Estimation of long memory in integrated variance

Eduardo Rossi
(Università di Pavia)

Paolo Santucci de Magistris
(Università di Padova e CREATEES)

# 17 (11-12)

Via San Felice, 5
I-27100 Pavia
http://epmq.unipv.eu/site/home.html

November 2012
Estimation of long memory in integrated variance

Eduardo Rossi ∗ Paolo Santucci de Magistris †
University of Pavia, Italy University of Padova and CREATEs

November 8, 2012

Abstract

A stylized fact is that realized variance has long memory. We show that, when the instantaneous volatility is a long memory process of order $d$, the integrated variance is characterized by the same long-range dependence. We prove that the spectral density of realized variance is given by the sum of the spectral density of the integrated variance plus that of a measurement error, due to the sparse sampling and market microstructure noise. Hence, the realized volatility has the same degree of long memory as the integrated variance. The additional term in the spectral density induces a finite-sample bias in the semiparametric estimates of the long memory. A Monte Carlo simulation provides evidence that the corrected local Whittle estimator of Hurvich et al. (2005) is much less biased than the standard local Whittle estimator and the empirical application shows that it is robust to the choice of the sampling frequency used to compute the realized variance. Finally, the empirical results suggest that the volatility series are more likely to be generated by a nonstationary fractional process.

Keywords: Realized variance, Long memory stochastic volatility, Measurement error, local Whittle estimator.

J.E.L: classification: C14, C22, C58.

∗ Dipartimento Economics and Management. Via San Felice 5, 27100 Pavia, Italy. Tel.: +39 0382/986207 Fax: +39 0382/304226. E-mail: erossi@eco.unipv.it

† Corresponding author: School of Economics and Management, Aarhus University, Bartholins Alle 10, 1322 University Park, DK-8000 Aarhus C, Denmark; phone: +45 89 421 547; email address: psantucci@creates.au.dk
1 Introduction

A well documented stylized fact is that the volatility of the financial returns is characterized by long-range dependence, or long memory, see, for instance, Baillie (1996), Bollerslev and Mikkelsen (1996), Dacorogna et al. (1993), Ding et al. (1993), Granger and Ding (1996). More recently Andersen et al. (2001a), Andersen et al. (2001b), Andersen et al. (2003), Martens et al. (2009) report evidence of stationary long memory in the realized variance (or realized volatility, \( RV \)) series.

In this paper, we theoretically study the long memory properties of the integrated variance (\( IV \)) and \( RV \), assuming that the instantaneous volatility, \( \sigma^2(t) \), is characterized by long memory of order \( d \).

The contributions of this paper are threefold. Firstly, we demonstrate that \( IV \) has the same fractional integration order of \( \sigma^2(t) \), since it has a pole at the zero frequencies that depends on the long memory parameter, \( d \). This result holds for both stationary and nonstationary long memory stochastic volatility models. Secondly, we show that when we consider sparse sampling and the presence of market microstructure noise, see Bandi and Russell (2008), Hansen and Lunde (2006) and for a recent survey McAleer and Medeiros (2008), the spectral density of \( RV \) is given by the spectral density of \( IV \) plus an additional constant term, which depends on the variance of the measurement error term. Therefore, \( RV \) is also a long-range dependent process and it has the same long memory of \( IV \) and \( \sigma^2(t) \). Moreover, in absence of microstructure noise, the spectral density of \( RV \) converges to that of \( IV \), as the sampling frequency increases.

Thirdly, we show by simulation that the local Whittle (LW) estimator of the long memory parameter is biased in finite samples as a consequence of the presence of the measurement error in the spectral density of \( RV \). In the context of our signal plus noise model, an alternative choice to the LW estimator is the corrected LW estimator of Hurvich et al. (2005), that explicitly accounts for the presence of the measurement error. We evaluate the impact that the choice of the sampling frequency and the variance of the measurement error have on the finite sample bias and the variability of both estimators. The results highlight
the dramatic reduction of the finite-sample bias of the corrected LW estimator with respect to the standard LW estimator.

The effectiveness of corrected LW estimator is also evident from the estimates of the degree of long memory of the volatility of 28 stocks traded on the NYSE. The results obtained with $RV$ at different sampling frequencies confirm the robustness of the corrected LW estimator to the presence of the measurement error, unlike the LW estimator, which is affected by the choice of the sampling frequency since this impacts on the variance of the measurement error. Indeed, the LW estimates decrease as the sampling frequency is getting smaller. The corrected LW estimates are not only larger than the uncorrected ones, but also rather constant with respect to the sampling frequency. This is particularly evident from the long memory signature plot of the $RV$, namely the plot of the long memory estimates obtained with $RV$ at different sampling frequencies. The corrected estimates are always larger than $1/2$, which is the upper bound of the stationary region. These results suggest that a deeper analysis of the nonstationary volatility models is called for.

This paper is organized as follows. In Section 2 we show that, when the instantaneous volatility is driven by a fractional Brownian motion, the degree of fractional integration of the $IV$ process is the same as the instantaneous volatility. Section 2.1 illustrates the characteristics of the measurement error. Section 3 discusses the semiparametric technique to obtain unbiased estimates of the long memory parameter of $IV$, based on a careful characterization of the spectral density of the realized variance. In Section 3.1 the results of the Monte Carlo simulations are illustrated and discussed. Section 4 compares the results obtained with the corrected and the uncorrected estimators of long memory, based on $RV$ estimated on the real data. Section 5 concludes.
2 Long memory in integrated and realized variance: theoretical results

Let \( P(t) \) be the price of an asset, where its logarithm, \( p(t) \), follows the stochastic differential equation:

\[
dp(t) = m(t)dt + \sigma(t)dW(t)
\]

where \( W(t) \) is a standard Brownian motion and \( m(t) \) is locally bounded and predictable. \( \sigma^2(t) \) is assumed to be independent of \( W(t) \) and càdlàg, see Barndorff-Nielsen and Shephard (2002a,b). Moreover, it is assumed that \( \sigma^2(t) \) is a long memory process, such that

\[
f_{\sigma^2}(\lambda) \sim c\lambda^{-2d} \quad \text{as} \quad \lambda \to 0
\]

where \( c \in \mathbb{R}_+ \) and \( f_{\sigma}(\lambda) \) is the spectral density of \( \sigma^2(t) \). When \( 0 < d < 1/2 \), the process \( \sigma^2(t) \) is stationary, while when \( 1/2 \leq d < 1 \), the process \( \sigma^2(t) \) is non-stationary, see Solo (1992), Velasco (1999a,b), Hurvich and Ray (1995, 2003) and Hurvich et al. (2005). In the nonstationary long memory case, the spectral density is not defined and it is replaced by the so called pseudo spectral density, which is the limit of the expectation of the sample periodogram. As noted by Hurvich and Ray (2003), the pseudo-spectral density plays a similar role as the ordinary spectral density in determining the properties of the periodogram, when \( d > 1/2 \), see also Hurvich et al. (2005, p.1288).

An example of a stationary long memory process for \( \sigma^2(t) \) that satisfies condition (2) is the fractional Ornstein-Uhlenbeck process of Comte and Renault (1998):

\[
d\ln \sigma^2(t) = -k \ln \sigma^2(t)dt + \gamma dW_d(t)
\]

where \( k > 0 \) is the drift parameter, while \( \gamma > 0 \) is the volatility parameter and
$W_d(t)$ is the fractional Brownian motion (fBm), which is defined\(^1\) as

$$W_d(t) = \frac{1}{\Gamma(1 + d)} \int_0^t (t - s)^d dW(s) + \int_0^0 [(t - s)^d - (-s)^d]dW(s) \quad (4)$$

The solution of (3) can be written as $\ln \sigma^2(t) = \int_0^t e^{-k(t-s)} \gamma dW_d(s)$. In particular, for $0 < d < 1/2$, Comte (1996) and Comte and Renault (1998) show that the process $\ln \sigma^2(t)$ has long memory of order $d$, if $k_\infty = \gamma d/k$ is finite and different from zero. Comte and Renault (1998) show that, when $k > 0$, the volatility process $\sigma^2(t)$ is asymptotically equivalent (in quadratic mean) to the stationary process\(^2\):

$$\hat{\sigma}^2(t) = \exp \left( \int_{-\infty}^t e^{-k(t-u)} \gamma dW_d(u) \right), \quad k > 0 \quad 0 < d < \frac{1}{2}, \quad (5)$$

They prove that the spectral density, $f_{\hat{\sigma}^2}(\lambda)$, of the process $\hat{\sigma}^2(t)$, is proportional to $\lambda^{-2d}$ as $\lambda \to 0$, so that the volatility process inherits the long-memory property induced by the fBm. More recently, Comte et al. (2010) show that also the class of affine fractional stochastic volatility models, satisfies condition (2).

In the following proposition, we show that the $IV$ has the same long memory degree of the instantaneous volatility.

**Proposition 1** Assuming that $\sigma^2(t)$ has a spectral density (or pseudo spectral density) that admits the representation in (2) around 0, then

$$\lim_{\lambda \to 0} \lambda^{2d} f_{IV}(\lambda) = c \in \mathbb{R}_+ \text{ where } f_{IV}(\lambda) \text{ is the spectral density (or pseudo spectral density) of } IV_t = \int_{t-1}^t \sigma^2(u)du.$$  

It is interesting to note that this result is proved under the assumption that $\sigma^2(t)$ has a spectral density (or pseudo spectral density) proportional to $\lambda^{-2d}$ at the origin as in (2). As a consequence of Proposition (1), at the origin the spectral density (or pseudo spectral density) of the $IV$ has the same behavior as that of the instantaneous volatility, so that $IV$ has the same degree of long memory as $\sigma^2(t)$. This result holds for stationary and nonstationary long memory instan-

---

\(^1\)The literature on long memory processes in econometrics distinguishes between type I and type II fractional Brownian motion. These processes have been carefully examined and contrasted by Marinucci and Robinson (1999) and Davidson and Hashimzade (2009).

