}(f_{n,t}) = 1$ and $Var[\bar{P}_{n,i}(f_{n,t})] = 1$ for $i \neq 0$; see Meddahi (2001) for further discussion of the general ESV class of models.[12]

Fixing $\rho_1 = \rho_2 = 0$, the Brownian motion in (15), $\{W_u, u \geq 0\}$, collapses to $\{W_{3,u}, u \geq 0\}$, which by assumption is independent of the Brownian motions driving the (latent) volatility factors, and in turn therefore also independent of $\{\sigma_u, u \geq 0\}$, as formally assumed in the previous section. However, for all other values of ρ_1 and ρ_2, the Brownian motion driving the price process, $\{W_u, u \geq 0\}$, and the volatility process, $\{\sigma_u, u \geq 0\}$, will be correlated.

The expression for σ_t^2 in equation (16) may appear somewhat arbitrary. However, any square-integrable function of the latent volatility factors, f_t, may be written as a (possibly infinite) linear combination of the corresponding eigenfunctions. Importantly, this ESV representation greatly facilitates the calculation of conditional expectations of the volatilities and/or squared returns due to some convenient properties of the eigenfunctions. In particular, the eigenfunctions associated with different eigenvalues are orthogonal and (for any nonconstant eigenfunction) centered at zero, that is $E[P_{i,j}(f_t)P_{k,l}(f_t)] = 0$ and $E[P_{i,j}(f_t)] = 0$, for $(i,j) \neq (0,0)$ and $(i,j) \neq (k,l)$. In addition, the dynamics of the eigenfunctions follow first order autoregressive processes, so that $E[P_{i,j}(f_{t+l}) \mid f_\tau, \tau \leq t] = \exp(-\lambda_{i,j}l)P_{i,j}(f_t)$ for $l > 0$.

[12] The infinitesimal generator, $\mathcal{A}_n$, associated with $f_{n,t}$ is formally defined by

$$\mathcal{A}_n\phi(f_{n,t}) \equiv m_n(f_{n,t})\phi'(f_{n,t}) + \frac{v_n(f_{n,t})}{2}\phi''(f_{n,t}),$$

for any square-integrable and twice differentiable function, $\phi(f_{n,t})$. The corresponding eigenfunctions, $\bar{P}_{n,i}(f_{n,t})$, and eigenvalues, $\bar{\lambda}_{n,i}$, satisfy

$$\mathcal{A}_n\bar{P}_{n,i}(f_{n,t}) = -\bar{\lambda}_{n,i}\bar{P}_{n,i}(f_{n,t}).$$

Importantly, as shown in Meddahi (2001), the eigenfunctions of the infinitesimal generator associated with $f_t \equiv (f_{1,t}, f_{2,t})$ are simply given by $P_{i,j}(f_t) \equiv \bar{P}_{1,i}(f_{1,t})\bar{P}_{2,j}(f_{2,t})$ with corresponding eigenvalues $\lambda_{i,j} \equiv \bar{\lambda}_{1,i}\bar{\lambda}_{2,j}$. For a more detailed discussion of the properties of infinitesimal generators see, e.g., Hansen and Scheinkman (1995) and Aït-Sahalia, Hansen and Scheinkman (2003).

Utilizing these properties of the ESV representation, it is possible to show that for $h \to 0$ (see the Appendix for details),

$$Var[U_t(h)] \;=\; 2h\, E\left[IQ_t\right] \;+\; o(h), \tag{18}$$

and

$$Cov[IV_t, U_t(h)] \;=\; 2h\; E[r_t]\; Cov[r_t, RV_t(h)] \;+\; o(h)\;. \tag{19}$$

Moreover, the expected integrated quarticity may be estimated by the expected realized quarticity,

$$E[RQ_t(h)] \;=\; E[IQ_t] \;+\; o(1)\;. \tag{20}$$

This latter results would, of course, be implied by the aforementioned consistency of $RQ_t(h)$ for IQ_t under appropriate uniform integrability conditions. However, the corresponding proof of consistency in Barndorff-Nielsen and Shephard (2003a) again rules out leverage effects.[13] Now, combining the results in (18), (19), and (20), it follows readily that

$$\begin{aligned} Var[IV_t] \;&=\; Var[RV_t(h)] \;-\; Var[U_t(h)] \;-\; 2Cov[U_t(h), IV_t] \\ &=\; Var[RV_t(h)] \;-\; 2h\; E[RQ_t(h)] \;-\; 4h\; E[r_t]\; Cov[r_t, RV_t(h)] \;+\; o(h). \end{aligned} \tag{21}$$

Hence, relative to (9), the leverage effect introduces the additional $4hE[r_t]Cov[r_t, RV_t(h)]$ term in the (feasible) asymptotic approximation to $Var[IV_t]$.

In actual empirical applications $2E[r_t]$ and $Cov[r_t, RV_t(h)]$ will both generally be orders of magnitude smaller than $E[RQ_t(h)]$ so that, invariably, the magnitude of the new adjustment term will be very small (negligible) relative to the $2hE[RQ_t(h)]$ term. To illustrate, consider the five-minute high-frequency S&P500 and U.S. T-Bond futures returns spanning the period from January 1990 through December 2002.[14] The relative importance of the leverage adjustment term, as measured by the daily $2E[r_t]Cov[r_t, RV_t(h)]/E[RQ_t(h)]$ ratios, equal -7.85×10^{-5} and -6.94×10^{-4} for each of the two markets. Also, for the DM/\$, Yen/\$, and Yen/DM half-hour returns underlying the empirical results in ABDL (2003) discussed below, these same daily ratios for the full December 1986 through June 1999 sample period equal 1.01×10^{-4}, -7.87×10^{-5}, and 3.54×10^{-4}, respectively. Clearly an inconsequential addition to the approximation for $Var[IV_t]$ in (9).

These empirical observations are further corroborated by the Monte Carlo simulation results for the leverage models with constant as well as time-varying drifts reported in Tables A2 and A3. The medians in the asymptotic approximations to $Var[IV_t]$, $Var[IV_t^{1/2}]$ and

[13] Related general results are available in Lepingle (1976) and the thesis by Becker (1998) in an abstract setting; see also Barndorff-Nielsen, Graversen and Shephard (2003).

[14] The data has previously been analyzed in Andersen, Bollerslev, Diebold and Vega (2003) from a very different perspective. We refer the reader to that study for a more detailed description of the data source and return construction.

$Var[\log(IV_t^{1/2})]$ in equations (9), (11), and (14), respectively, derived under the assumption of no leverage are all right-on the true medians (labelled $h = 1/\infty$). Moreover, as long as the frequency of the returns used in the calculation of the realized volatility and quarticity measures exceeds half-an-hour, or $h \leq 1/48$, the simulated distributions for the leverage models are indistinguishable from the corresponding distributions for the same models without leverage reported in Table A1. In short, the realized volatility measurement error adjustment procedures developed in the preceding section remain highly accurate in empirically realistic situations allowing for both leverage and time-varying drift. We next turn to a re-interpretation of the empirical evidence related to the Mincer-Zarnowitz volatility regressions reported in ABDL (2003) based on an application of these procedures.

