A Theoretical Comparison Between Integrated and Realized Volatility*

Nour Meddahi†

Université de Montréal

First version: December 7, 2001

This version: March 2002

Abstract

In this paper, we provide both qualitative and quantitative measures of the precision of measuring integrated volatility by realized volatility for a fixed frequency of observation. We start by characterizing for a general diffusion the difference between realized and integrated volatility for a given frequency of observation. Then we compute the mean and variance of this noise and the correlation between the noise and the integrated volatility in the Eigenfunction Stochastic Volatility model of Meddahi (2001a). This model has as special cases log-normal, affine and GARCH diffusion models. Using previous empirical results, we show that the noise is substantial compared with the unconditional mean and variance of integrated volatility, even if one employs five-minute returns. We also propose a simple approach to capture the information about integrated volatility contained in the returns through the leverage effect. We show that in practice, the leverage effect does not matter.

Key words: integrated volatility, realized volatility, infinitesimal generator, eigenfunction stochastic volatility models, leverage effect, exact moments.

---

*The author would like to thank Torben Andersen, Elena Andreou, Tim Bollerslev, Jean-Marie Dufour, René Garcia, Benoit Perron, Eric Renault and Neil Shephard for helpful comments on an earlier draft, the editors, and especially Bryan Campbell and two anonymuous referees for comments that improve the exposition of the paper. The author acknowledges financial support from FCAR, IFM2, and MITACS, and is solely responsible for any remaining errors.

†Département de sciences économiques, CRDE, CIRANO, Université de Montréal, and CEPR. Address: C.P. 6128, succursale Centre-ville, Montréal (Québec), H3C 3J7, Canada. E-mail: nour.meddahi@umontreal.ca.

1 Introduction

Several recent studies highlight the advantage of using high-frequency data to measure volatility of financial returns. These include Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Ebens (2001), Andersen, Bollerslev, Diebold and Labys (2001a, ABDL hereafter), Barndorff-Nielsen and Shephard (2001b-d), Taylor and Xu (1997) and Zhou (1996); for a survey of this literature, Andersen, Bollerslev and Diebold (2001), Barndorff-Nielsen, Nicolato and Shephard (2002) and Dacorogna et al. (2001) should be consulted. Typically, when one is interested in volatility over, say, a day, then these papers propose to study the estimation of this volatility by the sum of the intra-daily squared returns, such as returns over five or thirty minutes. This measure of volatility is called the realized volatility. The theoretical justification for this approach is that when the length of the intra-daily returns tends to zero, the sum tends in probability to the quadratic variation of the underlying diffusion process (ABDL, 2001a; Barndorff-Nielsen and Shephard, 2001a; Comte and Renault, 1998). Quadratic variation plays a central role in the option pricing literature. In particular, when there are no jumps, quadratic variation equals the integrated volatility highlighted by Hull and White (1987).

An important characteristic of high-frequency data is the presence of microstructure effects (Bai, Russell and Tiao, 2001; Andreou and Ghysels, 2001). Therefore, using data at the highest available frequency to measure volatility is not necessarily the best approach since the measure may be contaminated by microstructure effects. The solution adopted in the literature is to consider intra-daily returns over an intermediate frequency. For instance, when ABDL (2001b) address the issue of forecasting volatility through realized volatilities, the latter are based on intra-daily returns over thirty minutes.

The main objective of the paper is to provide both qualitative and quantitative measures of the precision of measuring integrated volatility by the realized volatility for a given frequency. In particular, we characterize the relative quality of the measures when one moves from one frequency to another.

Throughout the paper, we will assume the underlying data generating process is a continuous time, continuous sample-path model. We will derive the properties of the difference between integrated volatility and the realized volatility computed with intra-daily returns for a given frequency. The random variable defined as the realized volatility minus the integrated volatility is denoted the noise in what follows.

We start by characterizing this noise term in a general setting. The form of the noise allows us to give three of its qualitative characteristics. First, the unconditional mean of the noise is nonzero if and only if the drift of the diffusion characterizing the asset returns is nonzero. Second, the noise is heteroskedastic. Moreover, its conditional variance is correlated with the integrated and realized volatilities. Third, the noise is correlated with the integrated volatility if and only if there is a leverage effect or the drift depends on the instantaneous volatility.

In order to quantify these three characteristics, we consider a specific, yet general, class of continuous-time models. We assume that the underlying continuous-time process is an Eigenfunction Stochastic Volatility (ESV) model as presented in Meddahi (2001a). This class contains most of the popular SV models; in particular, the log-normal model of Hull and White (1987) and Wiggins (1987), the square-root and affine models of Heston (1993) and Duffie, Pan and Singleton (2000) respectively, and the GARCH diffusion model of Nelson (1990). In this setting, we derive explicitly the mean and the variance of the noise and its correlation with integrated volatility.

These theoretical results complement those of Barndorff-Nielsen and Shephard (2001b). These authors provide two important theoretical results. They give in a general setting a Central Limit Theorem of the convergence of the realized volatility to the integrated volatility when the length of the intra-daily returns tends to zero. Thus they provide the speed of convergence and the asymptotic variance of the noise term. This variance is stochastic, even in the limit. In the second more specific result, they characterize the mean and variance of the noise when the underlying instantaneous variance process is a linear combination of stationary covariance and autoregressive processes as the positive Lévy processes of Barndorff-Nielsen and Shephard (2001a).[1] In both cases, the authors ruled out leverage effects and assumed a driftless model in the second case. Thus our results extend the second results of Barndorff-Nielsen and Shephard (2001b) to the case where the underlying diffusion process governing the volatility is general and where there is both leverage effect and drift. Moreover, we provide also the first-order limit of the mean and variance of the noise term. Therefore, while it is not a Central Limit Theorem, our results complement those of Barndorff-Nielsen and Shephard (2001b). In particular, we show that this first-order limit does not depend on the leverage effect. This may suggest that the asymptotic result of Barndorff-Nielsen and Shephard (2001b) holds also when there is a leverage effect.[2]

We quantify values for the mean and variance of the noise and its correlation with integrated volatility by taking explicit examples from the literature. These examples include: i) the GARCH diffusion models without drift and leverage effect used by Andersen and Bollerslev (1998) and Andreou and Ghysels (2001); ii) the affine models with drift and leverage effect estimated by Andersen, Benzoni and Lund (2001) on the S&P500; iii) the log-normal model with drift and leverage effect also estimated by Andersen, Benzoni and Lund (2001) on the S&P500.

[1] As advocated by these authors, their results hold when the variance process is a marginalization of a vector of factors, where this vector admits a Vector Autoregressive representation of order one, VAR(1). Andersen (1994) firstly introduced such models in discrete time and called them the Square-Root Stochastic Autoregressive Volatility (SR-SARV) models while Meddahi and Renault (1996, 2000) introduced them in continuous time and showed their robustness against temporal and cross-sectional aggregations.

[2] Simulation results reported in Barndorff-Nielsen and Shephard (2002) suggest that this result holds also in the multivariate case. Note that in the previous version of this paper, Meddahi (2001c), we also provided some theoretical results in the multivariate case.

The main findings of empirical illustrations may be summarized as follows. First, the mean of the noise is very small relative to the mean of integrated volatility when one uses intra-daily observations. In particular, it is smaller (in absolute value) than .2%. Second, the standard deviation of the noise is relatively important with respect to the mean and the standard deviation of integrated volatility. In particular, when one uses realized volatility based on returns at five (resp. thirty) minutes, the ratio of the standard deviation of the noise over the mean of the integrated volatility is around 10% (resp. 25%). At the same frequencies, the ratio of the variance of the noise over the variance of the integrated volatility is around 5% (resp. 10%) and some times much more. These two ratios suggest that the noise is important even when one considers five-minute returns. Third, under leverage effects, the autocorrelation between the noise and integrated volatility is very small. Finally, we find that by using the simpler first-order asymptotic approximation, one obtains results that are very close those ones obtained by using exact formulas.

We also suggest an approach to extract through the leverage effect information concerning integrated volatility contained in the returns. It turns out that, in practice, this additional information is negligible.

The paper is organized as follows. In section 2, we characterize the noise and discuss its qualitative properties. In section 3, we recap the main properties of the ESV models of Meddahi (2001a). In the fourth section, we compute explicitly the mean and variance of the noise and the correlation between the noise and the integrated volatility. At each step, we give an empirical illustration of the importance of these terms. Section 5 suggests a solution for extracting information about integrated volatility contained in the returns through the leverage effect. The last section concludes, while all the proofs are provided in the Appendices.

2 The relationship between integrated and realized volatility

In this section, we characterize the relationship between integrated and realized volatility by using Ito lemma. Consider S_t a continuous time process representing the price of an asset or the exchange rate between two currencies. Assume that it is characterized by the following stochastic differential equation:

$$d\log(S_t) = m_t dt + \sigma_t dW_t \tag{2.1}$$

where W_t is a standard Brownian process. We assume that m_t is general and may depend, for instance, on σ_t and $log(S_t)$. The process σ_t is also general and we allow for leverage effect; i.e., if one assumes that σ_t^2 is characterized by

$$d\sigma_t^2 = \tilde{m}_t dt + \tilde{\sigma}_t d\tilde{W}_t,$$

then we allow $d\tilde{W}_t$ to be correlated with dW_t. We assume here (without loss of generality) that the time t is measured in units of one day. Consider a real h such that $1/h$ is a positive integer and define the realized volatility $RV_t(h)$ by

$$RV_t(h) \equiv \sum_{i=1}^{1/h} r_{t-1+ih}^{(h)2} \tag{2.2}$$

where $r_{t-1+ih}^{(h)}$ is the return over the period $[t-1+(i-1)h; t-1+ih]$, given by

$$r_{t-1+ih}^{(h)} \equiv \log\left(\frac{S_{t-1+ih}}{S_{t-1+(i-1)h}}\right). \tag{2.3}$$