\(^2\)The volatility process $\sigma^2(t)$ coincides almost surely with $\hat{\sigma}^2(t)$. 
taneous volatilities. Strictly speaking, finding a degree of long memory of $IV$ larger than $1/2$ implies an instantaneous volatility of the same fractional order. Although a model specification for the instantaneous volatility that allows for nonstationary long memory is potentially interesting and empirically relevant, nonstationary long memory continuous time stochastic volatility models are not yet investigated in literature.\textsuperscript{3} In Section 3.1, we propose a nonstationary long memory process for the instantaneous volatility, and we study by simulations the consequences of nonstationarity on the long memory of $IV$. In the next section, we analyze how the presence of long memory in the $IV$ translates in long-range dependence of the $RV$.

\section{The measurement error}

In this section we characterize the measurement error associated with the $RV$ estimator. To simplify the notation, we consider an equidistant partition $0 = t_0 < t_1 < \ldots < t_n = 1$, where $t_i = i/n$, and $\Delta = 1/n$, that is the interval is normalized to have unit length. Define $r_{i,\Delta} = p_{i\Delta,\Delta} - p_{(i-1)\Delta,\Delta}$. Adopting the notation of Hansen and Lunde (2005), the $RV^{\Delta}$ at sampling frequency $\Delta$ is

$$RV^{\Delta} = \sum_{i=1}^{n} r_{i,\Delta}^2$$

and Barndorff-Nielsen and Shephard (2002b) derived a distribution theory for $RV^{\Delta}$ when $n \to \infty$,

$$\sqrt{n}(RV^{\Delta} - IV) \xrightarrow{d} N(0, 2IQ),$$

where $IQ = \int_{0}^{1} \sigma^4(u) du$ is the integrated quarticity. In this paper we focus on the series of non-overlapping $IV$, $\{IV_t\}_{t=1}^{T}$, where $[0,T]$ represents our sampling period. Further, the time series of non-overlapping $RV^{\Delta}$ is composed by

$$\left\{ RV_t^{\Delta} = \sum_{i=1}^{n} r_{t,i,\Delta}^2 \right\}_{t=1}^{T},$$

where $r_{t,i,\Delta} = p_{t-1+i\Delta,\Delta} - p_{t-1+(i-1)\Delta,\Delta}$.\textsuperscript{4} We are interested in the estimation of the long memory of $IV_t$, using the observations on

\textsuperscript{3}For example, the estimates of long memory presented in Comte and Renault (1998) are larger than $1/2$.

\textsuperscript{4}In order to obtain non-overlapping $RV_t^{\Delta}$’s the first return included in the computation of $RV_t^{\Delta}$ is $r_{t,2,\Delta} = p_{t-1+2\Delta,\Delta} - p_{t-1+\Delta,\Delta}$.\textsuperscript{4}
2.1.1 No Noise Case

Barndorff-Nielsen and Shephard (2002b) and Meddahi (2002) characterize the discretization error, when $RV$ is used to measure the $IV$. While $RV^\Delta$ converges to $IV$ when $\Delta \to 0$, the difference may be not negligible for a given $\Delta > 0$. Following Barndorff-Nielsen and Shephard (2002b) and Meddahi (2002), we can decompose $RV^\Delta_t$, for a given $\Delta$, as

$$RV^\Delta_t = IV_t + u^\Delta_t. \quad (7)$$

with the discretization error equal to

$$u^\Delta_t = \sum_{i=1}^{n} u^\Delta_{t,i}, \quad (8)$$

where $u^\Delta_{t,i} = r^2_{t,i,\Delta} - \sigma^2_{t,i,\Delta}$ is the discretization error in the $i$-th subinterval, with $\sigma^2_{t,i,\Delta} = \int_{t-i+1}^{t-i+\Delta} \sigma^2(u)du$. When the drift $m(t)$ is non-zero, Meddahi (2002) proves that $u^\Delta_t$ has a non-zero mean. Furthermore, as pointed out by Barndorff-Nielsen and Shephard (2002b) and Meddahi (2002), the correlation between the $IV_t$ and the noise term is zero when there is no leverage effect, that is $dW$ in (1) and $dW_d$ in (3) are uncorrelated processes.

In the next Proposition we characterize the properties of the spectral density of $RV^\Delta$.

**Proposition 2** Consider the processes for $p(t)$ and $RV^\Delta_t$ defined in (1) and (6). Assume that condition (2) holds for $\sigma^2(t)$. Let $m(t) = \mu$ and assume no leverage effect, $\rho = \text{corr}(dW(t), dW_d(t)) = 0$,

i. For $\Delta > 0$, the spectral density (or pseudo spectral density) of $RV^\Delta_t$ is given by

$$f_{RV^\Delta}(\lambda) = f_{IV}(\lambda) + f_{u^\Delta}(\lambda) \quad (9)$$

and $\lim_{\lambda \to 0} \lambda^{2d} f_{RV^\Delta}(\lambda) = \lim_{\lambda \to 0} \lambda^{2d}(f_{IV}(\lambda) + f_{u^\Delta}(\lambda)) = c.$
ii. When \( \Delta \rightarrow 0 \), \( \text{Var}(u^\Delta_t) \rightarrow 0 \), then

\[
\lim_{\Delta \rightarrow 0} f_{RV^\Delta}(\lambda) = f_{IV}(\lambda).
\]  

(10)

And

\[
\lim_{\lambda \rightarrow 0} \left[ \lim_{\Delta \rightarrow 0} \lambda^{2d} f_{RV^\Delta}(\lambda) \right] = \lim_{\lambda \rightarrow 0} \lambda^{2d} f_{IV}(\lambda) = c.
\]  

(11)

For a given \( \Delta > 0 \), the spectral density (or pseudo spectral density) of \( RV^\Delta_t \) is equal to that of \( IV_t \) plus an additional term which depends on the variance of \( u^\Delta_t \). This means that \( RV^\Delta_t \) has the same degree of long memory of \( IV_t \) since \( f_u(\lambda) = \frac{\text{Var}(u^\Delta)}{2\pi} \) is constant with respect to \( \lambda \). In the ideal situation where prices are recorded continuously (\( \Delta \rightarrow 0 \)), the spectral density of \( RV^\Delta_t \) converges to that of the \( IV_t \) and, again, they share the same degree of long memory. It should be noted that the results in Proposition 2 extend to any stochastic volatility process for which the spectral density (or the pseudo spectral density) of \( IV_t \) exists. Note that the results in (9) and (11) are valid when \( m(t) = \mu \neq 0 \) in (1), since the discretization error remains an uncorrelated process. However, in this case, a closed from expression for \( \text{Var}[u^\Delta_t] \) becomes more involved than that provided by Barndorff-Nielsen and Shephard (2002a), because it depends also on \( m(t) \).

For example, when the instantaneous volatility follows the process in (3) and \( m(t) = 0 \), as in Barndorff-Nielsen and Shephard (2002a, p.257), the \( \text{Var}[u^\Delta_t] \) term is equal to

\[
\text{Var}[u^\Delta_t] = 2\Delta^{-1} \cdot \left\{ 2 \text{Var}[\hat{\sigma}^2(t)] \cdot \int_0^\Delta \int_0^v r(u) du dv + \Delta^2 E[\hat{\sigma}^2(t)]^2 \right\},
\]  

(12)

where

\[
E[\hat{\sigma}^2(t)] = \exp \left( \frac{\omega^2}{2} \right), \quad \text{Var}[\hat{\sigma}^2(t)] = [\exp(\omega^2) - 1] \exp(\omega^2)
\]  

(13)

with \( \omega^2 \equiv \text{Var}[\ln \sigma^2(t)] = \frac{\gamma^2}{\Gamma^2(1+\delta)k^{1+2\delta} \cos(\delta \pi)} \), see Casas and Gao (2008), and Meddahi (2002) derives a closed form expression for \( \text{Var}[u^\Delta_t] \) for the class of eigenfunction stochastic volatility models and assuming that \( m(t) = \mu \).
Estimation of Long Memory in Integrated Variance

\( r(\cdot) \) denotes the autocorrelation function of the process \( \hat{\sigma}^2(t) \). It is clear that the parameters in (3) affect \( E[\hat{\sigma}^2(t)] \), \( \text{Var}[\hat{\sigma}^2(t)] \), and \( \text{Cov}(\hat{\sigma}^2(t+h), \hat{\sigma}^2(t)) \), and through these, impact on the variance of the discretization error. It is hard to obtain closed-form expressions for the partial derivatives of \( \text{Var}[u^2_t] \) w.r.t. the parameters in (3), thus we investigate this point by simulations in Section 3.1.6

Differently from discrete-time stochastic volatility framework, where the variance of the measurement error is unrelated to the volatility parameters, in this setup the variance of the discretization error is a highly non-linear function of the instantaneous volatility process parameters. In the next section, we will characterize the measurement error and the spectral density of \( RV^\Delta_t \) when the prices are contaminated by the microstructure noise.

### 2.1.2 Microstructure Noise Case

Suppose now that the intraday price is observed with error, due to the presence of microstructure noise,

\[
\tilde{p}(t) = p(t) + \epsilon(t)
\]

where \( p(t) \) is the latent true, or efficient, price process that follows (1). The term \( \epsilon(t) \) is the noise around the true price, with mean 0 and finite fourth moment. In particular, \( \epsilon(t) \) is \( i.i.d. \) and it is independent of the efficient price and the true return process. Over periods of length \( \Delta \), we have

\[
\tilde{r}_{t,i,\Delta} = \left( p_{t-1+i}\Delta - p_{t-1+(i-1)\Delta}\Delta \right) + \left( \epsilon_{t,i,\Delta} - \epsilon_{t,i-1,\Delta} \right) = r_{t,i,\Delta} + \eta_{t,i,\Delta}.
\]

With discretization and microstructure noise, and \( m(t) = \mu \), the measurement error of \( RV^\Delta_t \) is given by

\[
\xi^\Delta = u^\Delta_t + \sum_{i=1}^n \eta^2_{t,i,\Delta} + 2 \left( \sum_{j=1}^n \sigma_{t,i,\Delta} z_{t,i,\eta_{t,i,\Delta}} \right) + 2 \Delta \mu \sum_{j=1}^n \eta_{t,i,\Delta}.
\]

As noted by Bandi and Russell (2006), while the efficient return is of order \( O_p(\sqrt{\Delta}) \), the microstructure noise is of order \( O_p(1) \) over any period of time.