3 ABDL (2003) Revisited

The forecast comparisons in ABDL (2003) are based on daily realized volatilities constructed from high-frequency half-hourly, or h=1/48, spot exchange rates for the U.S. dollar, the Deutschemark and the Japanese yen spanning twelve-and-a-half years.[15] Separate forecast evaluation regressions are reported for the "in-sample" period comprised of the 2,449 "regular" trading days from December 1, 1986 through December 1, 1996, and the shorter "out-of-sample" forecast period consisting of the 596 days from December 2, 1996 through June 30, 1999. Separate results are also reported for one-day-ahead and ten-days-ahead forecasts. However, for all series and both sample periods and forecast horizons, a simple AR(5) model estimated directly from the realized volatilities generally performs as well or better than any of the many alternative models considered, including several GARCH type models estimated directly to the high-frequency data (both with and without corrections for the pronounced intradaily seasonal pattern in volatility). The representative R^2's for the DM/\$, Yen/\$, and Yen/DM forecast regressions for $RV_{t+1}(1/48)$, $RV_{t+1}(1/48)^{1/2}$, $\log RV_{t+1}(1/48)^{1/2}$, $RV_{t+10,10}(1/48)$, $RV_{t+10,10}(1/48)^{1/2}$, and $\log RV_{t+10,10}(1/48)^{1/2}$, where $RV_{t+10,10}(1/48) \equiv RV_{t+1}(1/48) + RV_{t+2}(1/48) + ... + RV_{t+10}(1/48)$, as reported in ABDL (2003) and the accompanying appendix, are given in square brackets in Table 1.[16]

By failing to account for the measurement errors in the future realized volatilities, these R^2's understate the true degree of predictability in the (latent) integrated volatilities. This problem is rectified by the main entries in Table 1, which report the adjusted R^2's obtained

[15] The high-frequency data were generously provided by Olsen & Associates in Zürich, Switzerland; see Dacorogna, Gencay, Müller, Olsen and Pictet (2001) for further discussion of the data capture, filtering, and return construction.

[16] The out-of-sample period contains a "once-in-a-generation" move in the Japanese Yen on October 8, 1998. Somewhat higher R^2's, but qualitatively similar results, were obtained by excluding this and the neighboring two days; see ABDL (2003) and the accompanying appendix for further discussion and sensitivity analysis along these lines.

by applying the (feasible) asymptotic approximations in equations (9), (11), and (14), along with the relevant multiplicative adjustment factors.[17] The results are quite striking. For some of the forecasts horizons and rates, the "true" R^2's exceed the standard predictive R^2's, as reported in ABDL (2003), by up to *forty percent.* For instance, the in-sample, one-day-ahead R^2 for the DM/$ series given in the very first entry in the table equals 0.219, whereas the true (albeit estimated) R^2 is substantially higher at 0.314. As such, the results clearly highlight the importance of appropriately adjusting for measurement error when assessing the quality of volatility forecasts in practical empirical applications.

Table 1
ABDL (2003) Adjusted Predictive R^2's

	IV	$IV^{1/2}$	$\log IV^{1/2}$
In-Sample, One-Day-Ahead			
DM/$	0.314 [0.219]	0.399 [0.351]	0.482 [0.431]
Yen/$	0.315 [0.229]	0.412 [0.374]	0.476 [0.433]
Yen/DM	0.450 [0.361]	0.559 [0.499]	0.630 [0.567]
Out-of-Sample, One-Day-Ahead			
DM/$	0.200 [0.158]	0.296 [0.246]	0.350 [0.285]
Yen/$	0.230 [0.197]	0.366 [0.338]	0.419 [0.373]
Yen/DM	0.215 [0.189]	0.378 [0.344]	0.483 [0.424]
In-Sample, Ten-Days-Ahead			
DM/$	0.411 [0.374]	0.463 [0.436]	0.499 [0.473]
Yen/$	0.386 [0.355]	0.414 [0.396]	0.424 [0.407]
Yen/DM	0.536 [0.513]	0.606 [0.589]	0.653 [0.637]
Out-of-Sample, Ten-Days-Ahead			
DM/$	0.182 [0.168]	0.209 [0.195]	0.228 [0.213]
Yen/$	0.197 [0.187]	0.287 [0.279]	0.347 [0.336]
Yen/DM	0.186 [0.178]	0.301 [0.293]	0.401 [0.390]

Note: The table reports the adjusted predictive R^2's from the Mincer-Zarnowitz regressions of the realized volatilities on the AR(5) volatility forecasts in ABDL (2003), along with the corresponding unadjusted R^2's (in square brackets). The realized volatility measures are constructed from high-frequency half-hour returns. The "in-sample" period covers December 1, 1986 through December 1, 1996, while the "out-sample" period spans December 2, 1996 through June 30, 1999.

Interestingly, the numerical values for the adjusted R^2's for the DM-dollar series in Table 1 are quite close to the exact theoretical R^2's implied by the specific two-factor affine diffusion discussed in Andersen, Bollerslev and Meddahi (2003). This is especially noteworthy in so far the parameter values for this model are based on the identical DM-dollar sample

[17] The adjustments are constructed separately for each series and for the in-sample and out-of-sample periods using the corresponding realized volatility and quarticity series.

underlying the results reported on in Table 1. This suggests that the simple AR(5) models for the realized volatilities estimated in ABDL (2003) - when adjusted for the measurement error problem - capture a degree of predictability that is consistent with that implied by a conventional two-factor affine model. This type of benchmarking of the true predictive power of such reduced-form forecast procedures relative to that of a specific continuous-time volatility model would, of course, be impossible without the type of measurement error correction developed here.

4 Concluding Remarks

Building on the recent theoretical results of Barndorff-Nielsen and Shephard (2002a, 2003a), this note develops a set of simple and practically feasible expressions for calculating true measures of return volatility predictability relative to that of the corresponding underlying (latent) integrated volatility. The procedures are general and could be applied in the evaluation of any volatility forecasts. The analytical results for the eigenfunction stochastic volatility class of models and accompanying simulation based evidence confirm that the procedures work equally well in situations with pronounced leverage effects. On specifically applying the procedures to the ex-post forecast evaluation regressions reported in ABDL (2003), we document sizeable downward biases in terms of the previously reported predictive powers. More generally, the practical techniques developed here hold the promise for further development of new and improved easy-to-implement volatility forecasting procedures guided by proper benchmark comparisons. The techniques should also prove useful in more effectively calibrating the type of continuous-time models routinely employed in modern asset pricing theories.

Appendix 1: Monte Carlo Simulations

In order to assess the accuracy of the distributional assumptions and second-order Taylor series expansions underlying the asymptotic approximations in (9), (11), and (14) in empirically relevant specifications and sample sizes compatible with those of ABDL (2003), Tables A1-A3 report the simulated medians and ninety-percent confidence intervals (in square brackets) across 1,000 replications, each consisting of 2,500 "days." We report the results for a total of nine different continuous-time models along with $1/h = 288, 96, 48$, and 1, corresponding to the use of "5-minute," "15-minute," "half-hourly," and "daily" returns.

The first three models reported in Table A1 fix the mean returns at zero, and assume that the volatility and the Brownian motion driving the price process are independent, i.e., no leverage effects. In the notation of equation (15), $\mu = \rho_1 = \rho_2 = 0$,

$$d\log(S_t) = \sigma_t dW_{3,t}.$$

The numbers in the first panel refer to the GARCH(1,1) diffusion analyzed in Andersen and Bollerslev (1998),

$$d\sigma_t^2 \;=\; 0.035(0.636 - \sigma_t^2)dt \;+\; 0.144\;\sigma_t^2 dW_{1,t}\;.$$

The second panel gives the results for the two-factor affine diffusion estimated by Bollerslev and Zhou (2002), $\sigma_t^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2$, where

$$d\sigma_{1,t}^2 \;=\; 0.5708(0.3257 - \sigma_{1,t}^2)dt \;+\; 0.2286\;\sigma_{1,t} dW_{1,t}\;,$$

$$d\sigma_{2,t}^2 \;=\; 0.0757(0.1786 - \sigma_{2,t}^2)dt \;+\; 0.1096\;\sigma_{2,t} dW_{2,t}\;.$$

These parameter values were obtained from estimation based on the identical DM-dollar sample used in ABDL (2003). The third set of numbers refer to the log-normal diffusion reported in Andersen, Benzoni and Lund (2002) with volatility dynamics governed by

$$d\log(\sigma_t^2) \;=\; -0.0136[0.8382 \;+\; \log(\sigma_t^2)]dt \;+\; 0.1148\;dW_{1,t}\;.$$

All of the models in Table A1 satisfy the Barndorff-Nielsen and Shephard (2002a, 2003b) regularity conditions discussed in Sections 2.1 and 2.2.