When h goes to zero, the realized volatility converges in probability. This limit, which is independent of the discretization over the period $[t-1,t]$, defines the quadratic variation of the process $log(S_t)$ over the period $[t-1,t]$. In our context, the quadratic variation equals the integrated volatility denoted by IV_t and defined by[3]

$$IV_t \equiv \int_{t-1}^{t} \sigma_u^2 du. \tag{2.4}$$

Barndorff-Nielsen and Shephard (2001b-d) provide an asymptotic theory of the convergence of the realized volatility to the integrated volatility. In particular, they show that given the information $\sigma(\sigma_u, t-1 \leq u \leq t)$, we have that[4]

$$\sqrt{h^{-1}}(RV_t(h) - IV_t) \longrightarrow \mathcal{N}(0, 2\int_{t-1}^{t} \sigma_u^4 du). \tag{2.5}$$

While $RV_t(h)$ converges to IV_t when $h \to 0$, the difference may be not negligible for a given h. In order to study the difference, define $\mu_{t-1+ih}^{(h)}$ and $\varepsilon_{t-1+ih}^{(h)}$ by

$$\mu_{t-1+ih}^{(h)} \equiv \int_{t-1+(i-1)h}^{t-1+ih} m_u du \quad \text{and} \quad \varepsilon_{t-1+ih}^{(h)} \equiv \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u. \tag{2.6}$$

It is clear that

$$r_{t-1+ih}^{(h)} = \mu_{t-1+ih}^{(h)} + \varepsilon_{t-1+ih}^{(h)}. \tag{2.7}$$

Therefore,

$$(r_{t-1+ih}^{(h)})^2 = (\mu_{t-1+ih}^{(h)})^2 \; + \; 2\mu_{t-1+ih}^{(h)}\varepsilon_{t-1+ih}^{(h)} \; + \; (\varepsilon_{t-1+ih}^{(h)})^2.$$

Hence,

$$(r_{t-1+ih}^{(h)})^2 = \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du + (\mu_{t-1+ih}^{(h)})^2 + 2\mu_{t-1+ih}^{(h)}\varepsilon_{t-1+ih}^{(h)} + \left((\varepsilon_{t-1+ih}^{(h)})^2 - \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du\right).$$

To understand the properties of the third term, it is useful to rewrite it in terms of a stochastic integral. This is the purpose of the following Proposition, where we use Ito's Lemma to characterize the noise defined as the difference between the realized and integrated volatilities:

[3]If one incorporates jumps in (2.1), then the quadratic variation will be equal to the integrated volatility plus an additional term due to the jumps.

[4]As advocated by Barndorff-Nielsen and Shephard (2001b), since the variance of the asymptotic error in (2.5) is random, (2.5) is a mixed Gaussian limit theory.

Proposition 2.1 Characterizing the noise. *Let h be a positive real such that $1/h$ is an integer, i an integer and consider the processes S_t, $RV_t(h)$, $r^{(h)}_{t-1+ih}$, IV_t, $\mu^{(h)}_{t-1+ih}$ and $\varepsilon^{(h)}_{t-1+ih}$ defined respectively in (2.1), (2.2), (2.3), (2.4) and (2.6). Then:*

$$(r^{(h)}_{t-1+ih})^2 = \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du + u^{(h)}_{t-1+ih}, \text{ where} \tag{2.8}$$

$$u^{(h)}_{t-1+ih} = (\mu^{(h)}_{t-1+ih})^2 \ + \ 2\mu^{(h)}_{t-1+ih}\varepsilon^{(h)}_{t-1+ih} + 2\int_{t-1+(i-1)h}^{t-1+ih} (\int_{t-1+(i-1)h}^{u} \sigma_s dW_s)\sigma_u dW_u. \tag{2.9}$$

Hence,

$$RV_t(h) = IV_t + u_t(h) \tag{2.10}$$

where

$$u_t(h) = \sum_{i=1}^{1/h} u^{(h)}_{t-1+ih}. \tag{2.11}$$

We draw the following corollary from the proposition concerning the properties of the noise:

Corollary 2.1

a- The mean of $u^{(h)}_{t-1+ih}$ and $u_t(h)$ are in general nonzero when the drift m_u is nonzero.

b- The noise terms $u^{(h)}_{t-1+ih}$ and $u_t(h)$ are in general heteroskedastic.

c- Under leverage effect, $u^{(h)}_{t-1+ih}$ and $u_t(h)$ are correlated with integrated volatility IV_t.

These properties are clearly implied by the structure of the noise in (2.9). Note that the nonzero mean of the noise is not in contradiction with the asymptotic result (2.5) of Barndorff-Nielsen and Shephard (2001b) that implies

$$lim_{h\to 0}\frac{E[u_t(h)]}{\sqrt{h}} = 0.$$

Moreover, the heteroskedasticity of the noise, also implicit in (2.5), is problematic since the difference between the integrated and realized volatilities will have a higher variance when the instantaneous variance σ_t^2 and integrated volatility IV_t are high. Finally, under leverage effect, if the drift m_u depends on the volatility, the mean of the second term in (2.9) is nonzero.

Most of the previous remarks are well known. For instance, Barndorff-Nielsen and Shephard (2001b) pointed out that the mean of the noise is nonzero and that the noise is heteroskedastic. However, the impact of the leverage effect is not considered in the literature. We will consider explicit examples in the fourth section to quantify the characteristics of $u_t(h)$.

It is important to observe that we adopt a different approach than Barndorff-Nielsen and Shephard (2001a-b) to study the finite sample properties of $u_t(h)$. Their proofs are done given the sample path of the volatility. However, they exclude leverage effects and assume that the drift is an affine function of the variance. Our proofs use Ito lemma to study the leverage effect and more general formulations of the drift.

3 Eigenfunction Stochastic Volatility Models

In this section, we recap the main properties of the Eigenfunction Stochastic Volatility (ESV) models introduced in Meddahi (2001a). These models provide a convenient tractable framework, where many well-known models can be represented and in which analytic calculations can readily be performed. We will give a brief introduction to the general class of models, before indicating how common volatility models can be rewritten in this form.

3.1 General theory

The most popular stochastic volatility models like log-normal (Hull and White, 1987; Wiggins, 1987), square-root (Heston, 1993) and GARCH diffusion (Nelson, 1990) models have the following form:

$$d\log(S_t) = m_t dt + \sigma_t\,[\sqrt{1-\rho^2}dW_t^{(1)} + \rho dW_t^{(2)}], \quad \text{with} \tag{3.1}$$

$$\sigma_t^2 = g(f_t).$$

Here f_t is a state variable with simple dynamics that is characterized by

$$df_t = \mu(f_t)dt + \sigma(f_t)dW_t^{(2)}; \tag{3.2}$$

$g(.)$ is a known and *ad hoc* function; and $W_t^{(1)}$ and $W_t^{(2)}$ are two independent standard Brownian processes. In particular, we can represent:
1- Log-normal model: $\sigma_t^2 = \exp(f_t)$, $df_t = k[\theta - f_t]dt + \sigma dW_t^{(2)}$;
2- Square-root model: $\sigma_t^2 = f_t$, $df_t = k[\theta - f_t]dt + \sigma\sqrt{f_t}dW_t^{(2)}$;
3- GARCH diffusion model: $\sigma_t^2 = f_t$, $df_t = k[\theta - f_t]dt + \sigma f_t dW_t^{(2)}$.

Instead of taking an *ad hoc* function $g(.)$, Meddahi (2001a) proposes a flexible functional approach. More precisely, he assumes that the variance process σ_t^2 is given by

$$\sigma_t^2 = \sum_{i=0}^{p} a_i E_i(f_t), \tag{3.3}$$

where p is an integer, potentially infinite; a_i, $i = 0, ..., p$, are real numbers; and $E_i(f_t)$ are the eigenfunctions of the infinitesimal generator associated with f_t; see Hansen, Scheinkman (1995) and Ait-Sahalia, Hansen and Scheinkman (2001) for a review. In Appendix B, we recap the definition of this operator and some related properties. Such functions have some interesting properties:
i) two eigenfunctions $E_i(f_t)$ and $E_j(f_t)$ associated with two different eigenvalues are orthogonal, and any nonconstant eigenfunction is centered:

$$E[E_i(f_t)E_j(f_t)] = 0 \text{ and } E[E_i(f_t)] = 0; \tag{3.4}$$

ii) any eigenfunction is an autoregressive process of order one, in general heteroskedastic:

$$\forall h > 0,\ E[E_i(f_{t+h}) \mid f_\tau, \tau \le t] = \exp(-\delta_i h)E_i(f_t); \tag{3.5}$$

iii) any square-integrable function g, i.e. $E[g(f_t)^2] < \infty$, may be written as a linear combination of the eigenfunctions, i.e.

$$g(f_t) = \sum_{i=0}^{\infty} a_i E_i(f_t) \quad \text{where} \quad a_i = E[g(f_t)E_i(f_t)] \quad \text{and} \quad \sum_{i=0}^{\infty} a_i^2 = E[g(f_t)^2] < \infty. \tag{3.6}$$

Therefore, $g(f_t)$ is the limit in mean-square of $\sum_{i=0}^{p} a_i E_i(f_t)$ when p goes to $+\infty$.[5]

These three properties explain the powerfullness of the ESV approach. Consider any function of current or future values of returns. Given the Markovian nature of the joint process $(Log(S_t), f_t)$, a conditional expectation of any transformation of this variable, like the variance, is a function of f_t. Therefore, by using the third property, one can expand this function onto the eigenfunctions. The autoregressive features of these eigenfunctions (second property) allow for ready computation of the dynamics of this function. Finally, given the first property, it is easy to compute the covariance of two functions.