---

6The double integral in (12) can only be approximated for \( \Delta \to 0 \).
This means, that, when \( \Delta \to 0 \), the microstructure noise dominates over the true return process, and longer period returns are less contaminated by the noise than shorter period returns. Given the properties of \( \epsilon(t) \), then

**Proposition 3** Consider the processes for \( p(t) \), \( \tilde{p}(t) \) and \( RV_t^\Delta \) defined respectively in (1), (14) and (6). Let \( m(t) = \mu \) and assume no leverage effect, \( \rho = \text{corr}(dW(t), dW_d(t)) = 0 \),

i. For \( \Delta > 0 \),

\[
f_{RV}^\Delta(\lambda) = f_{IV}(\lambda) + f_{\xi}(\lambda)
\]

where \( f_{\xi}(\lambda) = \frac{\text{Var}(\xi^\Delta t)}{2\pi} \) is the spectral density of the measurement error term. It follows that

\[
\lim_{\lambda \to 0} \lambda^{2d} f_{RV}^\Delta(\lambda) = c,
\]

with \( c > 0 \).

ii. For \( \Delta \to 0 \), \( \text{Var}(\xi^\Delta t) \to \infty \) and the \( f_{RV}^\Delta(\lambda) \to \infty \), \( \forall \lambda \).

Proposition 3 extends the results of Proposition 2, namely \( RV_t^\Delta \) is characterized by the same degree of long memory of the \( IV_t \) even when the microstructure noise is present and \( \Delta > 0 \). When prices are observed with a microstructure noise, the long memory signal, \( IV_t \), turns out to be contaminated by a measurement error, whose variance is given by the sum of two components: the variance of the discretization error \( \text{Var}(u_{t,i}^\Delta) \), and the term due to the presence of the microstructure noise, which is given by \( \Delta^{-1}(E(u_{t,i}^\Delta) - \sigma_{\eta}^4) + 4\Delta^{-1} E \left[ \sigma^2_{t,i,\Delta} \sigma_{\eta}^2 \right] + 4\Delta \mu^2 \sigma_{\eta}^2 \), see (37) in Appendix A.3. In accordance with Proposition 3, the effect of the microstructure noise on the variance of \( \xi^\Delta_t \), diverges as \( \Delta \to 0 \), see Bandi and Russell (2006), and dominates the long memory signal which cannot be identified anymore. However, for a given \( \Delta > 0 \), the \( \text{Var}(\xi^\Delta_t) \) is finite, and \( RV_t^\Delta \) has the same long memory degree of \( IV_t \). On the other hand, the choice of \( \Delta \) impacts on the variance of \( \xi^\Delta_t \) and through this on the spectral density of \( RV^\Delta \).

If we decrease \( \Delta \), this reduces the variability of \( \xi^\Delta_t \) due to the discretization but increases the microstructure noise component, so that the net effect on \( \text{Var}(\xi^\Delta_t) \) is unknown a priori. This trade-off, which depends on the choice of \( \Delta \), will be
studied via simulations in Section 3.1.

3 Bias corrected estimation of long memory in integrated variance

Now, we turn our attention to the estimation of the long memory parameter $d$ by means of periodogram-based estimators using a LW criterion function. It is well known that the drawback of global long memory estimators is that they require unnecessary assumptions on the spectral density. Instead, a consistent estimate of $d$ can be obtained simply by specifying the shape of the spectral density at the origin. These methods are referred as local methods. Further, the semiparametric approach has the advantage, over the parametric ones, that it does not require a full specification of the dynamics of the process, but simply characterization of the spectrum as $\lambda \to 0$. This implies that semiparametric estimation is more robust to the misspecification of the dynamics.

In our case, we are interested in the estimation of the long memory of $IV_t$ (or $\sigma^2(t)$), using the observations on $RV_t^\Delta$. According to Proposition 3, the spectrum of $RV_t^\Delta$ for $\Delta > 0$ is that of $IV_t$, plus an additional term, which depends on the variance of the measurement error, see (37). This is a typical signal-plus-noise problem. Consequently, the quality of the semiparametric estimate of $d$, based on spectrum of $RV_t^\Delta$, can be dramatically affected in finite samples by the variability of the measurement error. In a recent paper, Hansen and Lunde (2010) note that, "even with the most accurate estimators of daily volatility, which can utilize thousands of high-frequency prices, the standard error for a single estimate is rarely less than 10%.

In this section, we discuss the effect of the variance of the measurement error on the semiparametric estimates of $d$. A large literature, see among others Deo and Hurvich (2001), Hurvich et al. (2005) and Haldrup and Nielsen (2007), discusses the properties of the semiparametric long memory estimators, such as the log-periodogram regression and the LW estimator, when the long memory
signal is contaminated by a noise term. Deo and Hurvich (2001) show that the Geweke and Porter-Hudak (1984) estimator is biased by a constant factor that depends on the variance of the noise term. Sun and Phillips (2003) introduce an additional nonlinear term in the log-periodogram regression, proportional to $\lambda^{2d}$, to account for the effect of the additive noise term, that is allowed to be weakly dependent. Arteche (2004) suggests that an optimal choice of the bandwidth is important to minimize the influence of the added noise term, since the variance of the measurement error heavily restricts the allowable bandwidth in finite samples. With a larger variance of the noise with respect to the signal, only the frequencies very close to the origin contain a valuable information. Arteche (2004) and Hurvich et al. (2005) show that, in the signal-plus-noise framework, the LW estimator is consistent for $d \in (0, 1)$ under general assumptions on the noise term. However, in finite samples, the estimates are downward biased.

A possible solution to this problem is provided by Hurvich et al. (2005). They consider a semiparametric specification of the spectral density, allowing also for possible correlation between the noise and the signal. In particular, it is required that the signal has an infinite moving average representation with mild conditions on the coefficients. For example, when $\sigma^2(t)$ is generated by the process in (3), a closed form expression of the spectral density of the signal, the $IV_t$, is hard to obtain. However, in this case $IV_t$ is a long memory stationary process which can be approximated by an infinite moving average representation. The modified LW objective function of Hurvich et al. (2005) is

$$Q(G, d, \beta)^* = \frac{1}{m} \sum_{j=1}^{m} \left\{ \ln \left[ G\lambda_j^{-2d}(1 + \beta \lambda_j^{2d}) \right] + \frac{\lambda_j^{2d} I_{RV}(\lambda_j)}{G(1 + \beta \lambda_j^{2d})} \right\},$$  \hspace{1cm} (19)

\footnote{In this Section, we will maintain the assumption that the noise term is dynamically uncorrelated with the signal and it is a white noise. As shown in section 2.1, this is the relevant when drift in price and leverage are excluded.}

\footnote{Comte et al. (2010) analyze the affine fractional stochastic volatility models and characterize the autocovariance function of the expected $IV_t$ process.}
where \( G \) is the spectrum at the origin.\(^9\) Concentrating \( G \) out, it yields

\[
R(d, \beta) = \frac{1}{m} \sum_{j=1}^{m} \ln \left( \lambda_j^{-2d} (1 + \beta \lambda_j^{2d}) \right) + \ln \left( \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I_{RV} (\lambda_j^2) \right), \quad (20)
\]

where the LW estimator is obtained setting \( \beta = 0 \) in the minimization of \( R \). The LW estimates of \( d \) and \( \beta \) are

\[
(d_c, \hat{\beta}) = \arg \min_{(d, \beta) \in \mathbb{D} \times \mathbb{B}} \hat{R}(d, \beta) \quad (21)
\]

where \( \mathbb{D} \) and \( \mathbb{B} \) are the admissible sets of \( d \) and \( \beta \), and \( m \) has to tend faster to \( \infty \) than \( T^{4d/(4d+1)} \). In the case of \( RV \), \( \hat{\beta} \) is interpreted as an estimate of the noise-to-signal ratio, \( \frac{\text{Var}(\xi)}{2 \pi f_{IV}(0)} \). The corrected estimator is consistent for \( d \in (0, 1) \) and asymptotically normal for \( d \in (0, 3/4) \) with asymptotic variance, in absence of correlation between the signal and the noise, equal to \( \frac{(1+d)^2}{16d^2m} \), which is a decreasing function of \( d \). It is interesting to note that, for all the admissible values of \( d \), the asymptotic variance of the bias-corrected LW estimator, \( \hat{d}_c \), is larger than the corresponding asymptotic variance of the LW estimator, \( \hat{d} \), that is \( \frac{1}{4m} \).

### 3.1 Simulations

In this section we present the results of the Monte Carlo analysis of the finite sample properties of the long memory estimation of \( IV_t \) based on \( RV_t^\Delta \). We want to investigate the impact that the measurement error has on the LW and corrected LW estimators of the long memory of \( IV_t \), disentangling the contribution of the discretization error from that due to the microstructure noise. The purpose is to evaluate if the corrected LW estimator provides superior performances, in terms of bias for a large range of values of the noise-to-signal ratio \( (n, sr) \), and for different choices of \( \Delta \).

\(^9\)Frederiksen et al. (2012) and Nielsen (2008) suggest to approximate the log-spectrum of the short-memory component of the signal and of the perturbation by means of an even polynomial term.
3.1.1 Setup

We consider two alternative setups for the generation of $\sigma^2(t)$. The first is the stationary one presented in equation (3), which has been discussed so far. However, given that there is some empirical evidence that the volatility may be nonstationary, see for example Comte and Renault (1998) and Harvey (1998), we also simulate $\sigma^2(t)$ from a nonstationary long memory process.