The results reported in Table A2 are based on the same three volatility specifications, but incorporate a positive drift and strong leverage effects. For the one-factor GARCH and log-normal diffusions,

$$d\log(S_t) \;=\; 0.0314\;dt \;+\; \sigma_t[-0.576\;dW_{1,t} + \sqrt{1 - 0.576^2} dW_{3,t}]\;,$$

where the values for the drift and leverage parameters are taken from Andersen, Benzoni and Lund (2002). For the two-factor affine model the instantaneous return dynamic is governed by,

$$d\log(S_t) \;=\; 0.0314\;dt \;+\; \sigma_t[0.9\;dW_{1,t} - 0.4\;dW_{2,t} + \sqrt{1 - 0.9^2 - 0.4^2} dW_{3,t}]\;,$$

with the two leverage parameters adapted from the estimates reported in Chernov, Gallant, Ghysels and Tauchen (2003).

In addition to the contemporaneous correlation between the return and volatility for the leverage models in Table A2, the last set of models in Table A3 also include a volatility feedback, or ARCH-in-mean, effect in the drift component. Specifically, for the two one-factor models,

$$d\log(S_t) \;=\; (0.0314 + 0.3\sigma_t^2)\;dt \;+\; \sigma_t[-0.576\;dW_{1,t} + \sqrt{1 - 0.576^2} dW_{3,t}]\;,$$

while for the two-factor model,

$$d\log(S_t) \;=\; (0.0314 + 0.3\sigma_t^2)\;dt \;+\; \sigma_t[0.9\;dW_{1,t} - 0.4\;dW_{2,t} + \sqrt{1 - 0.9^2 - 0.4^2} dW_{3,t}]\;.$$

The value of the slope coefficient in the drift is taken from Chernov (2003).

Table A1: Asymptotic Variance Approximations
Baseline Volatility Models

h	$Var[IV_t]$	$Var[IV_t^{1/2}]$	$Var[\log(IV_t^{1/2})]$
GARCH(1,1) Diffusion			
$1/\infty$	0.170	0.0647	0.138
	[0.117, 0.265]	[0.0518, 0.0853]	[0.112, 0.168]
1/288	0.170	0.0647	0.138
	[0.116, 0.266]	[0.0517, 0.0854]	[0.112, 0.168]
1/96	0.171	0.0648	0.138
	[0.116, 0.266]	[0.0520, 0.0859]	[0.112, 0.168]
1/48	0.170	0.0650	0.139
	[0.115, 0.268]	[0.0520, 0.0861]	[0.112, 0.169]
1	0.167	0.208	1.19
	[0.0923, 0.313]	[0.175, 0.248]	[1.08, 1.30]
Two-Factor Affine			
$1/\infty$	0.0259	0.0126	0.0261
	[0.0222, 0.0316]	[0.0111, 0.0145]	[0.0235, 0.0290]
1/288	0.0260	0.0126	0.0261
	[0.0222, 0.0316]	[0.0111, 0.0145]	[0.0234, 0.0291]
1/96	0.0260	0.0126	0.0263
	[0.0221, 0.0315]	[0.0111, 0.0146]	[0.0235, 0.0294]
1/48	0.0259	0.0127	0.0267
	[0.0219, 0.0315]	[0.0112, 0.0148]	[0.0238, 0.0302]
1	0.0245	0.136	1.07
	[0.00617, 0.0462]	[0.125, 0.149]	[0.973, 1.16]
Log-Normal Diffusion			
$1/\infty$	0.145	0.0544	0.109
	[0.0640, 0.333]	[0.0328, 0.0946]	[0.0764, 0.163]
1/288	0.144	0.0543	0.109
	[0.0643, 0.338]	[0.0330, 0.0943]	[0.0762, 0.163]
1/96	0.145	0.0546	0.109
	[0.0642, 0.337]	[0.0330, 0.0952]	[0.0766, 0.164]
1/48	0.144	0.0547	0.109
	[0.0635, 0.341]	[0.0331, 0.0953]	[0.0769, 0.165]
1	0.145	0.177	1.15
	[0.0529, 0.390]	[0.127, 0.252]	[1.05, 1.27]

Note: The table reports the simulated medians and ninety-percent confidence intervals (in square brackets) for the asymptotic approximations in equations (9), (11), and (14) across 1,000 replications, each consisting of 2,500 "days."

Table A2: Asymptotic Variance Approximations
Volatility Models with Leverage and Constant Drift

h	$Var[IV_t]$	$Var[IV_t^{1/2}]$	$Var[\log(IV_t^{1/2})]$
GARCH(1,1) Diffusion			
$1/\infty$	0.170	0.0647	0.138
	[0.117, 0.265]	[0.0518, 0.0853]	[0.112, 0.168]
1/288	0.170	0.0647	0.138
	[0.116, 0.262]	[0.0518, 0.0849]	[0.112, 0.168]
1/96	0.170	0.0647	0.138
	[0.116, 0.265]	[0.0520, 0.0852]	[0.112, 0.168]
1/48	0.170	0.0650	0.138
	[0.115, 0.268]	[0.0520, 0.0853]	[0.113, 0.168]
1	0.165	0.205	1.16
	[0.0964, 0.303]	[0.173, 0.247]	[1.07, 1.27]
Two-Factor Affine			
$1/\infty$	0.0259	0.0126	0.0261
	[0.0222, 0.0316]	[0.0111, 0.0145]	[0.0235, 0.0290]
1/288	0.0260	0.0126	0.0261
	[0.0221, 0.0317]	[0.0111, 0.0145]	[0.0234, 0.0292]
1/96	0.0261	0.0127	0.0263
	[0.0222, 0.0321]	[0.0112, 0.0145]	[0.0236, 0.0294]
1/48	0.0262	0.0129	0.0267
	[0.0221, 0.0323]	[0.0112, 0.0150]	[0.0239, 0.0301]
1	0.0370	0.139	1.07
	[0.0155, 0.0654]	[0.128, 0.154]	[0.973, 1.16]
Log-Normal Diffusion			
$1/\infty$	0.145	0.0544	0.109
	[0.0640, 0.333]	[0.0328, 0.0946]	[0.0764, 0.163]
1/288	0.144	0.0545	0.109
	[0.0640, 0.336]	[0.0329, 0.0941]	[0.0763, 0.162]
1/96	0.145	0.0545	0.109
	[0.0637, 0.337]	[0.0331, 0.0952]	[0.0763, 0.163]
1/48	0.146	0.0547	0.110
	[0.0635, 0.340]	[0.0335, 0.0943]	[0.0766, 0.162]
1	0.145	0.177	1.15
	[0.0515, 0.375]	[0.127, 0.251]	[1.04, 1.27]

Note: See Table A1.