3.2 Examples

3.2.1 The log-normal example

Consider the state variable f_t defined by, after a normalization,

$$df_t = -kf_t dt + \sqrt{2k}\ dW_t^{(2)}. \tag{3.7}$$

The eigenfunction associated with the Ornstein-Uhlenbeck process (3.7) are the Hermite polynomials H_i associated with the eigenvalues $\delta_i = ki$. These polynomials are given in Appendix B. Meddahi (2001a) shows that the log-normal model of Hull and White (1987) and Wiggins (1987) is an ESV model with

$$\sigma_t^2 = \sum_{i=0}^{\infty} a_i H_i(f_t), \quad \text{where} \quad a_i = \exp(\theta + \frac{\sigma^2}{4k}) \frac{(\sigma/\sqrt{2k})^i}{\sqrt{i!}}. \tag{3.8}$$

3.2.2 The square-root example

Consider the state variable f_t defined by, after a normalization,

$$df_t = k(\alpha + 1 - f_t)dt + \sqrt{2k}\sqrt{f_t}dW_t^{(2)} \text{ with } \alpha = \frac{2k\theta}{\eta^2} - 1. \tag{3.9}$$

The eigenfunctions associated with (3.9) are the Laguerre polynomials $L_i^{(\alpha)}(f_t)$ associated with the eigenvalues $\delta_i = ki$. The Laguerre polynomials are given in Appendix B. Meddahi (2001a) shows that the square-root model of Heston (1993) is an ESV model with

$$\sigma_t^2 = a_0 L_0^{(\alpha)}(f_t) + a_1 L_1^{(\alpha)}(f_t) \quad \text{where} \quad a_0 = \theta \quad \text{and} \quad a_1 = -\frac{\sqrt{\theta}\eta}{\sqrt{2k}}. \tag{3.10}$$

Note that this is also the case for the affine model of Duffie, Pan and Singleton (2000).

[5] Observe that we make a normalization assumption by specifying that $Var[E_i(f_t)] = 1$ for $i \neq 0$. Likewise, we assume that $E_0(f_t) = 1$.

3.2.3 The GARCH diffusion example

Consider the state variable f_t defined by

$$df_t = k(\theta - f_t)dt + \sigma f_t dW_t^{(2)}. \tag{3.11}$$

This process was first introduced by Wong (1964) and popularized by Nelson (1990). This process violates assumption **A2**, Appendix B. The main consequence is that in the expansion results (third property), one has to take an integral instead of a sum. We will not consider this approach in this paper. Instead, we assume that the variance is a GARCH diffusion model, i.e. $g(x) = x$, and that the second moment of the variance σ_t^2 is finite. These assumptions suffice to do all the calculations, since the first eigenfunction is an affine function given by

$$E_1(x) = \frac{\sqrt{1-\lambda}}{\theta\sqrt{\lambda}}(x - \theta) \quad \text{where} \quad \lambda = \sigma^2/2k, \tag{3.12}$$

and the variance depends only on E_0 and E_1. Indeed, we have:

$$\sigma_t^2 = a_0 E_0(f_t) + a_1 E_1(f_t) \quad \text{where} \quad a_0 = \theta \;\; \text{and} \;\; a_1 = \frac{\theta\sqrt{\lambda}}{\sqrt{1-\lambda}}. \tag{3.13}$$

Note that the second moment of the variance σ_t^2 is finite when λ is smaller than one. Andersen and Bollerslev (1998) and Andreou and Ghysels (2001) who consider this example also assume the existence of the second moment of σ_t^2 in order to use the weak GARCH results of Drost and Werker (1996).

3.3 The multifactor case

Meddahi (2001a) considers also the case where the variance is a function of several factors as in Bollerslev and Zhou (2001), Engle and Lee (1999) and Harvey, Ruiz and Shephard (1994) among others. Without loss of generality, we consider the two-factor case. Let $f_{1,t}$ and $f_{2,t}$ be two independent stochastic processes characterized by

$$df_{j,t} = \mu_j(f_{j,t})dt + \sigma_j(f_{j,t})dW_{j,t}, \;\; j = 1, 2, \tag{3.14}$$

where the eigenfunctions (resp eigenvalues) of the corresponding infinitesimal generator are denoted $E_{1,i}(f_{1,t})$ and $E_{2,i}(f_{2,t})$ (resp $\delta_{1,i}$ and $\delta_{2,i}$). Then the variance process σ_t^2 is defined by

$$\sigma_t^2 = \sum_{0\le i,j\le p} a_{i,j} E_{1,i}(f_{1,t}) E_{2,j}(f_{2,t}) \quad \text{where} \quad \sum_{0\le i,j\le p} a_{i,j}^2 < \infty.$$

It turns out that the properties of the eigenfunctions defined in (3.4), (3.5) and (3.6) also hold for the functions $E_{i,j}(f_t)$ defined by

$$E_{i,j}(f_t) \equiv E_{1,i}(f_{1,t}) E_{2,j}(f_{2,t}) \quad \text{where} \quad f_t \equiv (f_{1,t}, f_{2,t})'. \tag{3.15}$$

Hence, $E_{i,j}(f_t)$ are the eigenfunctions associated with the bivariate state variable $(f_{1,t}, f_{2,t})$.[6]

[6] See Chen, Hansen and Scheinkman (2000) for a general approach of eigenfunction modeling in the multivariate case.

4 Characteristics of the noise

In this section, we quantify the importance of the noise term. We start by computing its mean and variance. This generalizes Barndorff-Nielsen and Shephard (2001b) to allow a drift and leverage effect. At each step, we illustrate this importance by considering examples from the literature. Note that all the results that we will show later hold also when one considers a multifactor model without leverage effect.

4.1 Mean of the noise

We assume that the processes $\log(S_t)$, σ_t and f_t are defined by (3.1), (3.3) and (3.2). Besides, the drift m_u is assumed to be

$$m_u = \sum_{i=0}^{p} b_i E_i(f_u) \text{ with } \sum_{i=0}^{p} \mid b_i \mid < +\infty. \tag{4.1}$$

Observe that the condition (4.1) implies that $\sum_{i=0}^{p} b_i^2 < +\infty$ and, hence, we include any example where the drift is assumed to be a square-integrable function of f_t. In particular, if the drift is assumed to be an affine function of the variance, i.e.

$$m_u = c + d\sigma_u^2, \tag{4.2}$$

then the coefficients b_i are given by

$$b_0 = c + da_0, \quad b_i = da_i, \quad i \geq 1. \tag{4.3}$$

In the following propositions, for a given i, the reals $\{e_{i,j}\}$ and p_i are defined by the (L^2) expansion of $\sigma_t \sigma(f_t) E_i'(f_t)$ onto the eigenfunctions, i.e.

$$\sigma_t \sigma(f_t) E_i'(f_t) = \sum_{j=0}^{p_i} e_{i,j} E_j(f_t), \tag{4.4}$$

where $E_i'(.)$ is the first derivative of $E_i(.)$. Thus we assume that $\sigma_t \sigma(f_t) E_i'(f_t)$ is square-integrable. For further details, see Meddahi (2001b).

Proposition 4.1 Mean of the noise. *Let h be a positive real such that $1/h$ is an integer, and consider the processes $\log(S_t)$, σ_t, f_t, m_t, $RV_t(h)$, IV_t and $u_t(h)$ defined respectively in (3.1), (3.3), (3.2), (4.1), (2.2), (2.4) and (2.10). Then:*

$$E[u_t(h)] = hb_0^2 + \frac{2}{h}\left(\sum_{i=1}^{p} \frac{b_i(b_i + \rho e_{i,0})}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]\right). \tag{4.5}$$

As a consequence, when $h \to 0$, we obtain

$$E[u_t(h)] \sim h[b_0^2 + \sum_{i=1}^{p} b_i(b_i + \rho e_{i,0})]. \tag{4.6}$$

Corollary 4.1

a- The mean of $u_t(h)$ in nonzero when the drift is nonzero.

b- The mean of $u_t(h)$ depends on the leverage effect.

c- Equation (4.6) adds information to the asymptotic result (2.5) of Barndorff-Nielsen and Shephard (2001b) by implying that

$$\frac{E[u_t(h)]}{\sqrt{h}} \sim \sqrt{h}[b_0^2 + \sum_{i=1}^{p} b_i(b_i + \rho e_{i,0})] \text{ when } h \to 0.$$

To assess the importance of this mean, we consider the empirical results of Andersen, Benzoni and Lund (2001). These authors estimated several models on daily returns of the S&P500.[7] In particular, they estimated the square-root and log-normal models without and with leverage effect. They consider an affine drift as in (4.3). Although they rejected these models, we consider their empirical results in order to get a first-order approximation of the importance of (4.5) and (4.6).[8]

To measure the importance of the mean of the noise, we introduce the following criterion:

$$Ratio = 100\, \frac{E[u_t(h)]}{E[IV_t]}. \tag{4.7}$$

In other words, we consider the ratio, in percentages, of the mean of the noise term over the mean of the integrated volatility. Table 1 gives the results of the ratio by using the exact formula (4.5) and the asymptotic approximation (4.6).

Table 1. Mean of the noise: Affine and log-normal models with leverage

Model		Affine		Log-normal	
1/h	freq	Ratio-Ex	Ratio-As	Ratio-Ex	Ratio-As
1	day	.168	.168	.179	.179
24	1 hour	.00701	.00701	.00747	.00745
48	30 mn	.00351	.00351	.00373	.00373
96	15 mn	.00175	.00175	.00187	.00186
144	10 mn	.00117	.00117	.00124	.00124
288	5 mn	.000584	.000584	.000622	.000621

From Table 1,[9] it is clear that the results based on both exact and asymptotic formulae are the same. Moreover, the mean of the noise is almost the same in both affine and log-normal models. Finally, and more importantly, the mean of the noise is relatively negligible with intra-daily data, for instance when one uses returns based on hourly data or a higher frequency.[10]

[7] The sample period is 01/02/1953-12/31/1996.

[8] Models including jumps in (3.1) that we exclude in our study were not rejected by Andersen, Benzoni and Lund (2001). Note however that ESV models of Meddahi (2001a) can have jumps.

[9] The results of the log-normal model are based on the expansion (4.5) by taking the first 100 terms.

[10] As we mentioned in the introduction, we do not consider impact the of microstructure effects.