We assume that the log-price $p(t)$ follows:

$$dp(t) = \sigma(t)dW(t)$$  \hfill (22)

and the instantaneous volatility process $\sigma^2(t)$ is either:

$$d\ln \sigma^2(t) = k(\psi - \ln \sigma^2(t))dt + \gamma dW_d(t)$$  \hfill (23)

or

$$d\ln \sigma^2(t) = \gamma dW_d(t)$$  \hfill (24)

where $W_d$ is the fractional Brownian motion of order $d$ independent of $W(t)$. To simulate increments from the fractional Brownian motion we implement the Matlab routine by Yingchun Zhou and Stilian Stoev\(^\text{10}\) which is based on the circulant embedding algorithm for the values of interest of the Hurst’s exponent, $H = d + \frac{1}{2}$.

The specification in (23) has a stationary solution, such that the long memory is equal to $d$, while the one in (24) generates nonstationary volatility trajectories.

In fact, when $\gamma > 0$, in the neighborhood of $\gamma = 0$, the first order approximation of $\exp \{ \gamma W_d(t) \}$ is $1 + \gamma W_d(t)$, so that, following Solo (1992), the pseudo spectral density of $\sigma^2(t)$ has a pole in zero which is proportional to $\lambda^{-2\delta}$, where $\delta = 1 + d$. Therefore, when $d \in (-1/2, 0)$, then $\delta \in (\frac{1}{2}, 1)$, so that $\sigma^2(t)$ is a nonstationary (but “mean reverting”) long memory process. According to Proposition 1, the $IV_t$ also has a pseudo spectral density proportional to $\lambda^{-2\delta}$ at the origin and it is integrated of order $\delta > \frac{1}{2}$. We will evaluate in simulation

\(^{10}\)http://www.stat.lsa.umich.edu/~sstoev/code/ffgn.m
if $IV_t$ has the expected long memory degree.

We simulate from the Euler approximation of (22), (23) and (24), a set of discrete trajectories with a time step of 10 seconds for 6.5 hours per day, which roughly corresponds to the trading period of NYSE. Thus we have a total $6 \times 60 \times 6.5 = 2,160$ log-prices and log-instantaneous volatilities per day for 2,500 days, that is $Y_j = \{p_j, \ln \sigma^2_j\}_{j=1}^{2,160 \times 2,500}$. The generated price series are used to compute the RV series, with different $\Delta$. Note that the computational burden is due to the fact that we simulate for each Monte Carlo replication a trajectory of $Y_j$ which has 5,400,000 observations. Therefore, we treat the instantaneous volatilities generated at 10 seconds frequency as the true latent process. We estimate the long memory parameter of the $IV_t$, which is unobserved in practice, but known in a simulation study and computed for the day $t$ as $IV_t = \sum_{k=1}^{2,160} \sigma^2_{(t-1) \cdot 2,160 + k}$ for $t = 1, \ldots, 2,500$. The estimate of the long memory parameter of $IV_t$ is a natural benchmark for those based on $RV^\Delta_t$. Moreover, when sampling prices at 10 seconds, and constructing the RV measure, $RV^\text{all}$, is equivalent to letting $\Delta \to 0$. On the other hand, sampling at 1, 5, 10 and 30 minutes introduces the discretization error, mentioned in Section 2, which is the consequence of the sparse sampling. Finally when prices are recorded with noise and the sampling is at 1, 5, 10 and 30 minutes, then we have the joint effect of microstructure noise and sparse sampling.

The market microstructure noise is introduced in the simulations in the form of a bid-ask bounce, modeled as:

$$\tilde{p}(t) = p(t) + \frac{\zeta}{2} \mathbb{I}(t)$$

(25)

where $\zeta$ is the percentage spread, and the order-driven indicator variables $\mathbb{I}(t)$ are independently across $p$ and $t$ and identically distributed with $Pr\{\mathbb{I}(t) = 1\} = Pr\{\mathbb{I}(t) = -1\} = \frac{1}{2}$. This variable takes value 1 when the transaction is buyer-initiated, and $-1$ when it is seller-initiated. We adopt the simplest bid-ask bounce specification in order to make a comparison with the existing literature. Furthermore it is interesting to note that $d\tilde{p}(t)$ exhibits spurious volatility and
negative serial correlation, see for instance Nielsen and Frederiksen (2008). The parameter of the bid-ask spread, $\zeta$, is set according to the values found in Table 1 in Bandi and Russell (2006). We choose $\zeta = \{0.000, 0.001, 0.002\}$ which are common values to the most liquid stocks. Similar values for $\zeta$ are also used in Nielsen and Frederiksen (2008).

The two estimators of the long memory, the LW ($\hat{d}$) and the corrected LW ($\hat{d}_c$), are computed using the $RV_\Delta$ series obtained at different sampling frequencies, i.e. $\Delta = 10$ sec, 1 min, 5 min, 10 min and 30 min. In order to compare their finite sample performances, we compute, for each sampling frequency, the percentage relative bias from $S$ Monte Carlo simulations

$$\text{Bias}(\hat{d}) = \frac{100}{d} \left( \frac{1}{S} \sum_{s=1}^{S} (\hat{d}_s - d) \right),$$

and the RMSE

$$\text{RMSE}(\hat{d}) = \left( \frac{1}{S} \sum_{s=1}^{S} (\hat{d}_s - d)^2 \right)^{1/2}.$$  

\[26\]

\[27\]

3.1.2 Noise-to-signal ratio

A crucial quantity in the simulations is the $nsr$, $\frac{\text{Var}(\xi_\Delta)}{\text{Var}(IV_t)}$, which depends on the generating process parameters, as discussed in Section 2.1. To figure out this relationship, which obviously affects the simulation results, we use Monte Carlo simulations, and plot the simulated $nsr$ as a function of $d$, $\gamma$, $\zeta$, and $\Delta$. In this way, we can choose a combination of the structural parameters that resembles realistic values of the $nsr$, see for a discussion Meddahi (2002).

Figure 1(a) shows the simulated $nsr$ for different choices of $d$ and $\Delta$. It is evident that increasing $\Delta$ increases, for each choice of $d$, the $nsr$, provided that the microstructure noise is absent. For moderate choices of $\Delta$, 1 or 5 minutes, the impact of $d$ on the $nsr$ is rather limited.

When the $nsr$ is plotted for different $\gamma$’s, Figure 1(b), it is clear that, for a given $\Delta$, as $\gamma$ approaches 0, the innovation in the price process becomes the prevailing source of variability, so that the $nsr$ is shifted upwards. This is seen
simply noting that,

\[
\frac{\operatorname{Var}(\xi_{i,t}^\Delta)}{\operatorname{Var}(\sigma_{t,i}^2\Delta^2)} = \frac{\operatorname{Var}(\xi_{i,t}^\Delta)}{E[(\sigma_{t,i}^2\Delta^2)^2] - E[(\sigma_{t,i}^2\Delta^2)]^2}.
\]

As \(\gamma \to 0\), then \(E[(\sigma_{t,i}^2\Delta^2)^2] - E[(\sigma_{t,i}^2\Delta^2)]^2 \to 0\), so that \(\frac{\operatorname{Var}(\xi_{i,t}^\Delta)}{\operatorname{Var}(IV_{t})} \to \infty\), where \(\xi_{i,t}^\Delta\) is defined in (35). The nsr is increasing in \(\Delta\), starting from 1 minute frequency, while for 10 seconds the microstructure noise dominates.

In Figure 1(c), the nsr is plotted for different values of \(\zeta\), which is the bid-ask spread. It is fairly evident that sampling at 10 seconds, introduces a large microstructure noise such that the variance of the signal is totally dominated by the noise term. For \(\zeta = 0.001\) and \(\gamma = 0.5\), the nsr is equal to 1.95, when \(\Delta = 10\) seconds, and is 1.36 for \(\Delta = 30\) minutes.

### 3.1.3 Stationary instantaneous volatility

For the stationary case we report the results corresponding to the following set of parameter values: \(k = 0.9\), \(\psi = -9.2\), \(\gamma = \{0.5, 0.7\}\) and \(d = 0.4\). The parameter \(\psi\) is the long-run mean of the log-volatility and \(\psi = -9.2\) corresponds to an annualized volatility of approximately 16%. This combination of parameters generates a nsr which, in absence of microstructure noise, ranges between 2% when \(\Delta = 1\) minute, to 96% when \(\Delta = 30\) minutes. When \(\Delta = 5\) minutes, the nsr is between 10\% (\(\gamma = 0.7\)) and 20\% (\(\gamma = 0.5\)), and is consistent with the findings in Meddahi (2002).

Table 1 reports the percentage bias and RMSE of the estimated long memory parameter when \(d = 0.4\), obtained with the LW and the corrected LW estimators, see (21), for different choices of \(\Delta\), \(\gamma\) and \(\zeta\). In both panels, the estimates of \(d\) based on the \(IV_t\) are, as expected, the closest to the true value, and the percentage bias, smaller than 1\%, is due to the Monte Carlo variance. However, in the real world, \(IV_t\) is unobservable and we rely on the realized measures to conduct inference on the degree of long memory. When \(\zeta = 0\), the best LW estimates of \(d\) are obtained using all available returns, while the largest
(a) Noise-to-signal ratio, $\text{Var}(\xi_t^\Delta)/\text{Var}(IV_t)$, as a function of $\Delta \in (10\text{ sec}, 30\text{ min})$, with $\gamma = 0.5$ and $\zeta = 0$. Each line corresponds to a different value of $d$, i.e., $d = \{0, 0.1, 0.2, 0.3, 0.4, 0.45\}$.

(b) Noise-to-signal ratio, $\text{Var}(\xi_t^\Delta)/\text{Var}(IV_t)$, as a function of $\Delta \in (10\text{ sec}, 30\text{ min})$, with $\zeta = 0.001$ and $d = 0.4$. Each line corresponds to a different value of $\gamma$, i.e., $\gamma = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$.