Table A3: Asymptotic Variance Approximations
Volatility Models with Leverage and Time-Varying Drift

h	$Var[IV_t]$	$Var[IV_t^{1/2}]$	$Var[\log(IV_t^{1/2})]$
GARCH(1,1) Diffusion			
$1/\infty$	0.170	0.0647	0.138
	[0.117, 0.265]	[0.0518, 0.0853]	[0.112, 0.168]
1/288	0.170	0.0648	0.138
	[0.116, 0.262]	[0.0518, 0.0850]	[0.112, 0.168]
1/96	0.171	0.0647	0.138
	[0.116, 0.266]	[0.0521, 0.0853]	[0.112, 0.168]
1/48	0.171	0.0652	0.138
	[0.116, 0.270]	[0.0521, 0.0855]	[0.113, 0.168]
1	0.196	0.225	1.17
	[0.116, 0.398]	[0.189, 0.272]	[1.05, 1.27]
Two-Factor Affine			
$1/\infty$	0.0259	0.0126	0.0261
	[0.0222, 0.0316]	[0.0111, 0.0145]	[0.0235, 0.0290]
1/288	0.0261	0.0126	0.0262
	[0.0222, 0.0318]	[0.0111, 0.0146]	[0.0235, 0.0292]
1/96	0.0264	0.0128	0.0265
	[0.0225, 0.0324]	[0.0113, 0.0147]	[0.0238, 0.0296]
1/48	0.0268	0.0131	0.0272
	[0.0226, 0.0329]	[0.0115, 0.0152]	[0.0243, 0.0306]
1	0.0661	0.163	1.09
	[0.0362, 0.106]	[0.150, 0.180]	[0.998, 1.18]
Log-Normal Diffusion			
$1/\infty$	0.145	0.0544	0.109
	[0.0640, 0.333]	[0.0328, 0.0946]	[0.0764, 0.163]
1/288	0.145	0.0545	0.109
	[0.0641, 0.336]	[0.0329, 0.0942]	[0.0764, 0.163]
1/96	0.145	0.0546	0.109
	[0.0639, 0.337]	[0.0332, 0.0954]	[0.0764, 0.163]
1/48	0.146	0.0548	0.110
	[0.0636, 0.342]	[0.0336, 0.0948]	[0.0768, 0.163]
1	0.164	0.192	1.15
	[0.0608, 0.543]	[0.136, 0.281]	[1.04, 1.27]

Note: See Table A1.

Appendix 2: Technical Proofs

Proof of equations (10), (11), (12), (13), and (14). Let $f(\cdot)$ be a twice-differentiable function. By (6) and a second order Taylor approximation of $f(\cdot)$ at the point $RV_t(h)$ around IV_t, it follows that

$$f(RV_t(h)) \approx f(IV_t) + f'(IV_t)\sqrt{2hIQ_t}\; z_t + \frac{1}{2}f''(IV_t)2hIQ_t\; z_t^2. \tag{A.1}$$

Consequently,

$$\begin{aligned} E[f(RV_t(h))] &= E[f(IV_t)] + E[f'(IV_t)\sqrt{2hIQ_t}\; z_t] + \frac{1}{2}E[f''(IV_t)2hIQ_t\; z_t^2] + o(h) \\ &= E[f(IV_t)] + E[f'(IV_t)\sqrt{2hIQ_t}E[z_t \mid \sigma_u, t-1 \le u \le t]] \\ &+ \frac{1}{2}E[f''(IV_t)2hIQ_t E[z_t^2 \mid \sigma_u, t-1 \le u \le t]] + o(h) \\ &= E[f(IV_t)] + \frac{1}{2}E[f''(IV_t)2hIQ_t] + o(h), \end{aligned}$$

so that,

$$E[f(RV_t(h))] = E[f(IV_t)] + \frac{1}{2}E[f''(RV_t(h))2hRQ_t(h)] + o(h) \tag{A.2}$$

provided $E[f''(RV_t(h))RQ_t(h)] - E[f''(IV_t)IQ_t] = o(1)$. Equations (10), (12), and (13) follows by applying (A.1) to the functions $f_1(x) = x^{1/2}$, $f_2(x) = \log(x)$, and $f_3(x) = \log(x)^2$, where $f_1'(x) = 2^{-1}x^{-1/2}$, $f_1''(x) = -2^{-2}x^{-3/2}$, $f_2'(x) = x^{-1}$, $f_2''(x) = -x^{-2}$, $f_3'(x) = 2x^{-1}\log(x)$, and $f_3''(x) = 2x^{-2}(1-\log(x))$. By applying (A.2) to the function $f_1(\cdot)$, one gets (11). Similarly, by applying (A.2) to the functions $f_2(\cdot)$ and $f_3(\cdot)$, one gets (14).■

Proof of equation (18). By definition,

$$U_t(h) = \sum_{i=1}^{1/h} u_{t-1+ih}^{(h)} \quad \text{where} \quad u_{t-1+ih}^{(h)} \equiv r_{t-1+ih}^{(h)2} - \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du.$$

By similar arguments to Meddahi (2002), Proposition 4.2, it follows that the $u_{t-1+ih}^{(h)}$'s are uncorrelated, so that

$$Var[U_t(h)] = Var\left[\sum_{i=1}^{1/h} u_{t-1+ih}^{(h)}\right] = \frac{Var[u_{t-1+ih}^{(h)}]}{h}.$$

Hence, by Ito's Lemma,

$$u_{t-1+ih}^{(h)} = \mu^2 h^2 + 2\mu h \varepsilon_{t-1+ih}^{(h)} + 2Z_{t-1+ih}^{(h)}$$

where

$$\varepsilon_{t-1+ih}^{(h)} = \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u, \quad Z_{t-1+ih}^{(h)} = \int_{t-1+(i-1)h}^{t-1+ih}\left(\int_{t-1+(i-1)h}^{u} \sigma_s dW_s\right)\sigma_u dW_u.$$

Therefore,

$$Var[u_h^{(h)}] = 4\mu^2 h^2 Var[\varepsilon_h^{(h)}] + 4Var[Z_h^{(h)}] + 8\mu h Cov(\varepsilon_h^{(h)}, Z_h^{(h)})$$

In the sequel, we will show that for $h \to 0$,

$$Var[\varepsilon_h^{(h)}] = o(1), \tag{A.3}$$

$$Var[Z_h^{(h)}] = \frac{h^2}{2}\sum_{0\le i,j\le p} a_{i,j}^2 + o(h^2), \tag{A.4}$$

$$Cov(\varepsilon_h^{(h)}, Z_h^{(h)}) = o(h), \tag{A.5}$$

which in turn implies that

$$Var[u_h^{(h)}] = h^2\; 2\sum_{0\le i,j\le p} a_{i,j}^2 + o(h^2)\ ,$$

and thus

$$Var[h^{-1/2}U_t(h)] = 2\sum_{0\leq i,j\leq p} a_{i,j}^2 + o(1).$$

By formulas (4.2) and (4.5) in Andersen, Bollerslev and Meddahi (2003), $E[\sigma_t^4] = \sum_{0\leq i,j\leq p} a_{i,j}^2$, which combined with (7), achieves the proof of (18).

Proof of (A.3): By definition of $\varepsilon_h^{(h)}$, $\varepsilon_h^{(h)2} = \int_0^h \sigma_u^2 du$ and $E[\varepsilon_h^{(h)}] = 0$. Hence

$$Var[\varepsilon_h^{(h)}] = E[\varepsilon_h^{(h)2}] = \int_0^h E[\sigma_u^2]du = a_{0,0}h, \tag{A.6}$$

given that $E[P_{i,j}(f_u)] = 0$ when $i \neq 0$ or $j \neq 0$, which implies (A.3).