4.2 Variance of the noise

In the sequel, we will assume that the drift is constant, i.e.

$$m_u = b_0. \tag{4.8}$$

Proposition 4.2 Variance of the noise term. *Let h be a positive real such that $1/h$ is an integer, and consider the processes* $\log(S_t)$, σ_t, f_t, m_t, $RV_t(h)$, IV_t *and* $u_t(h)$ *defined respectively in (3.1), (3.3), (3.2), (4.1), (2.2), (2.4) and (2.10). Assume that the drift m_u is given by (4.8). Then:*

$$Var[u_{t+ih}^{(h)}] = 4a_0 b_0^2 h^3 + 8b_0 h\rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h] + 4\sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$

$$+8\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} \right]\right] \tag{4.9}$$

where $e_{i,j}$ and p_i are defined in (4.4) and under the convention

$$\frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i-\delta_j)} - \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} = -\frac{\exp(-\delta_i h)(1+\delta_i h) - 1}{\delta_i^2} \quad \text{when} \quad \delta_i = \delta_j.$$

Moreover, the random variables $u_{t+ih}^{(h)}$ are uncorrelated. Hence,

$$Var[u_t(h)] = \frac{Var[u_{t+ih}^{(h)}]}{h}. \tag{4.10}$$

Finally, when $h \to 0$, we have:

$$Var[u_t(h)] \;\sim\; h\; 2 \sum_{i=0}^{p} a_i^2. \tag{4.11}$$

Corollary 4.2

a- If there is neither drift nor leverage effect, $Var[u_t(h)]$ becomes

$$Var[u_t(h)] = \frac{4}{h}\left(\frac{a_0^2 h^2}{2} + \sum_{i=1}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h] \right).$$

b- We have $lim_{h\to 0} Var[u_t(h)] = 0$. Hence, realized volatility tends to integrated volatility in mean-square and in probability.

c- The asymptotic variance of the noise $u_t(h)$, i.e. $Var[\sqrt{h^{-1}}u_t(h)]$, does not depend on the leverage effect and equals the one of Barndorff-Nielsen and Shephard (2001b) since

$$E[\int_{t-1}^{t} \sigma_u^4 du] = \sum_{i=0}^{p} a_i^2. \tag{4.12}$$

Note that the formula established when there is neither leverage effect nor drift coincides with one of Barndorff-Nielsen and Shephard (2001b). Moreover, we already pointed out in section

2 the well-known convergence result. Finally, given that the asymptotic variance does not depend on the leverage effect, one can suggest that the asymptotic result of Barndorff-Nielsen and Shephard (2001b) holds also when there is leverage effect, at least unconditionally.

In order to quantify the importance of the variance of the noise for a given frequency, we consider several examples. We start by considering models without drift and leverage effect. The first examples are the square-root models estimated by Bollerslev and Zhou (2001) on daily exchange rate data using realized volatilities.[11] They estimated a model with one factor and another with two factors. They rejected only the first one. The second group of examples considered are the GARCH diffusions models of Andersen and Bollerslev (1998) and Andreou and Ghysels (2001). After that, we consider further examples with leverage effect where the drift is constant, in particular the square-root and log-normal models estimated by Andersen, Benzoni and Lund (2001).[12]

We use two criteria in order to measure the importance of the variance of the noise:

$$Ratio1 = 100 \, \frac{\sqrt{Var[u_t(h)]}}{E[IV_t]} \quad \text{and} \quad Ratio2 = 100 \, \frac{Var[u_t(h)]}{Var[IV_t]}. \tag{4.13}$$

The first criterion is clearly related to the length of the confidence interval of integrated volatility. The second criterion is appealing because of the randomness of the integrated volatility. Typically, when the noise is uncorrelated with the integrated volatility,[13] the variance of the realized volatility is the sum of the variances of the noise and the integrated volatility. This ratio is crucial when one considers filtering the integrated volatility from the realized volatilities.

Table 2-a. Variance of the noise: Affine model without leverage

		1 Fac.		2 Fac.		1 Fac.	2 Fac.	1 Fac.	2 Fac.
1/h	freq	Std-Ex	Std-As	Std-Ex	Std-As	Ratio1	Ratio1	Ratio2	Ratio2
1	day	1.29	1.31	.749	.753	249	149	295	2137
24	1 hour	.267	.268	.154	.154	51.7	30.5	12.7	89.9
48	30 mn	.189	.189	.109	.109	36.6	21.5	6.34	45.0
96	15 mn	.134	.134	.0768	.0768	25.9	15.2	3.17	22.5
144	10 mn	.109	.109	.0627	.0627	21.1	12.4	2.12	15.0
288	5 mn	.0773	.0773	.0444	.0444	14.9	8.80	1.06	7.50

In Table 2-a, we report the results based on the models estimated by Bollerslev and Zhou (2001). The first interesting result is that computing the standard deviation of the noise by using the exact formula or by using the asymptotic first order approximation is almost the same when one uses intra-daily data. Besides, the standard deviation of the noise is

[11] Galbraith and Zinde-Walsh (2001) and Maheu and McCurdy (2001) consider also estimation of GARCH models of Engle (1982) and Bollerslev (1986) by using realized volatilities.

[12] Notice that we do not take into account the affine term in the drift, i.e., we assume that the drift is constant.

[13] This holds when there is neither drift nor leverage effect; see the next subsection.

almost divided by two when one goes from the one factor model to the two factors one. Therefore, since the unconditional mean of the integrated volatility is almost the same for both models (.517 and .504 respectively), the first criterion is also divided by two when one goes from the one-factor model to the two-factor model. Consider the two factors model that was not rejected by Bollerslev and Zhou (2001). According to the first criterion, the results are 8.8% and 21.5% when one considers realized volatilities computed with five and thirty minutes returns respectively. This is clearly not negligible since it means that the length of the confidence interval of the integrated volatility is relatively large with respect to the integrated volatility. Of course this criterion does not take into account the dependence between the conditional standard deviation of the noise and the integrated volatility. Therefore, one has to be cautious in interpreting the results. However, by using the asymptotic theory developed in Barndorff-Nielsen and Shephard (2001b), Barndorff-Nielsen and Shephard (2001c) estimated empirically at each day the confidence interval of the integrated volatility and showed that its length is large and positively correlated with the integrated volatility.

Consider now the second criterion. For the two-factor model, the results are 7.5% and 45%, again when one considers realized volatilities computed with five and thirty minutes returns respectively. Again, this is not negligible, especially when one uses thirty minutes returns, and suggests that one has to filter the integrated volatility by using all the history of the realized volatility. Notice that for the one-factor model, the results for this criterion are relatively small since the integrated volatility is more volatile than for the two-factor model (the standard deviations are .751 and .162 respectively). Finally, the values of the second criterion are high even one uses thirty minutes returns for the following reason. In their inference procedure, Bollerslev and Zhou (2001) did not take into account the difference between the integrated and realized volatilities.[14] More precisely, they derived theoretical moment conditions for the integrated volatilities while they used the realized volatilities in the estimation procedure. By so doing, they incorporated the noise term in the variance process. Therefore, they obtained a high variance of the variance which influences crucially the second criterion. In particular, the variance of the variance clearly appears in (4.12) and (2.5).

Consider now the two GARCH diffusions models considered by Andersen and Bollerslev (1998) and Andreou and Ghysels (2001). They correspond to daily returns of DM-US$ and Yen-US$. The results are presented in Table 2-b.[15] Note again the small difference between using exact and first-order approximations for the standard deviation of the noise. Moreover, the results are almost the same for both DM-US$ and Yen-US$ returns. We consider only the results on DM-US$. The first criterion is still not negligible (around 10% with five-minute returns). Thus the length of the confidence intervals will be relatively important. However,

[14] An alternative estimation approach is considered by Barndorff-Nielsen and Shephard (2001b) that takes into account the noise.

[15] Note that the variance of the noise corresponds to the MSE computed by simulation in Andersen and Bollerslev (1998). The exact results are very close to their ones.

the second criterion is negligible at five minutes (2.37%) but not at thirty minutes (14.2%). In other words, filtering the integrated volatility when one uses realized volatility computed with thirty (resp. five) minutes returns will have a large (resp. small) impact on the quality of the measure of the integrated volatility.

Table 2-b. Variance of the noise: GARCH diffusion model

Model		DM-US$		Yen-US$		DM	Yen	DM	Yen
1/h	freq	Std-Ex	Std-As	Std-Ex	Std-As	Ratio1	Ratio1	Ratio2	Ratio2
1	day	1.07	1.07	.930	.934	168.	195.	681	421
24	1 hour	.219	.219	.191	.191	34.4	40.0	28.5	17.7
48	30 mn	.155	.155	.135	.135	24.3	28.3	14.2	8.84
96	15 mn	.109	.109	.0953	.0953	17.2	20.0	7.12	4.42
144	10 mn	.0893	.0893	.0778	.0778	14.0	16.3	4.75	2.95
288	5 mn	.0632	.0632	.0550	.0550	9.93	11.6	2.37	1.47

Consider now the results on the affine and log-normal models with leverage effect estimated by Andersen, Benzoni and Lund (2001) reported in Table 2-c. Consider the affine case. Again, the difference between the exact and asymptotic results is very small. The first criterion is still not negligible (around 10% with five minutes returns) while the second one is at five minutes (around 2.5%) but not at thirty minutes (around 15%). The same results hold for the log-normal model.

Table 2-c. Variance of the noise: Affine and log-normal models with leverage

Model		Aff.		Log-nor.		Aff.		Log-nor.	
1/h	freq	Std-Ex	Std-As	Std-Ex	Std-As	Ratio1	Ratio2	Ratio1	Ratio2
1	day	.891	.892	.992	.993	166.7	714.9	180.2	524.0
24	1 hour	.182	.182	.202	.203	34.1	29.8	36.8	21.8
48	30 mn	.129	.129	.143	.143	24.1	14.9	26.0	10.9
96	15 mn	.0910	.0910	.101	.101	17.0	7.45	18.4	5.46
144	10 mn	.0743	.0743	.0827	.0827	13.9	4.96	15.0	3.64
288	5 mn	.0526	.0526	.0585	.0585	9.82	2.48	10.6	1.82

4.3 Covariance between the noise and integrated volatility

Given that the variable of interest IV_t is observed with errors, it is important to characterize the covariance between IV_t and the noise term for estimation, filtering and forecasting purposes. This covariance is characterized in the following proposition.