(c) Noise-to-signal ratio, $\text{Var}(\xi_t^\Delta)/\text{Var}(IV_t)$, as function of $\Delta \in (10\text{ sec}, 30\text{ min})$, with $\gamma = 0.5$ and $d = 0.4$. Each line corresponds to a different $\zeta$, i.e., $\zeta = \{0, 0.0005, 0.001, 0.0015, 0.0020, 0.0025\}$.

Figure 1: Simulated noise-to-signal ratio.
Table 1: Bias and root mean squared error of Monte Carlo estimates of $d$

(a) $d = 0.4$ and $\gamma = 0.5$.

<table>
<thead>
<tr>
<th></th>
<th>$\zeta = 0.000$</th>
<th>$\zeta = 0.001$</th>
<th>$\zeta = 0.002$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias $\hat{d}$</td>
<td>RMSE($\hat{d}$)</td>
<td>Bias $\hat{d}_c$</td>
</tr>
<tr>
<td>$IV$</td>
<td>-0.67</td>
<td>0.0546</td>
<td>-0.67</td>
</tr>
<tr>
<td>$RV_{all}$</td>
<td>-0.74</td>
<td>0.0546</td>
<td>2.96</td>
</tr>
<tr>
<td>$RV^1$</td>
<td>-1.17</td>
<td>0.0546</td>
<td>2.50</td>
</tr>
<tr>
<td>$RV^5$</td>
<td>-3.09</td>
<td>0.0558</td>
<td>0.73</td>
</tr>
<tr>
<td>$RV^{10}$</td>
<td>-5.45</td>
<td>0.0583</td>
<td>-0.25</td>
</tr>
<tr>
<td>$RV^{30}$</td>
<td>-13.98</td>
<td>0.0778</td>
<td>-1.28</td>
</tr>
</tbody>
</table>

(b) $d = 0.4$ and $\gamma = 0.7$.

<table>
<thead>
<tr>
<th></th>
<th>$\zeta = 0.000$</th>
<th>$\zeta = 0.001$</th>
<th>$\zeta = 0.002$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias $\hat{d}$</td>
<td>RMSE($\hat{d}$)</td>
<td>Bias $\hat{d}_c$</td>
</tr>
<tr>
<td>$IV$</td>
<td>-0.94</td>
<td>0.0490</td>
<td>-0.94</td>
</tr>
<tr>
<td>$RV_{all}$</td>
<td>-1.02</td>
<td>0.0490</td>
<td>-0.82</td>
</tr>
<tr>
<td>$RV^1$</td>
<td>-1.35</td>
<td>0.0491</td>
<td>-0.17</td>
</tr>
<tr>
<td>$RV^5$</td>
<td>-2.98</td>
<td>0.0501</td>
<td>1.17</td>
</tr>
<tr>
<td>$RV^{10}$</td>
<td>-4.59</td>
<td>0.0521</td>
<td>1.54</td>
</tr>
<tr>
<td>$RV^{30}$</td>
<td>-11.29</td>
<td>0.0662</td>
<td>-0.41</td>
</tr>
</tbody>
</table>

Notes: $\hat{d}$ denotes the LW estimator of the long memory parameter, while $\hat{d}_c$ is the corrected LW estimator, see (21). The term Bias is referred to the relative percentage bias, defined in (26)-(27). The estimates are based on 1,000 samples of 2,500 daily observations from model (22)-(23) with parameter values indicated in the table and discretization step set to 10 seconds. The bandwidth used in the estimation of $d$ is $m = T^{0.65}$. 
negative biases are those obtained with $RV^{30}$. The negative bias becomes larger as $\gamma$ gets smaller, a result of the increase in the $nsr$ ratio, displayed in Figure 1(b). This is coherent with the fact that the only source of noise in this case is the discretization error, so that increasing the latter, produces more biased estimates. When $\zeta = 0.001$ or $\zeta = 0.002$, the largest negative bias of LW is that obtained with $RV^{all}$ (between -24% and -33%), while the bias of $RV^{30}$ is between -13% and -18%. In presence of microstructure noise, the best LW estimates of $d$ are obtained sampling at 1 and 5 minutes, and the bias is approximately $-10\%$, so that the average of the Monte Carlo estimates is approximately 0.36. Interestingly, correcting for the presence of the measurement error improves the quality of the estimates, in terms of bias and RMSE, for any choice of $\Delta$, and the relevance of the correction becomes evident as $\zeta$ increases. It is noteworthy that, despite the corrected LW estimator is asymptotically less efficient then the LW estimator, the RMSE of $\hat{d}_c$ is often smaller than that of $\hat{d}$. This means that the squared bias component of the RMSE prevails on the variance component, thus confirming the relevance of correcting for the measurement error.

### 3.1.4 Nonstationary instantaneous volatility

The simulated trajectories for the nonstationary $\sigma^2(t)$ are obtained setting $\gamma = 0.2$ and $d = \{-0.3, -0.4\}$. This implies that the fractional integration order, $\delta = 1 + d$, of the instantaneous and $IV_t$ is 0.7 and 0.6, respectively. We initialize each simulated path with $p(0) = \ln(100)$ and $\sigma^2(0) = \exp(\psi) > 0$. Figure 2 reports a simulated trajectory of the $IV_t$ under non-stationarity; it appears very realistic and, in particular, the model is able to generate long periods of high volatility which are typical of financial turmoils.

In the nonstationary setup, with $\delta = 0.6$, the $nsr$ ranges between 10% when $\Delta = 1$ minute, to 300% when $\Delta = 30$ minutes. With $\delta = 0.7$, the $nsr$ ranges between 3% when $\Delta = 1$ minute, to 82% when $\Delta = 30$ minutes. The reduction in the $nsr$ obtained when $\delta = 0.7$ is due to higher persistence of the signal. The $nsr$ corresponding to a $\Delta$ equal to 5 minutes, is 15%.

A similar evidence to that found in the stationary case emerges also when
the $\sigma^2(t)$ is long memory but nonstationary, which could be the relevant case in practice, see Table 2. Firstly, the LW estimate, $\hat{\delta}$, based on $IV_t$ is very close to the value $\delta = 1 + d$, which is the same as that of $\sigma^2(t)$. This is in accordance with Proposition 1, for long memory orders in the range $(0, 1)$. When $\zeta = 0$ and $\gamma = 0.2$, the variance of the noise dominates the signal as $\Delta$ increases, hence, the impact of the discretization error on the LW estimates of $\delta$ is very large. For example, when $\zeta = 0$ and $\delta = 0.6$, the bias of $\hat{\delta}$ based on $RV^{30}$ is negative and larger than 30%, and larger than 18% when $\delta = 0.7$.

As expected, the negative bias increases, as $\zeta$ increases, and we observe extremely large negative biases for $\zeta = 0.002$, so that $\hat{\delta}$, based on the $RV$, falls in the stationary region, even though the $IV_t$ is not stationary. For example, $RV^5$ has a negative bias equal to -28% when $\delta = 0.6$, meaning that $\hat{\delta} \approx 0.43$ on average. On the contrary, the corrected LW estimator, $\hat{\delta}_c$, provides unbiased estimates also in the nonstationary region, for both $\delta = 0.6$ and $\delta = 0.7$, and for all choices of $\Delta$ and $\zeta$.

3.1.5 Leverage

Finally, the last set of simulations, reported in Table 3, investigates the impact on the estimates of $d$ of the leverage effect, defined as the correlation between the innovation in the volatility process and that in the price process. For the stationary case the parameters are chosen as $k = 0.9$, $\psi = -9.2$, $\gamma = 0.5$, $d = 0.4$ and $\rho = -0.3$, where $\rho = corr(dW(t), dW_d(t))$, while for the nonstationary case,
Table 2: Bias and Root mean squared error of Monte Carlo estimates of $\delta = 1 + d$, when $\sigma^2(t)$ follows the process in (24)

(a) $\delta = 1 + d = 0.6$ AND $\gamma = 0.2$.

<table>
<thead>
<tr>
<th>$\zeta$</th>
<th>$0.000$</th>
<th>$\zeta$</th>
<th>$0.001$</th>
<th>$\zeta$</th>
<th>$0.002$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias $\hat{\delta}$</td>
<td>RMSE($\hat{\delta}$)</td>
<td>Bias $\hat{\delta}_c$</td>
<td>RMSE($\hat{\delta}_c$)</td>
<td>Bias $\hat{\delta}$</td>
</tr>
<tr>
<td>$IV$</td>
<td>1.02</td>
<td>0.0596</td>
<td>1.02</td>
<td>0.0596</td>
<td>1.02</td>
</tr>
<tr>
<td>$RV^{all}$</td>
<td>0.62</td>
<td>0.0596</td>
<td>2.10</td>
<td>0.0501</td>
<td>-18.84</td>
</tr>
<tr>
<td>$RV^1$</td>
<td>-1.32</td>
<td>0.0601</td>
<td>3.29</td>
<td>0.0639</td>
<td>-10.55</td>
</tr>
<tr>
<td>$RV^5$</td>
<td>-8.96</td>
<td>0.0797</td>
<td>2.38</td>
<td>0.0881</td>
<td>-14.38</td>
</tr>
<tr>
<td>$RV^{10}$</td>
<td>-16.11</td>
<td>0.1137</td>
<td>1.68</td>
<td>0.1026</td>
<td>-20.10</td>
</tr>
<tr>
<td>$RV^{30}$</td>
<td>-33.70</td>
<td>0.2105</td>
<td>1.23</td>
<td>0.1289</td>
<td>-35.61</td>
</tr>
</tbody>
</table>

(b) $\delta = 1 + d = 0.7$ AND $\gamma = 0.2$.