Proof of (A.4): Note that

$$\begin{aligned}
Var[Z_h^{(h)}] &= E\left[\int_0^h \left(\int_0^u \sigma_s dW_s\right)^2 \sigma_u^2 du\right] \\
&= E\left[\int_0^h \left(\int_0^u \sigma_s^2 ds\right)\sigma_u^2 du\right] + 2E\left[\int_0^h \left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right)\sigma_u^2 du\right] \\
&= E\left[\int_0^h \left(\int_0^u \sigma_s^2 ds\right)\sigma_u^2 du\right] + 2E\left[\int_0^h Z_u^{(u)}\sigma_u^2 du\right] \\
&= \int_0^h \left(\int_0^u E[\sigma_s^2\sigma_u^2]ds\right)du + 2\sum_{0\leq i,j\leq p} a_{i,j}\int_0^h E[Z_u^{(u)}P_{i,j}(f_u)]du.
\end{aligned}$$

Denoting $\lambda_{i,j} = \bar{\lambda}_{1,i}\bar{\lambda}_{2,j}$, it follows from (4.2) and (4.8) in Andersen, Bollerslev and Meddahi (2003) that

$$\forall s,u,\ s\leq u: \quad E[\sigma_s^2\sigma_u^2] = \sum_{0\leq i,j\leq p} a_{i,j}^2 \exp(-\lambda_{i,j}(u-s)),$$

and therefore

$$\int_0^h \left(\int_0^u E[\sigma_s^2\sigma_u^2]ds\right)du = \sum_{0\leq i,j\leq p} a_{i,j}^2\frac{1}{\lambda_{i,j}^2}[\exp(-\lambda_{i,j}h) - 1 + \lambda_{i,j}h] = \frac{h^2}{2}\sum_{0\leq i,j\leq p} a_{i,j}^2 + o(h^2).$$

Thus, in order to prove (A.4), it suffice to show that

$$\int_0^h E[Z_u^{(u)}P_{i,j}(f_u)]du = o(h^2). \tag{A.7}$$

Applying (A2) in Meddahi (2002) to the factors $f_{n,t}$, $n = 1,2$, it follows that

$$\bar{P}_{n,i}(f_{n,u}) = \exp(-\bar{\lambda}_{n,i}(u-s))\bar{P}_{n,i}(f_{n,s}) + \exp(-\bar{\lambda}_{n,i}(u-s))\int_s^u \exp(\bar{\lambda}_{n,i}(w-s))\sqrt{v_n(f_{n,w})}\bar{P}'_{n,i}(f_{n,w})\, dW_{n,w},$$

$$\begin{aligned}
P_{i,j}(f_u) &= \bar{P}_{1,i}(f_{1,u})\bar{P}_{2,j}(f_{2,u}) = \exp(-\lambda_{i,j}(u-s))\times \\
&\Big\{P_{i,j}(f_s) + \bar{P}_{1,i}(f_{1,s})\int_s^u \exp(\bar{\lambda}_{2,j}(w-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,w})\, dW_{2,w} \\
&+\bar{P}_{2,j}(f_{2,s})\int_s^u \exp(\bar{\lambda}_{1,i}(w-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,w})\, dW_{1,w} \\
&+\int_s^u \exp(\bar{\lambda}_{1,i}(w-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,w})\, dW_{1,w}\int_s^u \exp(\bar{\lambda}_{2,j}(w-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,w})\, dW_{2,w}\Big\}.
\end{aligned}$$

However, by Ito's Lemma,

$$\begin{aligned}
&\int_s^u \exp(\bar{\lambda}_{1,i}(w-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,w})\, dW_{1,w}\int_s^u \exp(\bar{\lambda}_{2,j}(w-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,w})\, dW_{2,w} \\
&= \int_s^u \exp(\bar{\lambda}_{1,i}(w-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,w})\left[\int_s^w \exp(\bar{\lambda}_{2,j}(z-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,z})\, dW_{2,z}\right] dW_{1,w}
\end{aligned}$$

$$+\int_s^u \exp(\bar{\lambda}_{2,j}(w-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,w})\left[\int_s^w \exp(\bar{\lambda}_{1,i}(z-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,z})\ dW_{1,z}\right]\ dW_{2,w}.$$

Thus,

$$\begin{aligned}
P_{i,j}(f_u) &= \exp(-\lambda_{i,j}(u-s))\times \\
&\Big\{P_{i,j}(f_s) + \bar{P}_{1,i}(f_{1,s})\int_s^u \exp(\bar{\lambda}_{2,j}(w-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,w})\ dW_{2,w} \\
&+\bar{P}_{2,j}(f_{2,s})\int_s^u \exp(\bar{\lambda}_{1,i}(w-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,w})\ dW_{1,w} \\
&+\int_s^u \exp(\bar{\lambda}_{1,i}(w-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,w})\left[\int_s^w \exp(\bar{\lambda}_{2,j}(z-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,z})\ dW_{2,z}\right]\ dW_{1,w} \\
&+\int_s^u \exp(\bar{\lambda}_{2,j}(w-s))\sqrt{v_2(f_{2,w})}\bar{P}'_{2,j}(f_{2,w})\left[\int_s^w \exp(\bar{\lambda}_{1,i}(z-s))\sqrt{v_1(f_{1,w})}\bar{P}'_{1,i}(f_{1,z})\ dW_{1,z}\right]\ dW_{2,w}\Big\}.
\end{aligned} \tag{A.8}$$

Consequently,

$$\begin{aligned}
E[Z_u^{(u)}P_{i,j}(f_u)] &= E\left[\int_0^u\left(\int_0^s \sigma_w dW_w\right)\sigma_s \exp(-\lambda_{i,j}(u-s))\bar{P}_{1,i}(f_{1,s})\sqrt{v_2(f_{2,s})}\bar{P}'_{2,j}(f_{2,s})\rho_2 ds\right] \\
&+E\left[\int_0^u\left(\int_0^s \sigma_w dW_w\right)\sigma_s \exp(-\lambda_{i,j}(u-s))\bar{P}_{2,j}(f_{2,s})\sqrt{v_1(f_{1,s})}\bar{P}'_{1,i}(f_{1,s})\rho_1 ds\right],
\end{aligned}$$

i.e., the first, fourth and fifth terms on the right-hand side of (A.8) do not contribute to $E[Z_u^{(u)}P_{i,j}(f_u)]$. Now define the reals $n_{i,j}$, $e_{i,j,k,l}$, $m_{i,j}$, and $d_{i,j,k,l}$ by the expansions in mean-square:

$$\sigma_s\bar{P}_{1,i}(f_{1,s})\sqrt{v_2(f_{2,s})}\bar{P}'_{2,j}(f_{2,s}) = \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}P_{k,l}(f_s),$$

$$\sigma_s\bar{P}_{2,j}(f_{2,s})\sqrt{v_1(f_{1,s})}\bar{P}'_{1,i}(f_{1,s}) = \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}P_{k,l}(f_s).$$

Then,

$$\begin{aligned}
E[Z_u^{(u)}P_{i,j}(f_u)] &= \rho_2 \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}E\left[\int_0^u\left(\int_0^s \sigma_w dW_w\right)\exp(-\lambda_{i,j}(u-s))P_{k,l}(f_s)ds\right] \\
&+\rho_1 \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}E\left[\int_0^u\left(\int_0^s \sigma_w dW_w\right)\exp(-\lambda_{i,j}(u-s))P_{k,l}(f_s)ds\right] \\
&=\rho_2 \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}\int_0^u\left(\int_0^s E[P_{k,l}(f_s)\sigma_w dW_w]\right)\exp(-\lambda_{i,j}(u-s))ds \\
&+\rho_1 \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}\int_0^u\left(\int_0^s E[P_{k,l}(f_s)\sigma_w dW_w]\right)\exp(-\lambda_{i,j}(u-s))ds.
\end{aligned}$$

Moreover,

$$\begin{aligned}
&E[P_{k,l}(f_s)\sigma_w dW_w] \\
&\quad=\exp(-\lambda_{k,l}(s-w))E\left[\rho_1\bar{P}_{2,l}(f_{2,w})\sqrt{v_1(f_{1,w})}\bar{P}'_{1,k}(f_{1,w})+\rho_2\bar{P}_{1,i}(f_{1,w})\sqrt{v_2(f_{2,w})}\bar{P}'_{2,l}(f_{2,w})\right]dw \\
&\quad=\exp(-\lambda_{k,l}(s-w))(\rho_1 d_{k,l,0,0}+\rho_2 e_{k,l,0,0})dw.
\end{aligned} \tag{A.9}$$