Proposition 4.3 Covariance between the noise and the integrated volatility. *Let h be a positive real such that $1/h$ is an integer, and consider the processes $\log(S_t)$, σ_t, f_t, m_t, $RV_t(h)$, IV_t and $u_t(h)$ defined respectively in (3.1), (3.3), (3.2), (4.1), (2.2), (2.4) and (2.10). Assume that the drift m_u is given by (4.8). Then:*

$$Cov(u_{t+ih}^{(h)}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du) = 2b_0 h\rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$

$$+2\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)}\right]\right]. \quad (4.14)$$

Besides,

$$Cov(u_t(h), IV_t) = \frac{1}{h} Cov(u_{t+ih}^{(h)}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du). \quad (4.15)$$

Finally, when $h \to 0$, the correlation between the noise $u_t(h)$ and integrated volatility IV_t is

$$Corr(u_t(h), IV_t) = \mathrm{O}(h^{3/2}). \quad (4.16)$$

Corollary 4.3

a- As pointed out by Barndorff-Nielsen and Shephard (2001b), the correlation of the noise with integrated volatility is zero when leverage effect is zero.

b- The correlation between the noise and integrated volatility tends to zero very quickly as one increases the frequency of intra-daily observations.

In order to assess the importance of this correlation, we consider models with leverage effect. In Table 3, we report this correlation for the affine and log-normal models estimated by Andersen, Benzoni and Lund (2001). Note that we provide two results for the log-normal case. The first one corresponds to one estimated by Andersen, Benzoni and Lund (2001). Given that the correlation is positive, we also take different values for the drift by fixing the other parameters. The second results correspond to $b_0 = .0334$ while the first ones correspond to $b_0 = .0314$. The difference between these two values is statistically small since the standard deviation of this parameter reported by Andersen, Benzoni and Lund (2001) is .0057. Note that the correlation is negative for the second results.

These results clearly establish that the correlation between the noise and the integrated volatility is very small, for instance -3.643 e-06 when one considers returns at thirty minutes in the affine case. Thus, this correlation can be ignored.

Table 3. Correlation between the noise and the integrated volatility

1/h	freq	Affine	Log-normal	
1	day	-0.001209	0.000119	-3.10e-05
24	1 hour	-1.030e-05	1.14e-06	-1.37e-07
48	30 mn	-3.643e-06	4.05e-07	-4.76e-08
96	15 mn	-1.288e-06	1.43e-07	-1.65e-08
144	10 mn	-7.019e-07	7.82e-08	-8.87e-09
2880	5 mn	-2.461e-07	2.84e-08	-2.42e-09

5 Leverage effect and integrated volatility

It is well known since Nelson (1991) that equity returns are negatively correlated with their instantaneous volatility. This is the so-called leverage effect. Thus, it is natural to extract the information about integrated volatility contained in the current daily return. This is the main purpose of this section.

We propose a simple approach for doing this extraction. We make the linear regression of the integrated volatility onto realized volatility, the current daily return and a constant. We calculate explicitly the coefficients of the regression and the corresponding R2. Of course, the approach holds when one knows that the underlying continuous time model is a ESV one. This is not always the case. Again, in this case, our approach may be viewed as a benchmark.

Before doing this regression, we first regress integrated volatility onto realized volatility and a constant (under leverage effect or not). This is important because when realized volatility is a very noisy estimator of the integrated volatility, it may be optimal in terms of mean square error to consider a weighted estimator of the realized volatility and the unconditional expectation of the integrated volatility. The best weighted estimator is obviously the linear regression of the integrated volatility onto the realized volatility and a constant, i.e.

$$IV_t = a(h) + b(h)RV_t(h) + \eta_t(h). \tag{5.1}$$

It is obvious that

$$b(h) = \frac{Cov(IV_t, RV_t(h))}{Var[RV_t(h)]}, \quad a(h) = E[IV_t] - b(h)E[RV_t(h)] \tag{5.2}$$

and that the R2 of the regression (5.1) is given by

$$R2(h) = \frac{Cov(IV_t, RV_t(h))^2}{Var[IV_t]Var[RV_t(h)]}. \tag{5.3}$$

Observe that when IV_t and $u_t(h)$ are not correlated, we have:

$$b(h) = \frac{1}{1 + Var[u_t(h)]/Var[IV_t]}, \quad R2(h) = b(h) \quad \text{and} \quad \frac{Var[\eta_t(h)]}{Var[u_t(h)]} = R2(h). \tag{5.4}$$

Table 4. Combining realized volatility with a constant

Model		Aff. 2 fac.	GARCH DM	Aff.	with	lev.	Log-nor.	with	lev.
1/h	freq	b(h)	b(h)	b(h)	R2(h)	Ratio	b(h)	R2(h)	Ratio
1	day	.0447	.128	.122	.122	.123	.160	.160	.160
24	1 hour	.527	.778	.770	.770	.770	.821	.821	.821
48	30 mn	.690	.875	.870	.870	.870	.902	.902	.902
96	15 mn	.816	.934	.931	.931	.931	.948	.948	.948
144	10 mn	.870	.955	.953	.953	.953	.965	.965	.965
288	5 mn	.930	.977	.976	.976	.976	.982	.982	.982

We report in Table 4 different values of b(h), R2(h) and the ratio $Var[\eta_t(h)]/Var[u_t(h)]$ as h varies. Note that $Var[\eta_t(h)]$ is the MSE of the regression (5.1), while $Var[u_t(h)]$ is the MSE that we considered in the previous section, i.e., when one considers the realized volatility as a measure for integrated volatility.

We report the result of the affine model with two factors, the GARCH diffusion for the DM-US$, the affine and log-normal models with leverage effect. The results of the other models are similar. Note that for models without leverage effect, we report only one column for b(h), R2(h) and the ratio $Var[\eta_t(h)]/Var[u_t(h)]$ since they are the same (see (5.4)).[16] From Table 4, it is clear that it is better to combine the constant and the realized volatility when one considers intra-daily returns at fifteen minutes or more. The ratio of the MSEs suggests that this improvement is important even if one considers five minutes returns.

However, one has to be careful with this criterion when both measures are very good. Consider three random variables y, x_1 and x_2. Let m_1 (resp m_2) be the best linear regression of y given x_1 (resp x_2), $R2_1$ and MSE_1 (resp $R2_2$ and MSE_2) the corresponding R2 and MSE. Then it is easy to show that

$$\frac{MSE_1}{MSE_2} = \frac{1 - R2_1}{1 - R2_2}.$$

Thus, the ratio of the MSEs may be high (or small) when $R2_1$ and $R2_2$ are close to one, that is the explanatory variables x_1 and x_2 explain well the variable y. For instance, if $R_21 = .98$ and $R2_2 = .99$, then the ratio of the MSEs is two. This is exactly what happens in Table 4 when h is very small.

We now reconsider the leverage effect case and study the regression

$$IV_t = a_1(h) + b_1(h)RV_t(h) + c_1(h)r_t + \psi_t(h) \tag{5.5}$$

where r_t is the daily return defined by

$$r_t \equiv \log(S_t/S_{t-1}). \tag{5.6}$$

Again, we have the following relationships

$$\begin{bmatrix} b_1(h) \\ c_1(h) \end{bmatrix} = \left(Var\begin{bmatrix} RV_t(h) \\ r_t \end{bmatrix}\right)^{-1} \begin{bmatrix} Cov(RV_t(h), IV_t) \\ Cov(r_t, IV_t) \end{bmatrix}. \tag{5.7}$$

The theoretical formulas of $b_1(h)$ and $c_1(h)$ are provided in Appendix D. Table 5 gives different values of $b(h)$, $b_1(h)$ and $c_1(h)$ for the affine and log-normal models with leverage effect estimated by Andersen, Benzoni and Lund (2001). We see that the contribution of the daily return r_t in explaining integrated volatility is very small when one uses realized volatility computed with intra-daily data. The result is not very surprising perhaps given the convergence

[16]Under leverage effect, they are not the same. However, in practice, ther are very close; see Table 4.

of realized volatility to integrated volatility when the length of intra-daily returns tends to zero. However, it is somewhat surprising for the daily frequency case.

The main reason of this small contribution is the following. Consider the autocorrelation between integrated volatility and daily return. By using results in Appendix D, one gets in the affine case that

$$Corr(IV_t, r_t) = \rho\sqrt{\frac{\exp(-k) - 1 + k}{k}}. \tag{5.8}$$

By using empirical results of Andersen, Benzoni and Lund (2001), one gets that this correlation is around (-.1). Thus the R2 of the regression of integrated volatility onto the daily return and a constant is 1%, which is very low. This is in contrast with the high (in absolute value) conditional correlation between the continuous-time innovations of the volatility and the price. The daily correlation is small since daily returns (and integrated volatility) are highly heteroskedastic and, hence, contain an important amount of noise not correlated with volatility. This is the same argument used by Andersen and Bollerslev (1998).

Table 5. Extracting the information from daily returns through the leverage effect

		Affine			Log-normal		
1/h	freq	$b_1(h)$	$c_1(h)$	R2(h)	$b_1(h)$	$c_1(h)$	R2(h)
1	day	.122	-.0162	.123	.160	-.0199	.162
24	1 hour	.770	-.00425	.770	.821	-.00424	.821
48	30 mn	.870	-.00240	.870	.901	-.00233	.902
96	15 mn	.931	-.00128	.931	.948	-.00123	.948
144	10 mn	.953	-.000876	.953	.965	-.000831	.965
288	5 mn	.976	-.000448	.976	.982	-.000423	.982

A potential limitation of this approach is that it could deliver negative integrated variance estimates. This is not the case in practice. The estimated integrated variance may be written as

$$\sum_{i=1}^{1/h}\left[b_1(h)r_{t-1+ih}^{(h)2} + c_1(h)r_{t-1+ih}^{(h)} + a_1(h)h\right]. \tag{5.9}$$

A sufficient condition ensuring the positivity of integrated volatility estimator is that each term in the sum (5.9) is positive. This positivity holds whatever the realization of $r_{t-1+ih}^{(h)}$ when $\Delta(h) = c_1(h)^2 - 4b_1(h)a_1(h)h$ si negative. It turns that this is the case when one uses the estimators we find for both affine and log-normal models. For instance, in the affine case, $\Delta(1) = -.229$, $\Delta(1/48) = -.005$ and $\Delta(h)$ tends toward zero when $h \to 0$.

