Proceedings of the World Congress on Engineering 2008 Vol II
WCE 2008, July 2 - 4, 2008, London, U.K.

The Taylor Effect Of Asset Returns: Stylized Fact Or Finite-Sample Behaviour?

Violetta Dalla *

Abstract—**Different power transformations of absolute returns of various financial assets have been found to display different magnitudes of sample autocorrelations, a property referred to as the Taylor effect. In this paper, we consider the long memory stochastic model for the returns under which, the asymptotic rate of decay of the autocorrelations of powers of absolute returns is governed by their long memory parameter. We show that, although the true long memory parameter of powers of absolute returns is the same across the different powers, the local Whittle estimator of the long memory parameter has finite-sample bias that differs across the power transformations chosen. A Monte-Carlo experiment provides evidence in support of our result that the reported differences in the long-run properties of the various power transformations of absolute returns could be due to finite-sample behaviour.**

Keywords: Taylor effect, stochastic volatility, local Whittle estimation

1 Introduction

In the empirical literature, asset returns are commonly found to be approximately uncorrelated over time while their non-linear transformations, such as powers of absolute returns and their logarithms, show significant autocorrelation over many lags. The degree of the latter autocorrelations has been found to vary across the different non-linear transformations chosen. This was first noted by [19], who found that for various financial series the sample autocorrelations are higher for the absolute returns than for the squared ones. In later studies, [7], [11], [6] and [12] examined a range of financial series and observed that the sample autocorrelations of the $p-$th power of absolute returns for various values of p tend to be highest when $p = 1$. This observation was termed in [11] as the Taylor effect and has been considered as one of the empirical stylized facts on asset returns.

Various models have been proposed in the literature to account for the strong sample autocorrelations of the powers of absolute returns. Among them is the long memory stochastic volatility (LMSV) model proposed independently by [2] and [14], and whose form has allowed for certain statistical results to be established across different powers p. In particular, the autocorrelation function of powers of absolute returns under the LMSV model was derived by [14]. The latter author found that the power that gives rise to the highest autocorrelations depends on the parameters in the underlying LMSV model and therefore, there will be certain LMSV models for which the autocorrelations of the powers of absolute returns are highest when $p = 1$. Hence, one cannot rule out that the empirical results found for the sample autocorrelations of powers of returns are simply driven by their population counterparts. In such a case, the Taylor effect would be due to the data generating mechanism and then the practitioner would want to choose a model that accounts for this effect.

However, one would wonder whether any of the observed differences in the sample autocorrelations of powers of absolute returns are driven by differences in the finite-sample behaviour of the sample autocorrelation due to the power transformation chosen. Indications of this can be found in studies related to the asymptotic rate of decay of the autocorrelations of powers of absolute returns. Under the LMSV model, [14] established that the asymptotic rate of decay of the autocorrelations of powers of absolute returns is the same for all powers p. This common rate of decay is governed by the long memory parameter in the LMSV model, so that powers of absolute returns have the same long memory parameter irrespective of the power p chosen. However, Monte-Carlo experiments with the LMSV model, performed by [20], [15], [5] and [3], suggest that semiparametric estimators of the long memory parameter of the $p-$th power of absolute returns display higher degree of negative finite-sample bias when $p = 2$ than when $p = 1$. As these semiparametric estimators are periodogram-based, the question arises as to whether when looking at powers of absolute returns the finite-sample properties of estimators based on second-order dependence are affected by the choice of the power p. In such a case, it could be possible that the Taylor effect is also driven by the finite-sample behaviour of the estimators used to identify the dependence in the $p-$th powers of absolute returns.

The main purpose of this paper is to investigate the effect

*I would like to thank L. Giraitis for helpful discussions and suggestions. Address: Department of Economics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. Tel: +44 1784 414 972, Email: violetta.dalla@rhul.ac.uk.