Hence,

$E[Z_u^{(u)} P_{i,j}(f_u)]$

$$= \rho_2 \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) \int_0^u \left(\int_0^s \exp(-\lambda_{k,l}(s-w))dw] \right) \exp(-\lambda_{i,j}(u-s))ds$$

$$+ \rho_1 \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) \int_0^u \left(\int_0^s \exp(-\lambda_{k,l}(s-w))dw] \right) \exp(-\lambda_{i,j}(u-s))ds$$

$$= \rho_2 \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) \left(\frac{1-\exp(-\lambda_{i,j}u)}{\lambda_{i,j}\lambda_{k,l}} - \frac{\exp(-\lambda_{k,l}u) - \exp(-\lambda_{i,j}u)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{k,l}} \right)$$

$$+ \rho_1 \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) \left(\frac{1-\exp(-\lambda_{i,j}u)}{\lambda_{i,j}\lambda_{k,l}} - \frac{\exp(-\lambda_{k,l}u) - \exp(-\lambda_{i,j}u)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{k,l}} \right). \quad \text{(A.10)}$$

Therefore,

$\int_0^h E[Z_u^{(u)} P_{i,j}(f_u)]du$

$$= \rho_2 \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) \int_0^h \left(\frac{1-\exp(-\lambda_{i,j}u)}{\lambda_{i,j}\lambda_{k,l}} - \frac{\exp(-\lambda_{k,l}u) - \exp(-\lambda_{i,j}u)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{k,l}} \right) du$$

$$+ \rho_1 \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) \int_0^h \left(\frac{1-\exp(-\lambda_{i,j}u)}{\lambda_{i,j}\lambda_{k,l}} - \frac{\exp(-\lambda_{k,l}u) - \exp(-\lambda_{i,j}u)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{k,l}} \right) du.$$

But

$$\int_0^h \left(\frac{1-\exp(-\lambda_{i,j}u)}{\lambda_{i,j}\lambda_{k,l}} - \frac{\exp(-\lambda_{k,l}u) - \exp(-\lambda_{i,j}u)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{k,l}} \right) du$$
$$= \frac{1}{\lambda_{k,l}} \left[\frac{h}{\lambda_{i,j}} - \frac{1-\exp(-\lambda_{i,j}h)}{\lambda_{i,j}^2} - \frac{1-\exp(-\lambda_{k,l}h)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{k,l}} + \frac{1-\exp(-\lambda_{i,j}h)}{(\lambda_{i,j}-\lambda_{k,l})\lambda_{i,j}} \right] = \frac{1}{6}h^3 + o(h^3),$$

which implies (A.7) and therefore (A.4).

Proof of (A.5): Since $E[\varepsilon_h^{(h)}] = 0$, $Cov(\varepsilon_h^{(h)}, Z_h^{(h)}) = E[\varepsilon_h^{(h)} Z_h^{(h)}]$. By Ito's Lemma,

$$\varepsilon_h^{(h)} Z_h^{(h)} = \int_0^h Z_u^{(u)} d\varepsilon_u^{(u)} + \int_0^h \varepsilon_u^{(u)} dZ_u^{(h)} + \int_0^h d[Z,\varepsilon]_u,$$

where $[Z,\varepsilon]_u$ denotes the quadratic covariation of $Z_u^{(u)}$ and $\varepsilon_u^{(u)}$ over the interval $[0,u]$, i.e.,

$$[Z,\varepsilon]_u = \int_0^u \left(\int_0^s \sigma_w dW_w \right) \sigma_s^2 ds = \int_0^u \varepsilon_s^{(s)} \sigma_s^2 ds.$$

Note that $E[\int_0^h Z_u^{(u)} d\varepsilon_u^{(u)}] = 0$ and $E[\int_0^h \varepsilon_u^{(u)} dZ_u^{(u)}] = 0$. Hence,

$$E[\varepsilon_h^{(h)} Z_h^{(h)}] = E\left[\int_0^h \varepsilon_u^{(u)} \sigma_u^2 du \right] = \sum_{0\le i,j\le p} a_{i,j} \int_0^h E[\varepsilon_u^{(u)} P_{i,j}(f_u)]du = \sum_{0\le i,j\le p} a_{i,j} \int_0^h \int_0^u E[\sigma_s dW_s P_{i,j}(f_u)]du. \quad \text{(A.11)}$$

Lastly, by (A.9),

$$E[\varepsilon_h^{(h)} Z_h^{(h)}] = \sum_{0\le i,j\le p} a_{i,j} \int_0^h \int_0^u \exp(-\lambda_{i,j}(u-s)) \left(\rho_2 e_{i,j,00} + \rho_1 d_{i,j,0,0} \right) dsdu$$
$$= \sum_{0\le i,j\le p} a_{i,j} \left(\rho_2 e_{i,j,00} + \rho_1 d_{i,j,0,0} \right) \frac{1}{\lambda_{i,j}^2} [\exp(-\lambda_{i,j}h) - 1 + \lambda_{i,j}h)]$$
$$= \frac{h^2}{2} \sum_{0\le i,j\le p} a_{i,j} \left(\rho_2 e_{i,j,00} + \rho_1 d_{i,j,0,0} \right) \quad + \quad o(h^2),$$

which implies (A.5), and achieves the proof of (18).

Proof of equation (20). By definition,

$$E[RQ_t(h)] = \frac{1}{3h} E\left[\sum_{i=1}^{1/h} r_{t-1+ih}^{(h)4}\right] = \frac{1}{3h^2} E[r_h^{(h)4}].$$

Thus, it suffice to show that

$$E[r_h^{(h)4}] = 3h^2 \sum_{0\le i,j\le p} a_{i,j}^2 + o(h^2). \tag{A.12}$$

Note that by (A.6), $r_h^{(h)} = \mu h + \varepsilon_h^{(h)}$ and $E[\varepsilon_h^{(h)2}] = a_{0,0}h$. Hence,

$$E[r_h^{(h)4}] = Var[r_h^{(h)2}] + E[r_h^{(h)2}]^2 = Var[r_h^{(h)2}] + (\mu^2h^2 + ha_{0,0})^2 = Var[r_h^{(h)2}] + a_{0,0}^2h^2 + o(h^2). \tag{A.13}$$

In addition, by using Ito's Lemma for $k = 1, ..., 1/h$,

$$r_{kh}^{(h)2} = \mu^2h^2 + \varepsilon_{kh}^{(h)2} + 2\mu h\varepsilon_{kh}^{(h)} = \mu^2h^2 + iv_{kh}^{(h)} + 2Z_{kh}^{(h)} + 2\mu h\varepsilon_{kh}^{(h)} \quad \text{where} \quad iv_{kh}^{(h)} = \int_{(k-1)h}^{kh} \sigma_u^2 du.$$

Therefore,

$$\begin{aligned} Var[r_h^{(h)2}] &= Var[iv_h^{(h)}] + 4Var[Z_h^{(h)}] + 4\mu^2h^2Var[\varepsilon_h^{(h)}] \\ &+4Cov(iv_h^{(h)}, Z_h^{(h)}) + 4\mu hCov(iv_h^{(h)}, \varepsilon_h^{(h)}) + 8\mu hCov(Z_h^{(h)}, \varepsilon_h^{(h)}). \end{aligned} \tag{A.14}$$

In the sequel, we will show that

$$Var[iv_h^{(h)}] = h^2 \sum_{0\le i,j\le p,(i,j)\ne(0,0)} a_{i,j}^2 + o(h^2), \tag{A.15}$$

$$Cov(iv_h^{(h)}, Z_h^{(h)}) = o(h^2), \tag{A.16}$$

$$Cov(iv_h^{(h)}, \varepsilon_h^{(h)}) = o(h). \tag{A.17}$$

Combining (A.13) and (A.14), along with (A.4), (A.5), (A.15), (A.16), and (A.17), achieves the proof of (A.12).