<table>
<thead>
<tr>
<th>$\zeta$</th>
<th>$0.000$</th>
<th>$\zeta$</th>
<th>$0.001$</th>
<th>$\zeta$</th>
<th>$0.002$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias $\hat{\delta}$</td>
<td>RMSE($\hat{\delta}$)</td>
<td>Bias $\hat{\delta}_c$</td>
<td>RMSE($\hat{\delta}_c$)</td>
<td>Bias $\hat{\delta}$</td>
</tr>
<tr>
<td>$IV$</td>
<td>-0.26</td>
<td>0.0639</td>
<td>-0.26</td>
<td>0.0639</td>
<td>-0.26</td>
</tr>
<tr>
<td>$RV^{all}$</td>
<td>-0.41</td>
<td>0.0638</td>
<td>0.07</td>
<td>0.0571</td>
<td>-0.07</td>
</tr>
<tr>
<td>$RV^1$</td>
<td>-1.12</td>
<td>0.0638</td>
<td>0.20</td>
<td>0.0465</td>
<td>-4.66</td>
</tr>
<tr>
<td>$RV^5$</td>
<td>-4.40</td>
<td>0.0703</td>
<td>-0.26</td>
<td>0.0452</td>
<td>-6.73</td>
</tr>
<tr>
<td>$RV^{10}$</td>
<td>-7.70</td>
<td>0.0837</td>
<td>-0.99</td>
<td>0.0470</td>
<td>-9.74</td>
</tr>
<tr>
<td>$RV^{30}$</td>
<td>-18.57</td>
<td>0.1442</td>
<td>-2.70</td>
<td>0.0769</td>
<td>-19.80</td>
</tr>
</tbody>
</table>

Notes: $\hat{\delta}$ denotes the LW estimator, while $\hat{\delta}_c$ is the corrected LW estimator, see (21). The term Bias is referred to the relative percentage bias, defined in equation (26)-(27). The estimates are based on 1,000 samples of 2,500 daily observations from model (22)-(23) with parameter values indicated in table and discretization step set to 10 seconds. The bandwidth used in the estimation of $\delta$ is $m = T^{0.65}$. When $d = -0.4$ and $d = -0.3$, then $\delta = 0.6$ and $\delta = 0.7$, respectively.
they are $\gamma = 0.2$, $d = -0.3$ and $\rho = -0.3$. The value $\rho = -0.3$ is chosen according to the findings in Andersen et al. (2002). Table 3 reports the simulation results for the case in which the innovations of the volatility process are correlated with the innovations of the price process (leverage effect). From Table 4(a) it emerges that, for intermediate choices of $\Delta$, the corrected LW estimator is generally robust to the presence of correlation between the signal and the noise. In particular, the correction works better for intermediate choices of $\Delta$. This indirectly confirms the finding in Meddahi (2002, p.493), that the correlation between the $IV_t$ and the measurement error increases non-linearly with $\Delta$, when $\rho \neq 0$. In the stationary case, the correlation is always positive and it ranges between 40% at 1 minute frequency and 60% at 30 minutes frequency. On the contrary, in the nonstationary case, the leverage effect produces a negative correlation that decreases with the sampling frequency, being -0.36% at 1 minute and -0.19% at 30 minutes. In this case, Table 4(a) shows that the leverage effect has a little impact on $\hat{\delta}_c$, which is well centered on the true parameter value also for large values of $\Delta$, while the LW estimator performs as in the no-leverage case.

Summarizing, the simulation results highlight the effectiveness, in terms of bias reduction, of the corrected LW estimator of long memory of the $IV_t$, even when the volatility signal is nonstationary and in presence of leverage effect. More importantly, the corrected estimator is robust to the choice of the sampling frequency used for the computation of $RV_t^{\Delta}$. A recent paper by Arteche (2012) proposes a modification of the Hurvich et al. (2005) estimator to account for the possible correlation between signal and noise, which may emerge under the presence of leverage. A detailed theoretical and empirical investigation of the consequences of the leverage on the estimation of long memory in $IV$ is left for future research.
Table 3: Bias and Root mean squared error of Monte Carlo estimates of $d$ and $\delta$, under leverage effect

(a) $d = 0.4$, $\gamma = 0.5$ AND $\rho = -0.3$.

<table>
<thead>
<tr>
<th></th>
<th>$\zeta = 0.000$</th>
<th>$\zeta = 0.001$</th>
<th>$\zeta = 0.002$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias $\hat{d}$</td>
<td>RMSE($\hat{d}$)</td>
<td>Bias $\hat{d}_c$</td>
</tr>
<tr>
<td>$IV$</td>
<td>-0.91</td>
<td>0.0554</td>
<td>-0.91</td>
</tr>
<tr>
<td>$RV^{all}$</td>
<td>-2.22</td>
<td>0.0568</td>
<td>3.21</td>
</tr>
<tr>
<td>$RV^1$</td>
<td>-7.50</td>
<td>0.0668</td>
<td>3.07</td>
</tr>
<tr>
<td>$RV^5$</td>
<td>-20.07</td>
<td>0.1081</td>
<td>-5.55</td>
</tr>
<tr>
<td>$RV^{10}$</td>
<td>-26.95</td>
<td>0.1380</td>
<td>-11.35</td>
</tr>
<tr>
<td>$RV^{30}$</td>
<td>-34.51</td>
<td>0.1682</td>
<td>-16.86</td>
</tr>
</tbody>
</table>

(b) $\delta = 1 + d = 0.7$, $\gamma = 0.2$ AND $\rho = -0.3$.

<table>
<thead>
<tr>
<th></th>
<th>$\zeta = 0.000$</th>
<th>$\zeta = 0.001$</th>
<th>$\zeta = 0.002$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias $\hat{\delta}$</td>
<td>RMSE($\hat{\delta}$)</td>
<td>Bias $\hat{\delta}_c$</td>
</tr>
<tr>
<td>$IV$</td>
<td>2.16</td>
<td>0.0352</td>
<td>2.16</td>
</tr>
<tr>
<td>$RV^{all}$</td>
<td>-0.57</td>
<td>0.0652</td>
<td>3.99</td>
</tr>
<tr>
<td>$RV^1$</td>
<td>-1.21</td>
<td>0.0655</td>
<td>-0.71</td>
</tr>
<tr>
<td>$RV^5$</td>
<td>-4.24</td>
<td>0.0711</td>
<td>1.15</td>
</tr>
<tr>
<td>$RV^{10}$</td>
<td>-7.60</td>
<td>0.0843</td>
<td>0.44</td>
</tr>
<tr>
<td>$RV^{30}$</td>
<td>-18.18</td>
<td>0.1435</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Notes: $\hat{d}$ and $\hat{\delta}$ denote the LW estimators, while $\hat{d}_c$ and $\hat{\delta}_c$ are the corrected LW estimators, see (21). The term Bias is referred to the relative percentage bias, defined in equation (26)-(27). The estimates are based on 1,000 samples of 2,500 daily observations from models (22), (23) and (24) with parameter values indicated in table and discretization step set to 10 seconds. The bandwidth used in the estimation of $d$ and $\delta$ is $m = T^{0.65}$. In Panel (a), the $IV$ process is stationary, with $d = 0.4$. In Panel (b), the $IV$ process is nonstationary, with $\delta = 0.7$. 
4 Empirical Analysis

We estimate the long memory of $IV_t$, based on the $RV$ series of 28 stocks traded on NYSE. The sample period ranges from January 2, 2001 to December 31, 2007, for a total of 1760 trading days. The $RV$ series are computed with alternative sample frequencies, say 1 minute, 5 minutes, 10 minutes, 15 minutes and 30 minutes.

Table 4 reports the estimates of the long memory parameter $d$. The corrected estimates signal that volatilities are generated by a nonstationary long memory process. Firstly, on average, the estimates obtained correcting for the measurement error are higher than those obtained with the LW estimator, which lies in the stationary region. Secondly, the corrected estimates are relatively constant with respect to the choice of the sampling frequency used for the computation of the $RV$. Instead, the non corrected ones are characterized by a downward trend with respect to $\Delta$. The dispersion of the corrected estimates, as measured by $\sigma(\hat{d})$ and $\hat{d}_1 - \hat{d}_{30}$, is smaller than that observed for the LW estimates. This evidence is clear from a visual inspection of Figure 3, which reports the long memory signature plot for four stocks in the sample. The LW estimates based on the $RV$ series (dashed line) fall in most cases in the stationary region, while it is evident the downward trend with respect to $\Delta$. On the other hand, the corrected estimates of $d$ (solid line) are always above the Whittle estimates and are constant across different choices of $\Delta$, in line with the simulation results.

We think that it is important to stress the fact that the corrected LW estimator always lies in the nonstationary region, suggesting that the volatility process could be a nonstationary process. From this point of view, the fact that the LW estimates of $d$, based on $RV_t^{\Delta}$, turn out to be less than 0.5, namely a stationary long memory process, is mainly due to the role of the measurement error.

This means that using a biased long memory estimator leads to wrong conclusions on the stationarity of the integrated and instantaneous volatility processes.