of the power p of absolute returns on the finite-sample behaviour of estimation based on second-order dependence. We consider the LMSV model for the returns and examine estimation of the long memory parameter for powers of absolute returns. We choose the local Whittle (LW) estimator of the long memory parameter introduced by [16]; for the $p-$th power of absolute returns under the LMSV model its consistency was established by [3] and its finite-sample properties were examined in some of the aforementioned Monte-Carlo experiments. We show that the finite-sample properties of the LW estimator applied to the $p-$th power of absolute returns under the LMSV model differ across the power p chosen. In particular, we prove that the dominant term in its finite-sample bias depends on the power p and that for certain cases the latter dependence is quadratic in p. We also conduct a Monte-Carlo experiment as in [15] and [3] extending the range of the powers p chosen and find that the finite-sample bias of the LW estimator of the $p-$th power of absolute returns is severely affected by the choice of the power p. The rest of the paper is as follows. Section 2 discusses the LMSV model for the returns and the LW estimator of the long memory parameter. Section 3 contains our theoretical results on the finite-sample properties of the LW estimator applied to powers of absolute returns under the LMSV model of Section 2. A Monte-Carlo study is contained in Section 4, while Section 5 concludes. The proofs of Section 3 are found in Appendix A and all figures are given in Appendix B.

2 LMSV model and LW estimation

We consider a version of the LMSV model of [14] and [2]. We assume that the returns of an asset, denoted by $\{r_t\}$, satisfy

$$\begin{aligned} r_t &= u_t\sigma_t \\ &= \sigma u_t \exp\left(\sigma_h h_t\right), \end{aligned} \tag{1}$$

where σ and σ_h are positive constants, and that the following assumptions hold:

A.1 $\{u_t\}$ is a standard Gaussian i.i.d. sequence.

A.2 $\{h_t\}$ is a standard Gaussian sequence.

A.3 $\{h_t\}$ and $\{u_t\}$ are mutually independent.

A.4 $\{h_t\}$ is a long memory sequence with long memory parameter $0 < \alpha_h < 1$, whose spectral density function $f_h(.)$ satisfies

$$f_h(\lambda) = \lambda^{-\alpha_h}\left(c_{0,h} + c_{1,h}\lambda^2 + o(\lambda^2)\right), \quad \lambda \to 0+,$$

and autocorrelation function $\rho_h(.)$ has the property

$$\rho_h(\tau) \sim c_h \tau^{-1+\alpha_h}, \quad \tau \longrightarrow +\infty.$$

Assumptions A.1-A.3 are as those in [14] and [2]. Under these assumptions, the latter authors derived various properties for the returns and their non-linear transformations. They also considered that $\{h_t\}$ is a stationary $ARFIMA(p, \alpha_h/2, q)$ model, see [1] and [13]. Our assumption A.4 is satisfied by stationary $ARFIMA(p, \alpha_h/2, q)$ models and is one of the conditions employed in [3] to examine the consistency of the LW estimator applied to powers of absolute returns under the LMSV model (1).

Following [14], we have under assumptions A.1 and A.3 that the autocorrelation function $\rho_p(.)$ of $\{|r_t|^p\}$ satisfies

$$\rho_p(\tau) = \frac{\exp\left(p^2\sigma_h^2\rho_h(\tau)\right) - 1}{\kappa_p \exp\left(p^2\sigma_h^2\right) - 1}, \quad \tau \geq 1, \tag{2}$$

where

$$\kappa_p = \frac{\sqrt{\pi}\Gamma\left(p+\frac{1}{2}\right)}{\Gamma^2\left(\frac{p+1}{2}\right)},$$

with $\Gamma()$ denoting the gamma function, and hence

$$\rho_p(\tau) \sim \frac{p^2\sigma_h^2}{\kappa_p \exp\left(p^2\sigma_h^2\right) - 1}\rho_h(\tau), \quad \tau \longrightarrow \infty,$$

so that for big lags τ the autocorrelation function of $\{|r_t|^p\}$ is proportional to that of $\{h_t\}$. If furthermore we use assumption A.4, we have that

$$\rho_p(\tau) \sim \frac{p^2\sigma_h^2 c_h}{\kappa_p \exp\left(p^2\sigma_h^2\right) - 1}|\tau|^{-1+\alpha_h}, \quad \tau \longrightarrow \infty. \tag{3}$$

As noted in [14], it is clear from (2) that it is not possible to make statements about which power p maximizes the autocorrelation function of $\{|r_t|^p\}$. This does not exclude the possibility that the autocorrelation function of $\{|r_t|^p\}$ is maximized when $p = 1$. The author also adds that higher values of σ_h^2 are associated with lower powers p maximizing the autocorrelation function $\rho_p(.)$. Expression (3) implies that, for big lags, the autocorrelation function of $\{|r_t|^p\}$ decays at the same rate for all powers p, and this rate of decay is controlled by the long memory parameter α_h.