6 Conclusion

In this paper, we provide qualitative and quantitative results about the characteristics of the difference between the realized and integrated volatility for a given frequency of observations and termed the noise. The main findings are threefold. First, under leverage effect or time varying drift, the mean of the noise is nonzero but negligible compared to the mean of the integrated volatility. Second, the noise is heteroskedastic and its standard deviation is not negligible with respect to the mean and the standard deviation of the integrated volatility even if one considers returns at five-minute intervals. Third, under leverage effect, the correlation of the noise with integrated volatility is nonzero but very small.

We also show that combining the realized volatility with the constant or some other variables reduces the noise. In particular, it is better to consider the linear regression of the integrated volatility on the constant and the realized volatility. Moreover, under leverage effect, we can add the daily return to extract the information that it contains about the integrated volatility through the leverage effect. It turns out that the improvement is small.

Our work can be extended in different directions. We need to take into account parameter uncertainty, since, in practice, these parameters must be estimated. Moreover, we ignore microstructures effects. A simple approach is to assume that one of the factors is a continuous-time Markov chain since such processes also admit an eigenfunctions decomposition.

Two other major extensions are currently under investigation. The first one incorporates jumps in the price or its volatility. Assuming the characteristics of the jumps, i.e. their intensity and sizes, are functions of the same state variable we consider will be very useful. This is exactly what happens in the affine models with jumps of Duffie, Pan and Singleton (2000). The second extension is related to the realized power variations considered by Barndorff-Nielsen and Shephard (2001d). These authors showed that $\sum_{i=1}^{1/h} \mid r_{t-1+ih}^{(h)} \mid^{\gamma}$ converges toward $\int_{t-1}^{t} \sigma_u^{\gamma} du$ when $h \to 0$ and provided a Central Limit Theorem as well. They also showed that the difference between the previous quantities, up to the asymptotic scaling parameter, is smaller when one considers γ equal to one (for instance) instead of two. Thus, finding the optimal γ that reduces the importance of the microstructure effects is of interest. Interestingly, the eigenfunctions expansion is still valid in this case and very easy indeed. For instance, if one considers the log-normal model, then we have

$$\sigma_t^{\gamma} = \sum_{i=0}^{\infty} a_i(\gamma) H_i(f_t), \quad \text{where} \quad a_i(\gamma) = \exp\left(\frac{\theta\gamma}{2} + \frac{\sigma^2\gamma^2}{16k}\right) \frac{(\sigma\gamma/\sqrt{8k})^i}{\sqrt{i!}}.$$

Thus, using our approach will be useful in this case.

References

Ait-Sahalia, Y., L.P. Hansen and J. Scheinkman (2001), "Operator Methods for Continuous-Time Markov Models," in Y. Ait-Sahalia and L.P. Hansen (Eds.), *Handbook of Financial Econometrics,* forthcoming.

Andersen, T.G. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling," *Mathematical Finance*, 4, 75-102.

Andersen, T.G., L. Benzoni and J. Lund (2001), " An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, forthcoming.

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

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 (2001), "Parametric and Nonparametric Measurements of Volatility," in Y. Ait-Sahalia and L.P. Hansen (Eds.), *Handbook of Financial Econometrics,* forthcoming.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001), "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001a), "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 (2001b), "Modeling and Forecasting Realized Volatility," unpublished manuscript, Northwestern University.

Andreou, E. and E. Ghysels (2001), "Rolling-Sampling Volatility Estimators: Some New Theoretical, Simulation and Empirical results," *Journal of Business and Economic Statistics,* forthcoming.

Bai, X., J. Russell and G.C. Tiao (2001), "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., E. Nicolato and N. Shephard (2002), " Some Recent Developments is Stochastic Volatility Modelling," *Quantitative Finance*, 2, 11-23.

Barndorff-Nielsen, O.E. and N. Shephard (2001a), "Non-Gaussian OU based Models and some of their uses in Financial Economics," *Journal of the Royal Statistical Society, B*, 63, 167-241.

Barndorff-Nielsen, O.E. and N. Shephard (2001b), "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, B*, forthcoming.

Barndorff-Nielsen, O.E. and N. Shephard (2001c), " Estimating Quadratic Variation Using Realised Variance," *Journal of Applied Econometrics*, forthcoming.

Barndorff-Nielsen, O.E. and N. Shephard (2001d), " Realised Power Variation and Stochastic Volatility," unpublished manuscript, Nuffield College, Oxford.

Barndorff-Nielsen, O.E. and N. Shephard (2002), " Econometric Analysis of Realised Covariation: High Frequency Covariance, Regression and Correlation in Financial Economics," unpublished manuscript, Nuffield College, Oxford.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. and H. Zhou (2001), "Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility," *Journal of Econometrics*, forthcoming.

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

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

Dacorogna, M.M., R. Gençay, U. Muller, R.B. Olsen and O.V. Pictet (2001), *An Introduction to High-Frequency Finance,* Academic Press.

Drost, F.C. and B.J.M. Werker (1996), "Closing the GARCH Gap: Continuous Time GARCH Modeling," *Journal of Econometrics*, 74, 31-58.

Duffie, D., J. Pan and K. Singleton (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1342-1376.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987-1007.

Engle, R.F. and G.G.J. Lee (1999), "A Permanent and Transitory Component Model for Stock Return Volatility," in *Cointegration, Causality and Forecasting: a festschrift in Honour of Clive W.J. Granger*, R.F. Engle and H. White eds., Chapter 20, 475-497, Oxford University Press.

Galbraith, J.W. and V. Zinde-Walsh (2001), "Properties of Estimates of Daily GARCH Parameters Based on Intra-Day Observations," CIRANO DP 2001s-15.

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

Hansen, L.P., J. Scheinkman and N. Touzi (1998), "Spectral Methods for Identifying Scalar Diffusions," *Journal of Econometrics*, 86, 1-32.

Harvey, A.C., E. Ruiz and N. Shephard (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247-264.

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.

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

Maheu, J. and T.H. McCurdy (2001), "Nonlinear Features of Realised FX Volatility," *Review of Economics and Statistics*, forthcoming.

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

Meddahi, N. (2001b), "Moments of Continuous Time Stochastic Volatility Models," in preparation.

Meddahi, N. (2001c), "A Theoretical Comparison Between Integrated and Realized Volatilities," CIRANO working paper, 2001s-71.

Meddahi, N. and E. Renault (1996), "Aggregation and Marginalization of GARCH and Stochastic Volatility Models," GREMAQ DP 96.30.433, Université de Toulouse and CRDE DP 3597, Université de Montréal.

Meddahi, N. and E. Renault (2000), "Temporal Aggregation of Volatility Models," *Journal of Econometrics*, forthcoming, and CIRANO working paper 2000s-22.

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.

Taylor, S.J. and X. Xu (1997), "The Incremental Volatility Information in One Million Foreign Exchange Quotations," *Journal of Empirical Finance*, 4, 317-340.

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

Wong, E. (1964), "The Construction of a Class of Stationary Markoff Processes," in Belleman, R., (Ed.), Sixteenth Symposium in Applied Mathematics, Stochastic Processes in Mathematical Physics and Engineering, American Mathematical Society, Providence, 264-276.

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

Appendix A

We start the Appendix by giving some lemmas.

Lemma A1 *Let f_t defined by (3.2). Then:*

$$dE_i(f_u) = -\delta_i E_i(f_u)du + E_i'(f_u)\sigma(f_u)dW_u^{(2)}, \tag{A.1}$$

$$\forall s < u, \quad E_i(f_u) = \exp(-\delta_i(u-s))E_i(f_s) + \exp(-\delta_i(u-s))\int_s^u \exp(\delta_i(w-s))\sigma(f_w)E_i'(f_w)dW_w^{(2)}, \tag{A.2}$$

$$E[\int_0^u E_i(f_u)\sigma_s dW_s] = \rho e_{i,0}\frac{1-\exp(-\delta_i u)}{\delta_i}, \tag{A.3}$$

where $e_{i,0}$ is given by $e_{i,0} = E[\sigma_s \sigma(f_s)E_i'(f_s)]$ (and defined in (4.4)). Thus,

$$E[\int_0^h (\int_0^u E_i(f_u)\sigma_s dW_s)du] = \rho e_{i,0}\frac{\exp(-\delta_i h) + \delta_i h - 1}{\delta_i^2}. \tag{A.4}$$

Proof: By Ito's Lemma, we get

$$dE_i(f_t) = \mathcal{A}E_i(f_t)dt + E_i'(f_t)\sigma(f_t)dW_t^{(2)}.$$

By defintion, $\mathcal{A}E_i(f_t) = -\delta_i E_i(f_t)$. Thus, we get (A.1).

Define z_u by $z_u = \exp(\delta_i u)E_i(f_u)$. By using Ito's Lemma we get $dz_u = \exp(\delta_i u)E_i'(f_u)\sigma(f_u)dW_u^{(2)}$. Hence, $z_u = z_s + \int_s^u \exp(\delta_i w)E_i'(f_w)\sigma(f_w)dW_w^{(2)}$. We then get (A.2).