Proof of (A.15): By (4.5) in Andersen, Bollerslev and Meddahi (2003) it follows that

$$Var[iv_h^{(h)}] = 2 \sum_{0\le i,j\le p,\ (i,j)\ne(0,0)} \frac{a_{i,j}^2}{\lambda_{i,j}^2}[\exp(-\lambda_{i,j}h) + \lambda_{i,j}h - 1] = h^2 \sum_{0\le i,j\le p,\ (i,j)\ne(0,0)} a_{i,j}^2 + o(h^2).$$

Proof of (A.16): The quadratic covariation of $iv_h^{(h)}$ and $Z_h^{(h)}$ is zero. Hence, by Ito's Lemma

$$iv_h^{(h)}Z_h^{(h)} = \int_0^h Z_u^{(u)} div_u^{(u)} + \int_0^h iv_u^{(u)} dZ_u^{(u)} = \int_0^h Z_u^{(u)}\sigma_u^2 du + \int_0^h iv_u^{(u)} \left(\int_0^u \sigma_s dW_s\right)\sigma_u dW_u,$$

which implies (A.16) since

$$Cov(iv_h^{(h)}, Z_h^{(h)}) = E(iv_h^{(h)}Z_h^{(h)}) = \int_0^h E[Z_u^{(u)}\sigma_u^2]du = \sum_{0\le i,j\le p} a_{i,j}\int_0^h E[Z_u^{(u)}P_{i,j}(u)]du = o(h^2),$$

where the last equality holds by (A.7).

Proof of (A.17): By similar arguments to the ones above, the quadratic covariation between $iv_h^{(h)}$ and $\varepsilon_h^{(h)}$ equals zero. Hence, using Ito's Lemma

$$E\left[iv_h^{(h)}\varepsilon_h^{(h)}\right] = E\left[\int_0^h \varepsilon_u^{(u)} div_u^{(u)} + \int_0^h iv_u^{(u)} d\varepsilon_u^{(u)}\right] = E\left[\int_0^h \varepsilon_u^{(u)}\sigma_u^2 du\right] = E[\varepsilon_h^{(h)}Z_h^{(h)}],$$

where the last equality holds by the first equality in (A.11), so that (A.5) implies (A.17).

Proof of equation (19). Note that $\varepsilon_1^{(1)} = r_1 - \mu$, where $\mu = E[r_t]$ and $\varepsilon_1^{(1)} = \sum_{k=1}^{1/h} \varepsilon_{kh}^{(h)}$. Thus,

$$Cov(U_t(h), IV_t) = Cov(U_1(h), IV_1) = Cov(\sum_{k=1}^{1/h} u_{kh}^{(h)}, IV_1) = Cov\left(\sum_{k=1}^{1/h} Z_{kh}^{(h)} + 2\mu h \varepsilon_1^{(1)}, IV_1\right)$$
$$= Cov\left(\sum_{k=1}^{1/h} Z_{kh}^{(h)}, IV_1\right) + 2\mu h Cov(IV_1, \varepsilon_1^{(1)}).$$

In the sequel, we will show that

$$Cov\left(\sum_{k=1}^{1/h} Z_{kh}^{(h)}, IV_1\right) = o(h), \tag{A.18}$$
$$Cov(IV_1, \varepsilon_1^{(1)}) = Cov(RV_1(h), \varepsilon_1^{(1)}) + o(1), \tag{A.19}$$

which achieve the proof of (19).
Proof of (A.18): Since $E[Z_{kh}^{(h)} \mid f_s, W_s, s \leq (k-1)h] = 0.$, it follows that

$$Cov\left(\sum_{k=1}^{1/h} Z_{kh}^{(h)}, IV_1\right) = Cov(\sum_{k=1}^{1/h} Z_{kh}^{(h)}, \sum_{k=1}^{1/h} iv_{kh}^{(h)}) = \sum_{k=1}^{1/h} Cov(Z_{kh}^{(h)}, iv_{kh}^{(h)}) + \sum_{k=1}^{1/h} \sum_{l=k+1}^{1/h} Cov(Z_{kh}^{(h)}, iv_{lh}^{(h)}).$$

Thus,

$$Cov(\sum_{k=1}^{1/h} Z_{kh}^{(h)}, IV_1) = h^{-1} Cov(Z_h^{(h)}, iv_h^{(h)}) + \sum_{k=1}^{1/h-1} (h^{-1} - k) Cov(Z_h^{(h)}, iv_{(k+1)h}^{(h)})$$

But (A.16) implies that $Cov(Z_h^{(h)}, iv_h^{(h)}) = o(h^2)$, so that (A.18) follows from

$$\sum_{k=1}^{1/h-1} (h^{-1} - k) Cov(Z_h^{(h)}, iv_{(k+1)h}^{(h)}) = o(h). \tag{A.20}$$

Now using the AR(1) structure of the eigenfunctions,

$$E[iv_{(k+1)h}^{(h)} \mid f_s, W_s, s \leq h] = \sum_{0 \leq i,j \leq p} a_{ij} \int_0^h E[P_{ij}(f_{kh+u}) \mid f_s, W_s, s \leq h] du$$
$$= \sum_{0 \leq i,j \leq p} a_{ij} \int_0^h \exp(-\lambda_{ij}(u + (k-1)h)) du \;\; P_{ij}(f_h) = \sum_{0 \leq i,j \leq p} a_{ij} \exp(-\lambda_{ij} kh) \frac{1 - \exp(-\lambda_{ij} h)}{\lambda_{ij}} \; P_{ij}(f_h).$$

Hence,

$$\begin{aligned} Cov(Z_h^{(h)}, iv_{(k+1)h}^{(h)}) &= Cov(Z_h^{(h)}, E[iv_{(k+1)h}^{(h)} \mid f_s, W_s, s \leq h]) \\ &= \sum_{0 \leq i,j \leq p} a_{ij} \exp(-\lambda_{ij} kh) \frac{1 - \exp(-\lambda_{ij} h)}{\lambda_{ij}} \; Cov(Z_h^{(h)}, P_{ij}(f_h)) \\ &= \sum_{0 \leq i,j \leq p, \; (i,j) \neq (0,0)} a_{ij} \exp(-\lambda_{ij} kh) \frac{1 - \exp(-\lambda_{ij} h)}{\lambda_{ij}} \; E(Z_h^{(h)} P_{ij}(f_h)). \end{aligned}$$

Note that

$$\begin{aligned} \frac{1 - \exp(-\lambda_{i,j} u)}{\lambda_{i,j} \lambda_{k,l}} - \frac{\exp(-\lambda_{k,l} u) - \exp(-\lambda_{i,j} u)}{(\lambda_{i,j} - \lambda_{k,l}) \lambda_{k,l}} &= \frac{u - \lambda_{i,j} u^2/2 + o(u^2)}{\lambda_{k,l}} \\ &\quad - \frac{-\lambda_{k,l} u + \lambda_{k,l}^2 u^2/2 + \lambda_{i,j} u - \lambda_{i,j}^2 u^2/2 + o(u^2)}{(\lambda_{i,j} - \lambda_{k,l}) \lambda_{k,l}} \\ &= \frac{u^2}{2} + o(u^2). \end{aligned}$$

Moreover, by (A.10)

$$E[Z_h^{(h)} P_{ij}(f_h)] = b_{ij} h^2 + o(h^2),$$

where

$$b_{i,j} = \frac{1}{2}\left(\rho_2 \sum_{0\le k,l\le n_{i,j}} e_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0}) + \rho_1 \sum_{0\le k,l\le m_{i,j}} d_{i,j,k,l}(\rho_1 d_{k,l,0,0} + \rho_2 e_{k,l,0,0})\right).$$

Therefore

$$Cov(Z_h^{(h)}, iv_{(k+1)h}^{(h)}) = \sum_{0\le i,j\le p,\ (i,j)\neq(0,0)} a_{ij}(b_{ij}h^2 + o(h^2))\exp(-\lambda_{ij}kh)\frac{1-\exp(-\lambda_{ij}h)}{\lambda_{ij}},$$

and

$$\sum_{k=1}^{1/h-1}(h^{-1}-k)Cov(Z_h^{(h)}, iv_{(k+1)h}^{(h)})$$

$$= \sum_{0\le i,j\le p,\ (i,j)\neq(0,0)} a_{ij}(b_{ij}h^2+o(h^2))\left(\sum_{k=1}^{1/h-1}(h^{-1}-k)\exp(-\lambda_{ij}kh)\right)\frac{1-\exp(-\lambda_{ij}h)}{\lambda_{ij}}.$$