For the estimation of the long memory parameter, we use the LW estimator, see [16] and [17]. Given a generic set of data $\{x_1, ..., x_n\}$, the LW estimator $\widehat{\alpha}_x$ is defined as

$$\widehat{\alpha}_x = \underset{\alpha\in[-1,1]}{\arg\min} U_n(\alpha),$$

of the objective function

$$U_n(\alpha) = \log\left(\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{\alpha} I_x(\lambda_j)\right) - \frac{\alpha}{m}\sum_{j=1}^{m}\log\left(\lambda_j\right),$$

where $\lambda_j = \frac{2\pi j}{n}$, $j = 1, .., n$ denote the Fourier frequencies,

$$I_x(\lambda) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} x_t e^{it\lambda}\right|^2$$

is the periodogram of the data and $m = m_n$ is the bandwidth parameter for which it is assumed that

$$m \to \infty \quad \text{and} \quad m = o(n), \quad \text{as } n \to \infty.$$

Here, we consider the LW estimator applied to $\{|r_t|^p\}$, which we denote by $\widehat{\alpha}_p$. In the LMSV model with assumptions A.1-A.4, [3] showed that the long memory parameter of $\{|r_t|^p\}$ is always given by α_h, and moreover established the consistency of the LW estimator $\widehat{\alpha}_p$. Under further assumptions, they also established the asymptotic distribution of $\widehat{\alpha}_p$, which was found do be independent of the power p. However, the Monte-Carlo experiments in [15] and [3] for the cases $p = 1$ and $p = 2$ suggest that the finite-sample behaviour, in particular the finite-sample bias, of the LW estimator $\widehat{\alpha}_p$ is heavily affected by the choice of the power p. These results point towards the possibility that the choice of the power p affects the finite-sample behaviour of periodogram-based estimators applied to $\{|r_t|^p\}$.

3 Power transformations of the LMSV model and LW estimation

In this section we provide our theoretical results on the finite-sample behaviour of the LW estimator $\widehat{\alpha}_p$ applied to $\{|r_t|^p\}$. Recall from the discussion above that for all powers $p > 0$, the true long memory parameter of $\{|r_t|^p\}$ is equal to α_h. We use the notation $B_\beta = (2\pi)^\beta \frac{\beta}{(\beta+1)^2}$ and $Q_{m,p} = \frac{1}{m}\sum_{j=1}^{m}\left(\log\left(\frac{j}{m}\right)+1\right)\left(c_1^2(p)c_{0,h}\right)^{-1}\lambda_j^{\alpha_h} I_{|r|^p}(\lambda_j)$, where $c_1(p)$ is given in (12). The proof of the following theorem is found in Appendix 5.

Theorem 1 *Suppose that the returns $\{r_t\}$ follows the LMSV model (1) and that assumptions A.1-A.4 are satisfied.*

a) If $\alpha_h > \frac{1}{2}$, then

$$\begin{aligned}\widehat{\alpha}_p - \alpha_h &= -\left(\frac{m}{n}\right)^{1-\alpha_h}\frac{\sigma_h^2 c_{0,h} C_2 B_{1-\alpha_h}}{2}p^2 \\ &\quad -(Q_{m,p} - E(Q_{m,p}))(1+o_p(1)) \\ &\quad +o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^{1-\alpha_h}\right), \qquad (4)\end{aligned}$$

with $C_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |1-l|^{-\alpha_h}|l|^{-\alpha_h}\,dl.$

b) If $\alpha_h < \frac{1}{2}$, then

$$\begin{aligned}\widehat{\alpha}_p - \alpha_h &= -\left(\frac{m}{n}\right)^{\alpha_h}\frac{B_{a_h}}{2\pi c_{0,h}}C(p) \\ &\quad -(Q_{m,p} - E(Q_{m,p}))(1+o_p(1)) \\ &\quad +o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^{\alpha_h}\right), \qquad (5)\end{aligned}$$

where

$$\begin{aligned}C(p) &= C_1(p) + C_2(p) \\ &= \sum_{k=2}^{\infty}\frac{p^{2(k-1)}\sigma_h^{2(k-1)}}{k!}S_k \\ &\quad + \left(\frac{\sqrt{\pi}\,\Gamma\left(p+\frac{1}{2}\right)}{\Gamma^2\left(\frac{p+1}{2}\right)} - 1\right)\frac{\exp\left(p^2\sigma_h^2\right)}{p^2\sigma_h^2}, \qquad (6)\end{aligned}$$