We have:

$$\begin{aligned}
E\left[\int_0^u E_i(f_u)\sigma_s dW_s\right] &= E\left[\int_0^u \exp(-\delta_i(u-s))E_i(f_s)\sigma_s dW_s\right] \\
&\quad + E\left[\int_0^u \exp(-\delta_i(u-s))\left(\int_s^u \exp(\delta_i(w-s))\sigma(f_w)E_i'(f_w)dW_w^{(2)}\right)\sigma_s dW_s\right] \\
&= 0 + \int_0^u \exp(-\delta_i(u-s))[\exp(\delta_i(s-s))\sigma(f_s)E_i'(f_s)\sigma_s\rho ds] \\
&= \rho e_{i,0}\int_0^u \exp(-\delta_i(u-s))ds = \rho e_{i,0}\frac{1-\exp(-\delta_i u)}{\delta_i},
\end{aligned}$$

i.e. (A.3). From (A.3), one gets (A.4).□

Proof of Proposition 2.1. We have

$$r_{t-1+ih}^{(h)} = \int_{t-1+(i-1)h}^{t-1+ih} m_u du \; + \; \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u.$$

Therefore,

$$r_{t-1+ih}^{(h)2} = \left(\int_{t-1+(i-1)h}^{t-1+ih} m_u du\right)^2 + 2\left(\int_{t-1+(i-1)h}^{t-1+ih} m_u du\right)\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u\right) + \left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u\right)^2.$$

Let us consider $\int_0^h \sigma_u dW_u$ and compute its square by using Ito's Lemma. We have

$$(\int_0^h \sigma_u dW_u)^2 = 2\int_0^h (\int_0^u \sigma_s dW_s)\sigma_u dW_u + \int_0^h \sigma_u^2 du.$$

Hence,

$$\left(\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u\right)^2 = 2\int_{t-1+(i-1)h}^{t-1+ih}(\int_{t-1+(i-1)h}^{u} \sigma_s dW_s)\sigma_u dW_u + \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du.$$

As a consequence, we get (2.8) and, hence, (2.7).□

Appendix B

Let $\mathcal{A}$ be the infinitesimal generator oprerator associated with f_t:

$$\mathcal{A}\phi(f_t) \equiv \mu(f_t)\phi'(f_t) + \frac{\sigma^2(f_t)}{2}\phi''(f_t), \tag{B.1}$$

where $\phi(f_t)$ is a square-integrable function and twice differentiable. Let $E_i(f_t)$, $i = 0, 1, ...$, be the set of the eigenfunctions of $\mathcal{A}$ with corresponding eigenvalues $(-\delta_i)$, i.e.

$$\mathcal{A}E_i(f_t) = -\delta_i E_i(f_t). \tag{B.2}$$

Here, we assume that the eigenvalues are real numbers and that the spectrum, i.e. the set of the eigenvalues, is discrete:

Assumption A1. The stationary process $\{f_t\}$ is time reversible.

Assumption A2. The spectrum of the infinitesimal generator operator $\mathcal{A}$ of $\{f_t\}$ is discrete and denoted $\{-\delta_i, i \in \mathbf{N}\}$ with $\delta_0 = 0$ and $\delta_0 < \delta_1 < \delta_2 < ... < \delta_i < \delta_{i+1}...$; the corresponding eigenfunctions are denoted $E_i(f_t)$, $i \in \mathbf{N}$.

Hansen, Scheinkman and Touzi (1998) show that under some appropriate boundary protocol, stationary scalar diffusions are time-reversible. Hence, assumption **A1** is not restrictive when one considers a volatility model that depends on one factor. It is, however, when one considers a multivariate vector f_t. This assumption is ensured when the factors are independent as in the volatility literature. Assumption **A2** is true for both log-normal and square-root models but not for the GARCH diffusion model. A sufficient assumption that ensures **A2** is that the operator $\mathcal{A}$ is compact.

Hermite Polynomials:

$$H_0(x) = 1,\ H_1(x) = x \text{ and } \forall i > 1, H_i(x) = \frac{1}{\sqrt{i}}\{xH_{i-1}(x) - \sqrt{i-1}H_{i-2}(x)\}. \tag{B.3}$$

Laguerre Polynomials:

$$\begin{aligned} \begin{pmatrix} i+\alpha \\ i \end{pmatrix}^{1/2} i\ L_i^{(\alpha)}(x) = & \begin{pmatrix} i-1+\alpha \\ i-1 \end{pmatrix}^{1/2} (-x + 2i + \alpha - 1)L_{i-1}^{(\alpha)}(x) \\ & - \begin{pmatrix} i-2+\alpha \\ i-2 \end{pmatrix}^{1/2} (i + \alpha - 1)L_{i-2}^{(\alpha)}(x), \text{ where} \end{aligned} \tag{B.4}$$

$$L_0^{(\alpha)}(x) = 1, \quad L_1^{(\alpha)}(x) = \frac{1 + \alpha - x}{\sqrt{1+\alpha}}.$$

Appendix C

Proof of Proposition 4.1. Let μ_h and ε_h defined respectively by

$$\mu_h \equiv \int_0^h m_u du \quad \text{and} \quad \varepsilon_h \equiv \int_0^h \sigma_u dW. \tag{C.1}$$

By Ito's Lemma, we have:
$\mu_h^2 = 2\int_0^h \mu_u d\mu_u + \int_0^h d[\mu,\mu]_u = 2\int_0^h \mu_u m_u du = 2\int_0^h (\int_0^u m_s m_u ds) du$. Hence,
$E[\mu_h^2] = 2\int_0^h (\int_0^u E[m_s m_u] ds) du$. But, for $u \geq s$:
$E[m_s m_u] = \sum_{0\leq i,j\leq p} b_i b_j E[E_i(f_s)E_j(f_u)] = \sum_{0\leq i,j\leq p} b_i b_j \exp(-\delta_j(u-s)) E[E_i(f_s)E_j(f_s)]$
$= \sum_{i=0}^p b_i^2 \exp(-\delta_i(u-s))$. Hence,
$\int_0^u E[m_s m_u] ds = \sum_{i=0}^p \frac{b_i^2}{\delta_i}[1-\exp(-\delta_i u)]$. As a consequence,
$E[\mu_h^2] = 2\sum_{i=0}^p \frac{b_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$.
Let $\tilde{\mu}_h$ defined by

$$\tilde{\mu}_h \equiv 2\int_0^h m_u du \int_0^h \sigma_u dW_u. \tag{C.2}$$

By Ito's Lemma, we have:
$\tilde{\mu}_h = 2\int_0^h (\int_0^u \sigma_s dW_s) m_u du + 2\int_0^h (\int_0^u m_s ds)\sigma_u dW_u$. Hence,
$E[\tilde{\mu}_h] = 2E[\int_0^h (\int_0^u \sigma_s dW_s) m_u du] = 2\sum_{i=0}^p b_i \int_0^h E(\int_0^u E_i(f_u)\sigma_s dW_s)) du.$
$= 2\rho \sum_{i=0}^p \frac{b_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) + \delta_i h - 1]$ by (A.4).
The expectation of the third term in $u_{t+ih}^{(h)}$ is zero. Hence
$E[u_{t+ih}^{(h)}] = E[\mu_h^2] + E[\tilde{\mu}_h] = 2\sum_{i=0}^p \frac{b_i^2 + \rho b_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$
$= h b_0^2 + 2\sum_{i=1}^p \frac{b_i^2 + \rho b_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$ since $e_{0,0} = 0$.
Therefore we get (4.5) since $E[u_t(h)] = \frac{1}{h}E[u_{t+ih}^{(h)}]$.
For a small h, since $\delta_i \neq 0$, we have: $[\exp(-\delta_i h) - 1 + \delta_i h] \sim \delta_i^2 h^2/2$. Thus, (4.6) is deduced.□

Proof of Proposition 4.2. We have $Var[u_{t+ih}] = Var[b_0^2 h^2 + 2b_0 h\varepsilon_h + 2\tilde{Z}_h]$, where $\tilde{Z}_h \equiv \int_0^h (\int_0^u \sigma_s dW_s)\sigma_u dW_u$. Thus,
$Var[u_{t+ih}] = 4b_0^2 h^2 Var[\varepsilon_h] + 4Var[\tilde{Z}_h] + 8b_0 h Cov[\varepsilon_h, \tilde{Z}_h] = 4b_0^2 h^2 E[\varepsilon_h^2] + 4E[\tilde{Z}_h^2] + 8b_0 h E[\varepsilon_h \tilde{Z}_h]$.
We will compute the three terms:
i) We have: $E[\varepsilon_h^2] = E[\int_0^h \sigma_u^2 du] = a_0 h$.
ii) By Ito's Lemma, we have:
$\tilde{Z}_h^2 = 2\int_0^h \tilde{Z}_u\, d\tilde{Z}_u + \int_0^h d[\tilde{Z},\tilde{Z}]_u = 2\int_0^h \tilde{Z}_u\, (\int_0^u \sigma_s dW_s)\sigma_u dW_u + \int_0^h (\int_0^u \sigma_s dW_s)^2 \sigma_u^2 du.$
Therefore,

$E[\tilde{Z}_h^2] = E\left[\int_0^h (\int_0^u \sigma_s dW_s)^2 \sigma_u^2 du\right] = \int_0^h E\left[\left(2\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s + \int_0^u \sigma_s^2 ds\right)\sigma_u^2 du\right].$

We have to compute the two terms. Consider the second one. We have:

$$\int_0^h E\left[\left(\int_0^u \sigma_s^2 ds\right)\sigma_u^2\right] du = \sum_{i=0}^p a_i \int_0^h \left(\int_0^u E[E_i(f_u)\sigma_s^2] ds\right) du$$
$$= \sum_{i=0}^p a_i \int_0^h \left(\int_0^u \exp(-\delta_i(u-s)) E[E_i(f_s)\sigma_s^2] ds\right) du = \sum_{i=0}^p a_i^2 \int_0^h \left(\int_0^u \exp(-\delta_i(u-s)) ds\right) du$$
$$= \sum_{i=0}^p \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h].$$

Consider now the first term. Let us compute at a first step $E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right) E_i(f_u)\right]$.