Using a similar argument, but in a discrete-time domain framework, Hansen and

\textsuperscript{11}We avoid the possible upward bias in the semiparametric estimates of $d$, due to the presence of large shifts as generated by changing bull and bear markets, during the 2008-2009 financial crisis.
Table 4: Long memory estimates based on $RV_t^\Delta$ of 28 NYSE stocks

<table>
<thead>
<tr>
<th>Company</th>
<th>Local Whittle</th>
<th>Corrected Local Whittle</th>
<th>$d = d_0$</th>
<th>$\bar{d}$</th>
<th>$\sigma(\hat{d})$</th>
<th>$\hat{d}<em>1 - \hat{d}</em>{30}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.5722</td>
<td>0.5196</td>
<td>0.5039</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>BA</td>
<td>0.5378</td>
<td>0.5165</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>CAT</td>
<td>0.5465</td>
<td>0.5172</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>C</td>
<td>0.5637</td>
<td>0.5177</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>DD</td>
<td>0.5637</td>
<td>0.5177</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>EMR</td>
<td>0.5245</td>
<td>0.4933</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>F</td>
<td>0.4820</td>
<td>0.4515</td>
<td>0.4245</td>
<td>0.4323</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>GE</td>
<td>0.5291</td>
<td>0.4799</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>GS</td>
<td>0.5157</td>
<td>0.4557</td>
<td>0.4245</td>
<td>0.4323</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>HD</td>
<td>0.5619</td>
<td>0.5187</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>HON</td>
<td>0.4636</td>
<td>0.4598</td>
<td>0.4245</td>
<td>0.4323</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>IBM</td>
<td>0.5278</td>
<td>0.4971</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>INTC</td>
<td>0.5615</td>
<td>0.5157</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>MCD</td>
<td>0.4934</td>
<td>0.4581</td>
<td>0.4245</td>
<td>0.4323</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>NEM</td>
<td>0.5094</td>
<td>0.4690</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>PFE</td>
<td>0.5704</td>
<td>0.4953</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>RCG</td>
<td>0.5907</td>
<td>0.5221</td>
<td>0.4865</td>
<td>0.4945</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>TWX</td>
<td>0.4825</td>
<td>0.4515</td>
<td>0.4245</td>
<td>0.4323</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>UPS</td>
<td>0.5232</td>
<td>0.4696</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>XOM</td>
<td>0.5292</td>
<td>0.4998</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
<tr>
<td>CAN</td>
<td>0.5101</td>
<td>0.4760</td>
<td>0.4567</td>
<td>0.4645</td>
<td>0.0065</td>
<td>-0.0122</td>
</tr>
</tbody>
</table>

Notes: The names of the companies corresponding to the stock tickers reported in Table can be found at [http://www.nyse.com](http://www.nyse.com).

The $RV_t^\Delta$ is computed with different choices for $\Delta$, that is 1, 5, 10, 15 and 30 minutes. Table reports the estimates of $d$ obtained using the LW estimator and the corrected LW, see (20). The bandwidth used for the estimation is $m = T^{0.8}$. $\bar{d}$ is the average of the estimates, with respect to the different choices of $\Delta$. $\sigma(\hat{d})$ is the sample standard deviation of the estimates of $d$ computed with different $RV_t^\Delta$. $\hat{d}_1 - \hat{d}_{30}$ is the difference between the estimates obtained with the RV at 1 minute and 30 minutes. AVG is the average for each column.
Figure 3: Long memory signature plots: Long memory parameter estimates for different sampling frequencies (1, 5, 10, 15 and 30 minutes). Dashed lines represent the LW estimator of the memory parameter (obtained minimizing the function in (20) concentrated with respect to $G$ with $\beta = 0$). Solid lines represent the corrected LW estimator (see (21)).
Lunde (2010) have proposed an instrumental variable estimator of the persistence of the signal when the latter is a unit root process. To the best of our knowledge, the consequences of a fractional, but nonstationary, volatility process are not studied yet in the literature and the evidence reported here deserves a more detailed analysis.

5 Conclusions

A stylized fact is that $RV$ has long memory. In this paper, we investigate the dynamic properties and the source of the long-range dependence of $RV$. First, we find that, when the instantaneous volatility is driven by a fractional Brownian motion, the $IV_t$ is characterized by the same degree of long-range dependence, $d$. As a consequence, the $RV$ inherits this property, since the spectral density of $RV$ is equal to the spectral density of $IV$, plus a term which depends on the variance of the measurement error.

The additional term in the spectral density of $RV$ impacts on the finite sample properties of the semiparametric estimates of $d$, since they crucially depend on the use of $RV$ in place of the unobservable $IV$. In absence of microstructure noise, the $RV$ spectral density converges to the spectral density of $IV$, as $\Delta \to 0$. When the presence of microstructure noise prevents us from using all the available price observations, the additional component in the spectral density, which depends on the discretization error and on the microstructure noise, significantly affects the finite sample bias of semiparametric estimates of $d$.

We adopt a correction of the LW estimator along the lines of Hurvich et al. (2005). A Monte Carlo experiment confirms that the correction of the local Whittle estimator is robust to the measurement error for all choices of $\Delta$. Thus the trade-off between discretization error and microstructure noise is neutralized by adopting a corrected version of the LW estimator. Finally, the estimation of the long memory of 28 NYSE stocks emphasizes the practical importance of considering the measurement error when estimating the degree of long memory of $IV$.
The corrected estimates of $d$ suggest that the $IV$ and the instantaneous volatility can be nonstationary processes. In this study we have not considered the role of jumps in prices and their potential effect on the estimation of long memory in $IV$. This is left for future research.
A Proofs

A.1 Proof of Proposition 1

Given that $IV_t = \int_{t-1}^{t} \hat{\sigma}^2(s)ds$. Following Chambers (1996) we express the integral operator in the definition of $IV$ as a simple filter that has transfer function

$$T(\lambda) = \int_{0}^{1} e^{-i\lambda u} du = \frac{1}{(-i\lambda)}[e^{-i\lambda} - 1].$$

Therefore the spectral density (or pseudo spectral density) of $IV$ is given by

$$f_{IV}(\lambda) = |T(\lambda)|^2 f_{\sigma^2}(\lambda). \quad (28)$$

The limit of $f_{IV}(\lambda)$ for $\lambda \to 0$ is

$$\lim_{\lambda \to 0} f_{IV}(\lambda) = \lim_{\lambda \to 0} |T(\lambda)|^2 f_{\sigma^2}(\lambda) \quad (29)$$

Since $|T(\lambda)|^2 = \frac{2(1-\cos(\lambda))}{|\lambda|^2}$ and $(1-\cos(\lambda)) \approx |\lambda|^2/2$ as $\lambda \to 0$, then $\lim_{\lambda \to 0} |T(\lambda)|^2 = 1$. Thus,

$$\lim_{\lambda \to 0} \lambda^{2d} f_{IV}(\lambda) = \lim_{\lambda \to 0} \lambda^{2d} f_{\sigma^2}(\lambda) = c, \quad (30)$$

that is $IV$ has the same degree of long memory of $\sigma^2(t)$, which is equivalent to $\sigma^2(t)$.

A.2 Proof of Proposition 2

Assume that the processes for $p(t)$ and $RV_{t}\Delta$ are those in (1) and (6). Assume also that $\sigma^2(t)$ is such that condition (2) is verified, and that $m(t) = \mu$ and no leverage effect, then

$$u_{t,i}^\Delta \equiv \frac{\xi}{\sigma_{t,i}\Delta} (z_{t,i}^2 - 1) + \Delta^2 \mu^2 + 2\Delta \mu \sigma_{t,i} \Delta z_{t,i}. \quad (31)$$

It is easy to show that

(i) $E(u_{t,i}^\Delta) = \Delta \mu^2$;
(ii) \( \text{Var}(u_t^\Delta) = 2\Delta^{-1}E\left[(\sigma_{t,i}^2)\right] + 4\Delta\mu^2E\left[\sigma_{t,i}^2\right]; \)

(iii) \( u_t^\Delta \) is dynamically uncorrelated, i.e., \( \text{Cov}(u_t^\Delta, u_{t+h}^\Delta) = 0, \) for any integer \( h \neq 0; \)

(iv) The error term \( u_t^\Delta \) is uncorrelated with \( IV_t; \)

(v) \( \text{Cov}(RV_t^\Delta, RV_{t-h}^\Delta) = \text{Cov}(IV_t, IV_{t-h}), \) for any integer \( h \neq 0. \)

Thus, for \( \Delta > 0 \) and \( 0 < d < 1/2 \) the spectral density of \( RV_t^\Delta \) is given by

\[
f_{RV^\Delta}(\lambda) = \frac{1}{2\pi} \left\{ \text{Var}(IV_t) + \text{Var}(u_t^\Delta) + 2\sum_{j=1}^{\infty} [\text{Cov}(IV_t, IV_{t-j}) \cos(\lambda j)] \right\}
= f_{IV}(\lambda) + f_{u^\Delta}(\lambda)
\]

(32)

When \( \Delta > 0 \) and \( 1/2 \leq d < 1, \) the pseudo spectral density of \( RV_t^\Delta \) is given by the expectation of its sample periodogram, i.e.

\[
f_{RV^\Delta}(\lambda) \equiv E(I_{RV^\Delta}(\lambda)) = E(I_{IV^\Delta+u^\Delta}(\lambda))
= E(I_{IV}(\lambda)) + E(I_{u^\Delta}(\lambda))
= f_{IV}(\lambda) + f_{u^\Delta}(\lambda)
\]

(33)

where \( I_\cdot(\lambda) \) is the sample periodogram.

Therefore, for \( 0 < d < 1, \) \( \lim_{\lambda \to 0} \lambda^{2d}f_{RV^\Delta}(\lambda) = \lim_{\lambda \to 0} \lambda^{2d}(f_{IV}(\lambda)+f_{u^\Delta}(\lambda)) = c \) with \( c > 0, \) where \( f_{u^\Delta}(\lambda) = \frac{\text{Var}(u_t^\Delta)}{2\pi}. \)

Given that \( \text{Var}(u_t^\Delta) \) converges to zero as \( \Delta \to 0, \) so that \( f_{u^\Delta}(\lambda) \to 0. \) This implies that

\[
\lim_{\Delta \to 0} f_{RV^\Delta}(\lambda) = f_{IV}(\lambda)
\]

(34)

the proof then follows from Proposition 1, and multiplying both sides by \( \lambda^{2d}, \)

letting \( \lambda \to 0. \)

### A.3 Proof of Proposition 3

Consider the processes \( \tilde{p}(t), RV_t^\Delta, \xi_t^\Delta \) defined respectively in (14), (6) and (16). Let \( m(t) = \mu \) and assume no leverage effect. Assume also that \( \sigma^2(t) \) is such that
Estimation of Long Memory in Integrated Variance

condition (2) is verified. First, in order to characterize the spectral density of $RV_t^\Delta$, we need to obtain the moments of the measurement error, $\xi_t^\Delta = \sum_{i=1}^{n} \xi_{t,i}^\Delta$.