But

$$\frac{1-\exp(-\lambda_{ij}h)}{\lambda_{ij}} = h + o(h)\ ,$$

and

$$\sum_{k=1}^{1/h-1}(h^{-1}-k)\exp(-\lambda_{ij}kh) = h^{-1}\sum_{k=1}^{1/h-1}(1-kh)\exp(-\lambda_{ij}kh) = \int_0^1 (1-x)\exp(-\lambda_{ij}x)dx + o(1) = c_{ij} + o(1),$$

where

$$c_{ij} = \int_0^1 (1-x)\exp(-\lambda_{ij}x)dx.$$

Hence,

$$\begin{aligned}\sum_{k=1}^{1/h-1}(h^{-1}-k)Cov(Z_h^{(h)}, iv_{(k+1)h}^{(h)}) &= \sum_{0\le i,j\le p,\ (i,j)\neq(0,0)} a_{ij}(b_{ij}h^2+o(h^2))(c_{ij}+o(1))(h+o(h))\\ &= \sum_{0\le i,j\le p,\ (i,j)\neq(0,0)} a_{ij}b_{ij}c_{ij}h^3 + o(h^3),\end{aligned}$$

which implies (A.20), and achieves the proof of (A.18).
Proof of (A.19): Note that $Cov(RV_1(h), \varepsilon_1^{(1)}) = Cov(IV_1, \varepsilon_1^{(1)}) + Cov(U_1(h), \varepsilon_1^{(1)})$. Since the vector $(Z_{kh}^{(h)}, \varepsilon_{kh}^{(h)})^\top$ is a martingale difference sequence,

$$Cov(U_1(h), \varepsilon_1^{(1)}) = 2\mu h Var[\varepsilon_1^{(1)}] + 2Cov(\sum_{k=1}^{1/h} Z_{kh}^{(h)}, \sum_{k=1}^{1/h}\varepsilon_{kh}^{(h)}) = 2\mu h Var[\varepsilon_1^{(1)}] + 2h^{-1}Cov(Z_h^{(h)}, \varepsilon_h^{(h)}).$$

Thus, by (A.5),

$$Cov(U_1(h), \varepsilon_1^{(1)}) = 2\mu h Var[\varepsilon_1^{(1)}] + 2h^{-1}o(h) = o(1),$$

i.e., (A.19) and in turn (19).■

References

Aït-Sahalia, Y., L.P. Hansen and J. Scheinkman (2003), "Operator Methods for Continuous-Time Markov Models," in Y. Ait-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics,* forthcoming.

Andersen, T.G., L. Benzoni and J. Lund (2002), "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239-1284.

Andersen, T.G. and T. Bollerslev (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885-905.

Andersen, T.G., T. Bollerslev and F.X. Diebold (2003), "Parametric and Nonparametric Volatility Measurement," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics,* forthcoming.

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

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003), "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579-625.

Andersen, T.G., T. Bollerslev, F.X. Diebold and C. Vega (2003), "The Evolving Effects of Macroeconomic News on Global Stock, Bond and Foreign Exchange Markets," unpublished manuscript, University of Pennsylvania.

Andersen, T.G., T. Bollerslev, and N. Meddahi (2003), "Analytic Evaluation of Volatility Forecasts," *International Economic Review*, forthcoming.

Andreou, E. and E. Ghysels (2002), "Rolling-Sampling Volatility Estimators: Some New Theoretical, Simulation and Empirical Results," *Journal of Business and Economic Statistics,* 20, 363-376.

Bai, X., J.R. Russell, and G.C. Tiao (2000), "Beyond Merton's Utopia: Effects of Non-Normality and Dependence on the Precision of Variance Estimates using High-frequency Financial Data," unpublished manuscript, University of Chicago.

Barndorff-Nielsen, O.E., S.E. Graversen and N. Shephard (2003), "Power Variation and Stochastic Volatility: A Review and Some New Results," *Journal of Applied Probability*, forthcoming.

Barndorff-Nielsen, O.E. and N. Shephard (2001), "Non-Gaussian Ornstein-Uhlenbeck-based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society, B*, 63, 167-207.

Barndorff-Nielsen, O.E. and N. Shephard (2002a), "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, B*, 64, 253-280.

Barndorff-Nielsen, O.E. and N. Shephard (2002b), "Estimating Quadratic Variation Using Realised Variance," *Journal of Applied Econometrics*, 17, 457-477.

Barndorff-Nielsen, O.E. and N. Shephard (2003a), "Realised Power Variation and Stochastic Volatility," *Bernoulli*, 9, 243-265.

Barndorff-Nielsen, O.E. and N. Shephard (2003b), "How Accurate is the Asymptotic Approximation to the Distribution of Realised Volatility?," in D. Andrews, J. Powell, P. Ruud, and J. Stock (eds.) *Identification and Inference for Econometric Models: A Festschrift for Thomas J. Rothenberg*, Cambridge University Press, forthcoming.

Becker, E. (1998), "Théorèmes Limites pour des Processus Discrétisés," Thèse de Doctorat, Université de Paris VI.

Billingsley, P. (1994), *Probability and Measure*, Third Edition, New-York: John Wiley & Sons.

Black, F. (1976), "Studies in Stock Price Volatility Changes," in *Proceedings of the 1976 Business Meeting of the Business and Statistics Section, American Statistical Association*, 177-181.

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

Chen, X., L.P. Hansen and J. Scheinkman (2000), "Principal Components and the Long Run," unpublished manuscript, University of Chicago.

Chernov, M. (2003), "Empirical reverse Engineering of the Pricing Kernel," *Journal of Econometrics*, 116, 329-364.

Chernov, M., R. Gallant, E. Ghysels and G. Tauchen (2003), "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225-257.

Chong, Y.Y. and D. Hendry (1986), "Econometric Evaluation of Linear Macro-Economic Models," *Review of Economic Studies*, 53, 671-690.

Comte, F. and E. Renault (1998), "Long Memory in Continuous Time Stochastic Volatility Models," *Mathematical Finance*, 8, 291-323.

Dacorogna, M.M., R. Gencay, U.A. Müller, R.B. Olsen, and O.V. Pictet (2001), *An Introduction to High-Frequency Finance*, San Diego: Academic Press.

Garcia, R., E. Ghysels and E. Renault (2003), "Option Pricing Models," in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics,* forthcoming.

Hansen, L.P. and J. Scheinkman (1995), "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes," *Econometrica* , 63, 767-804.

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.

Hoffman-Jørgensen, J. (1994), *Probability with a View Toward Statistics*, Volume 1, Chapman and Hall Probability Series, New-York: Chapman & Hall.

Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, 42, 281-300.

Lepingle, D. (1976), "La variation d'Ordre p des Semi-Martingales," *Zeitschrift fur Wahrscheinlichkeitstheorie verw. Gebiete*, 36, 295-316.

Meddahi, N. (2001), "An Eigenfunction Approach for Volatility Modeling," CIRANO working paper, 2001s-70.

Meddahi, N. (2002), "A Theoretical Comparison Between Integrated and Realized Volatility," *Journal of Applied Econometrics*, 17, 479-508.

Mincer, J. and V. Zarnowitz (1969), "The Evaluation of Economic Forecasts," in J. Mincer (Ed.), *Economic Forecasts and Expectation*, National Bureau of Research, New York.

Nelson, D.B. (1990), "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7-39.

Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347-370.

Protter, P. (1995), "'Stochastic Integration and Differential Equations: A New Approach," Springer Verlag.

Wiggins, J.B. (1987), "Options Values under Stochastic Volatility: Theory and Empirical Estimates", *Journal of Financial Economics*, 19, 351-372.