This term is zero when $i = 0$ since $E_0(.) = 1$. For $i \neq 0$, we have:

$E[\tilde{Z}_u E_i(f_u)] = E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right) E_i(f_u)\right] = E\left[\int_0^u \left(\int_0^s \sigma_w dW_w\right) E_i(f_u)\sigma_s dW_s\right]$ by (A.2),

$$= E\left[\int_0^u \left(\int_0^s \sigma_w dW_w\right) \rho \exp(-\delta_i(u-s)) E_i(f_u)'\sigma(f_s)\sigma_s ds\right] \text{ by (4.4),}$$
$$= \rho E\left[\int_0^u \left(\int_0^s \sigma_w dW_w\right) \exp(-\delta_i(u-s))(\sum_{j=0}^{p_i} e_{i,j} E_j(f_s)) ds\right]$$
$$= \rho \sum_{j=0}^{p_i} e_{i,j} \int_0^u \left(\int_0^s E[E_j(f_s)\sigma_w dW_w]\right) \exp(-\delta_i(u-s)) ds$$
$$= \rho \sum_{j=0}^{p_i} e_{i,j} \int_0^u \left(\int_0^s E[\rho \exp(-\delta_j(s-w)) E_j(f_w)'\sigma(f_w)\sigma_w dw]\right) \exp(-\delta_i(u-s)) ds \text{ by (A.2),}$$
$$= \rho^2 \sum_{j=0}^{p_i} e_{i,j} \int_0^u \left(\int_0^s \exp(-\delta_j(s-w)) e_{j,0} dw\right) \exp(-\delta_i(u-s)) ds$$
$$= \rho^2 \sum_{j=1}^{p_i} e_{i,j} \int_0^u \left(\int_0^s \exp(-\delta_j(s-w)) e_{j,0} dw\right) \exp(-\delta_i(u-s)) ds \quad \text{since } e_{0,0} = 0$$
$$= \rho^2 \sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j}\left[\frac{1-\exp(-\delta_i u)}{\delta_i} - \frac{\exp(-\delta_j u) - \exp(-\delta_i u)}{\delta_i - \delta_j}\right]$$

Hence,

$$E[\int_0^h \tilde{Z}_u \sigma_u^2 du] = \int_0^h E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right)\sigma_u^2 du\right]$$
$$= \sum_{i=1}^p a_i \int_0^h E\left[\left(\int_0^u \left(\int_0^s \sigma_w dW_w\right)\sigma_s dW_s\right) E_i(f_u) du\right]$$
$$= \rho^2 \sum_{i=1}^p a_i \int_0^h \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j}\left[\frac{1-\exp(-\delta_i u)}{\delta_i} - \frac{\exp(-\delta_j u) - \exp(-\delta_i u)}{\delta_i - \delta_j}\right]\right] du$$

$$= \rho^2 \sum_{i=1}^p a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j}\left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i - \delta_j)}\right]\right]. \quad \text{(C.3)}$$

As a summary,

$$E[\tilde{Z}_h^2] = \sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$

$$+ 2\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} \right] \right] \quad \text{(C.4)}$$

iii) $\varepsilon_h \tilde{Z}_h = \int_0^h \tilde{Z}_u d\varepsilon_u + \int_0^h d\tilde{Z}_u \varepsilon_u + \int_0^h d[\tilde{Z}, \varepsilon]_u = \int_0^h \tilde{Z}_u \sigma_u dW_u + \int_0^h \varepsilon_u^2 \sigma_u dW_u + \int_0^h \sigma_u^2 \varepsilon_u du.$

Thus,

$E[\varepsilon_h \tilde{Z}_h] = E[\int_0^h \sigma_u^2 \varepsilon_u du] = \sum_{i=0}^{p} a_i \int_0^h \int_0^u E[E_i(f_u)\sigma_s dW_s]$

$= \sum_{i=0}^{p} a_i \int_0^h \rho \frac{e_{i,0}}{\delta_i}(1 - \exp(-\delta_i u))du = \rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$ since $e_{0,0} = 0$.

Hence,

$Var[u_{t+ih}] = 4a_0 b_0^2 h^3 + 8b_0 h \rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h] + 4 \sum_{i=0}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$

$+8\rho^2 \sum_{i=1}^{p} a_i \left[\sum_{j=1}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} \right] \right]$, i.e. (4.9).

The random variables $u_{t-1+(i-1)h}^{(h)}$ are uncorrelated since $\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u dW_u$ and $\int_{t-1+(i-1)h}^{t-1+ih} (\int_{t-1+(i-1)h}^{u} \sigma_s dW_s)\sigma_u dW_u$ are martingale difference sequences.

Thus, $Var[u_t(h)] = Var[u_{t-1+ih}^{(h)}]/h$.

For a small h, we have:

i) $[\exp(-\delta_i h) - 1 + \delta_i h] \sim \delta_i^2 h^2/2$;

ii) $\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} = \frac{h^3 \delta_j}{6} + o(h^3)$.

Hence, for a small h, the dominant term in (4.9) is the third one. Thus, we get (4.11).□

Proof of Proposition 4.3. We have:

$Cov(u_{t+ih}^{(h)}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du) = E[u_{t+ih}^{(h)} \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du] - E[u_{t+ih}^{(h)}]E[\int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du]$

$= E[u_{t+ih}^{(h)} \int_{t-1+(i-1)h}^{t-1+ih} \sigma_u^2 du] - h^3 a_0 b_0^2 = E[(b_0^2 h^2 + 2b_0 h \varepsilon_h + 2\tilde{Z}_h) \int_0^h \sigma_u^2 du] - h^3 a_0 b_0^2$

$= 2E[(b_0 h \varepsilon_h + \tilde{Z}_h) \int_0^h \sigma_u^2 du]$.

We have: $E[\varepsilon_h \int_0^h \sigma_u^2 du] = E[\int_0^h \varepsilon_u \sigma_u^2 du] = \sum_{i=0}^{p} a_i \int_0^h E[E_i(f_u) \int_0^u \sigma_s dW_s]$

$= \rho \sum_{i=1}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$ since $e_{0,0} = 0$ and (A.4).

Moreover: $E[\tilde{Z}_h \int_0^h \sigma_u^2 du] = E[\int_0^h (\int_0^u \sigma_u^2 du) d\tilde{Z}_u] + E[\int_0^h \tilde{Z}_u \sigma_u^2 du] = E[\int_0^h \tilde{Z}_u \sigma_u^2 du]$

$= \rho^2 \sum_{i=0}^{p} a_i \left[\sum_{j=0}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j} \left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)} \right] \right]$ by (C.3).

Hence,

$$Cov(u^{(h)}_{t+ih}, \int_{t-1+(i-1)h}^{t-1+ih} \sigma^2_u du) = 2b_0 h\rho \sum_{i=0}^{p} \frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$$
$$+2\rho^2 \sum_{i=0}^{p} a_i \left[\sum_{j=0}^{p_i} e_{i,j} \frac{e_{j,0}}{\delta_j}\left[\frac{h}{\delta_i} - \frac{1-\exp(-\delta_i h)}{\delta_i^2} - \frac{1-\exp(-\delta_j h)}{\delta_j(\delta_i - \delta_j)} + \frac{1-\exp(-\delta_i h)}{\delta_i(\delta_i-\delta_j)}\right]\right].\Box$$

Appendix D

In this Appendix, we compute some variables used in the text or the Tables.

1- Mean and variance of the integrated volatility: Meddahi (2001b) shows that:

$$E[IV_t] = a_0 \text{ and } Var[IV_t] = 2\sum_{i=1}^{p} \frac{a_i^2}{\delta_i^2}[\exp(-\delta_i) - 1 + \delta_i]. \tag{D.1}$$

2- Coefficients $e_{i,j}$: Meddahi (2001b) shows that:
For an affine model: $e_{1,0} = \sqrt{2k\theta}$ and $e_{1,1} = \sigma$.
For a log-normal model: $e_{i,0} = \sqrt{2k}\sqrt{i}\exp\left(\frac{\theta}{2} + \frac{\sigma^2}{16k}\right)\frac{(\sigma/\sqrt{8k})^{i-1}}{\sqrt{(i-1)!}}$ and

$$e_{i,j} = \sqrt{2k}\sqrt{i}\exp(\frac{\theta}{2} + \frac{\sigma^2}{16k})$$

$$\sum_{|j-i+1|\le s\le j+i-1;\frac{j-i+1+s}{2}\in\mathbf{N}} \frac{\sigma^s}{(8k)^{s/2}s!}\begin{pmatrix} s \\ \frac{j-i+1+s}{2}\end{pmatrix}\left(\prod_{q=0}^{(j-i+1+s)/2-1}(j-q)\prod_{q=0}^{(i-1-j+s)/2-1}(i-1-q)\right)^{1/2}.$$

3- Coefficients $b_1(h)$ $c_1(h)$ We have to compute the coefficients in the right part of (5.7). The coefficients $Cov(RV_t, IV_t)$ and $Var[RV_t(h)]$ are already computed.
i) $Cov(r_t, IV_t) = Cov(\int_{t-1}^t \sigma_u dW_u, \int_{t-1}^t \sigma_u^2 du) = E(\int_{t-1}^t \sigma_u dW_u \int_{t-1}^t \sigma_u^2 du)$
$= E(\int_{t-1}^t(\int_{t-1}^u \sigma_s dW_s)\sigma_u^2 du + \int_{t-1}^t(\int_{t-1}^u \sigma_s^2 ds)\sigma_u dW_u)$ by Ito's Lemma
$= E(\int_{t-1}^t(\int_{t-1}^u \sigma_s dW_s)\sigma_u^2 du) = \rho\sum_{i=1}^{p}\frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i) - 1 + \delta_i]$ by (A.4).
ii) $Cov(r_t, RV_t(h)) = Cov(r_t, IV_t) + Cov(r_t, u_t(h))$. But:
$Cov(r_t, u_t(h)) = Cov(\sum_{i=1}^{1/h} r^{(h)}_{t-1+ih}, \sum_{i=1}^{1/h} u^{(h)}_{t-1+ih}) = \frac{1}{h}Cov(r^{(h)}_{t-1+h}, u^{(h)}_{t-1+h}) = \frac{1}{h}E(r^{(h)}_{t-1+h}u^{(h)}_{t-1+h})$.
By using Ito's Lemma we get:
$E[r^{(h)}_{t-1+h}u^{(h)}_{t-1+h}] = E\left(\int_{t-1}^{t-1+h}\sigma_u dW_u\right)\left(\int_{t-1}^{t-1+h}(\int_{t-1}^u \sigma_s dW_s)\sigma_u dW_u\right)$
$= E\int_{t-1}^{t-1+h}(\int_{t-1+h}^u \sigma_s dW_s)\sigma_u^2 du = \rho\sum_{i=1}^{p}\frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$ by (A.4).
Hence, $Cov(r_t, u_t(h)) = \frac{\rho}{h}\sum_{i=1}^{p}\frac{a_i e_{i,0}}{\delta_i^2}[\exp(-\delta_i h) - 1 + \delta_i h]$.
iii) $Var(r_t) = E[\int_{t-1}^t \sigma_u^2 du] = a_0$.