(a) $\xi_{t,i}^\Delta$ is defined as

$$\xi_{t,i}^\Delta \equiv \sigma_{t,i,\Delta}^2 (z_{t,i}^2 - 1) + \Delta^2 \mu^2 + \eta_{t,i,\Delta}^2 + 2(\sigma_{t,i,\Delta} z_{t,i,\Delta} \eta_{t,i,\Delta}) + 2\Delta \mu \eta_{t,i,\Delta} + 2\Delta \mu \sigma_{t,i,\Delta} z_{t,i},$$

hence

$$E(\xi_{t,i}^\Delta) = E\left[\sigma_{t,i,\Delta}^2 (z_{t,i}^2 - 1) + \Delta^2 \mu^2 + \eta_{t,i,\Delta}^2\right] + 2E(\sigma_{t,i,\Delta} z_{t,i,\Delta} \eta_{t,i,\Delta}) + 2\Delta \mu E(\eta_{t,i,\Delta}) + 2\Delta \mu E(\sigma_{t,i,\Delta} z_{t,i})$$

$$= \Delta^2 \mu^2 + E(\eta_{t,i,\Delta}^2)$$

$$= \Delta^2 \mu^2 + \sigma_{\eta}^2,$$

where $\sigma_{\eta}^2 = \text{Var}[\eta_{t,i,\Delta}] = 2 \text{Var}[\epsilon_{t,i,\Delta}]$. Because $\sigma_{t,i,\Delta}$, $z_{t,i}$, and $\eta_{t,i,\Delta}$ are mutually independent, $E(\sigma_{t,i,\Delta} z_{t,i} \eta_{t,i,\Delta}) = E(\sigma_{t,i,\Delta}) \cdot E(z_{t,i}) \cdot E(\eta_{t,i,\Delta}) = 0$.

It follows that $E\left(\sum_{i=1}^{n} \xi_{t,i}^\Delta\right) = \Delta^{-1} \sigma_{\eta}^2 + \Delta \mu^2$.

(b) The covariance between $\xi_{t,i}^\Delta$ and $\xi_{t,j}^\Delta$ can be written as

$$\text{Cov}\left(\xi_{t,i}^\Delta, \xi_{t,j}^\Delta\right) = E\left[u_{t,i}^\Delta u_{t,j}^\Delta\right] + E\left[u_{t,i}^\Delta \eta_{t,j,\Delta}^2\right] + 2E\left[u_{t,i}^\Delta (\sigma_{t,j,\Delta} z_{t,j,\Delta} \eta_{t,j,\Delta})\right]$$

$$+ 2\Delta \mu E\left[u_{t,i}^\Delta \eta_{t,j,\Delta}\right] + 4\Delta \mu \left[\sigma_{t,i,\Delta} z_{t,i,\Delta} \eta_{t,i,\Delta}\right] \eta_{t,j}\right)$$

$$+ E\left[\eta_{t,i,\Delta}^2 u_{t,j}^\Delta\right] + E\left[\eta_{t,i,\Delta}^2 \eta_{t,j,\Delta}^2\right] + 2E\left[\eta_{t,i,\Delta}^2 (\sigma_{t,j,\Delta} z_{t,j,\Delta} \eta_{t,j,\Delta})\right]$$

$$+ 2E\left[\sigma_{t,i,\Delta} z_{t,i,\Delta} \eta_{t,i,\Delta}\right] + 2E\left[\sigma_{t,i,\Delta} z_{t,i,\Delta} \eta_{t,i,\Delta}\right] \eta_{t,j,\Delta}\right]$$

$$+ 4E\left[\sigma_{t,i,\Delta} z_{t,i,\Delta} \eta_{t,i,\Delta}\right] \eta_{t,j,\Delta}\right]$$

$$+ 2\Delta \mu \left[\eta_{t,i,\Delta}^2 \eta_{t,j,\Delta}^2\right] + 2\Delta \mu \left[\eta_{t,i,\Delta}^2 \eta_{t,j,\Delta}^2\right] + 2\Delta \mu \left[\eta_{t,i,\Delta} (\sigma_{t,j,\Delta} z_{t,j,\Delta} \eta_{t,j,\Delta})\right]$$

$$+ 2\Delta \mu E\left[\eta_{t,i,\Delta}^2 \eta_{t,j,\Delta}^2\right] + 4\Delta^2 \mu^2 E[\eta_{t,i,\Delta} \eta_{t,j,\Delta}] - \sigma_{\eta}^4 - 2\Delta^2 \mu^2 \sigma_{\eta}^2 - \Delta^4 \mu^4$$

$$= \sigma_{\eta}^4 + 2\Delta^2 \mu^2 \sigma_{\eta}^2 + \Delta^4 \mu^4 - \sigma_{\eta}^4 - 2\Delta^2 \mu^2 \sigma_{\eta}^2 - \Delta^4 \mu^4 = 0 \quad \forall i \neq j$$
The covariance of $\xi^\Delta_t$ and $\xi^\Delta_{t+h}$ is equal to

$$\text{Cov} \left( \sum_{i=1}^{n} \xi^\Delta_{t,i}, \sum_{j=1}^{n} \xi^\Delta_{t+h,j} \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov} (\xi^\Delta_{t,i}, \xi^\Delta_{t+h,j}) = 2n^2 \cdot 0 = 0 \quad \text{for any integer } h \neq 0.$$ \hfill (36)

(c) The variance of $\xi^\Delta_{t,i}$ is,

$$\text{Var} (\xi^\Delta_{t,i}) = \text{Var}(u^\Delta_{t,i}) + \text{Var}(\eta^2_{t,i}) + 4 \text{Var} (\sigma^2_{t,i,\Delta} z_{t,i} \eta_{t,i,\Delta}) + 4 \Delta^2 \mu^2 \text{Var}(\eta_{t,i})$$

$$= 2E \left[ (\sigma^2_{t,i,\Delta})^2 \right] + 4 \Delta^2 \mu^2 E \left[ \sigma^2_{t,i,\Delta} \right] + E(\eta^4_{t,i,\Delta}) - \sigma^4_{\eta} + 4 \sigma^2_{\eta} E \left[ \sigma^2_{t,i,\Delta} \right] + 4 \Delta^2 \mu^2 \sigma^2_{\eta},$$

it follows that the variance of $\xi^\Delta_t$ is

$$\text{Var} (\xi^\Delta_t) = \text{Var} \left( \sum_{i=1}^{n} \xi^\Delta_{t,i} \right) = \sum_{i=1}^{n} \text{Var} (\xi^\Delta_{t,i})$$

$$= 2\Delta^{-1} E \left[ (\sigma^2_{t,i,\Delta})^2 \right] + 4 \Delta^2 \mu^2 E \left[ (\sigma^2_{t,i,\Delta})^2 \right] + \Delta^{-1}(E(\eta^4_{t,i,\Delta}) - \sigma^4_{\eta})$$

$$+ 4 \Delta^{-1} \sigma^2_{\eta} E \left[ \sigma^2_{t,i,\Delta} \right] + 4 \Delta^2 \mu^2 \sigma^2_{\eta} \quad \text{(37)}$$

i. For $\Delta > 0$ and $0 < d < 1/2$, the spectral density of $RV^\Delta_t$ is therefore given by:

$$f_{RV^\Delta}(\lambda) = \frac{1}{2\pi} \left\{ \text{Var}(IV_t) + \text{Var}(\xi^\Delta_t) + 2 \sum_{j=1}^{\infty} [\text{Cov}(IV_t, IV_{t-j}) \cos(\lambda j)] \right\}$$

$$= f_{IV}(\lambda) + f_{\xi^\Delta}(\lambda).$$

When $\Delta > 0$ and $1/2 \leq d < 1$, the pseudo spectral density of $RV^\Delta_t$ is given by the expectation of the sample periodogram of $RV$, i.e.

$$f_{RV^\Delta}(\lambda) \equiv E(I_{RV^\Delta}(\lambda)) = E(I_{IV^\Delta}(\lambda))$$

$$= E(I_{IV}(\lambda)) + E(I_{\xi^\Delta}(\lambda))$$

$$= f_{IV}(\lambda) + f_{\xi^\Delta}(\lambda) \quad \text{(38)}$$

The proof then follows from Proposition 1, and multiplying both sides by $\lambda^{2d}$, letting $\lambda \to 0$.

ii. It is evident that when $\Delta \to 0$, $\text{Var} (\xi^\Delta_t) \to \infty$, so that $f_{RV^\Delta}(\lambda) \to \infty \forall \lambda.$


Acknowledgements

We are grateful to Kris Boudt, Niels Haldrup, Roberto Renò, an Associate Editor and two anonymous referees and two anonymous referees for useful comments and suggestions which greatly helped to improve the paper. We thank Asger Lunde for having provided the data. We also thank the participants at the ICEEE’11 conference, the Econometric Society European Meeting 2011 in Oslo, and the XII Workshop on Quantitative Finance for useful comments. We acknowledge the research support of CREATES (funded by the Danish National Research Foundation).
References


Haldrup, N. and Nielsen, M. O. (2007). Estimation of fractional integration
in the presence of data noise. *Computational Statistics & Data Analysis*,

Hansen, P. R. and Lunde, A. (2005). A realized variance for the whole day
based on intermittent high-frequency data. *Journal of Financial Econometrics*,


Hansen, P. R. and Lunde, A. (2010). Estimating the persistence and the auto-
correlation function of a time series that is measured with error. Research
Paper 2010-8, CREATES.

and Satchell, S., editors, *Forecasting Volatility in Financial Markets*. London:

stationary or noninvertible fractionally integrated processes. *Journal of Time
Series Analysis*, 16:17–42.


Hurvich, C. M. and Ray, B. K. (2003). The local whittle estimator of long-
470.
Estimation of Long Memory in Integrated Variance