with $S_k = \sum_{l\in\mathbb{Z}}\rho_h^k(l)$, *for* $k = 2, 3, ...$

For any power $p > 0$, equations (4) and (5) provide the deviation of the estimator $\widehat{\alpha}_p$ from its true value α_h. For relatively large fixed n, the last term in (4) and (5) will become negligible. Notice also that in both equations (4) and (5), the first term is non-stochastic and the second term has zero mean. Therefore, for any $p > 0$, the dominant term in the finite-sample bias of the estimator $\widehat{\alpha}_p$ is given by

$$-\left(\frac{m}{n}\right)^{1-\alpha_h}\frac{\sigma_h^2 c_{0,h} C_2 B_{1-\alpha_h}}{2}p^2, \quad \text{if } \alpha_h > \frac{1}{2}, \qquad (7)$$

or by

$$-\left(\frac{m}{n}\right)^{\alpha_h}\frac{B_{a_h}}{2\pi c_{0,h}}C(p), \quad \text{if } \alpha_h < \frac{1}{2}. \qquad (8)$$

It is clear from (7) and (8) that the $p-$th power transformation that we apply to absolute returns affects the finite-sample bias of the LW estimator. If $\alpha_h > \frac{1}{2}$, the effect of p is quadratic. If $\alpha_h < \frac{1}{2}$, the effect of p comes through the quantity $C(p)$. Notice that $C(p)$ is a function of the power p, the variance σ_h^2 and also of the autocorrelation function of $\{h_t\}$ through the quantities S_k in $C_1(p)$. Moreover, the autocorrelation function of $\{h_t\}$ will have to be non-negative in order for the autocorrelation function of $\{|r_t|^p\}$ to be non-negative. Since $\alpha_h < \frac{1}{2}$, we have that for all $k = 2, 3, ...$ S_k is a positive finite quantity that does depend on p. So, $C_1(p)$ is a strictly increasing function in $p > 0$. On the other hand, $C_2(p)$ depends on p and σ_h^2, and one can easily show that it is strictly increasing in $p > \frac{1}{\sigma_h^2}$. For $p \le \frac{1}{\sigma_h^2}$, $C_2(p)$ can be increasing and/or decreasing in p depending on the value of σ_h^2, see Figures 1 and 2. Therefore, the function $C(p)$ is strictly increasing in p for all $p > \frac{1}{\sigma_h^2}$. For $p \le \frac{1}{\sigma_h^2}$, the shape of $C(p)$ depends on the value of σ_h^2 and whether $C_1(p)$ dominates $C_2(p)$ or not. However, even if σ_h^2 were known, there are practical obstacles in the calculation of the derivative of $C_1(p)$, since $C_1(p)$ depends also on the autocorrelation function of $\{h_t\}$ for which no parametric model has been specified[1].

[1] Even if we chose the simple $ARFIMA(0, \frac{\alpha_h}{2}, 0)$ model for $\{h_t\}$ and knew the true value of σ_h, we would need to compare for fixed α_h the derivative of $C_1(p) = \sum_{k=2}^{\infty}\frac{p^{2(k-1)}\sigma_h^{2(k-1)}}{k!}\left(\frac{\Gamma(1-\frac{\alpha_h}{2})}{\Gamma(\frac{\alpha_h}{2})}\right)^k \sum_{l\in\mathbb{Z}}\left(\frac{\Gamma(l+\frac{\alpha_h}{2})}{\Gamma(l+1-\frac{\alpha_h}{2})}\right)^k$ with that of $C_2(p)$.

ISBN:978-988-17012-3-7

Corollary 1 *In the LMSV model (1) with assumptions A.1-A.4, we have that the finite-sample bias of the LW estimator applied to* $|r_t|^p$ *depends on the power transformation* $p > 0$ *chosen. In particular, when* $\alpha_h > \frac{1}{2}$ *or when* $\alpha_h < \frac{1}{2}$ *and* $p > \frac{1}{\sigma_h^2}$*, the dominant term in the finite-sample bias of the LW estimator* $\widehat{\alpha}_p$ *increases (in absolute terms) as p increases.*

The results of the corollary help explain the results of the Monte-Carlo experiments conducted by [15] and [3], which showed that the LW estimator $\widehat{\alpha}_p$ displays a higher degree of negative finite-sample bias when $p = 2$ than when $p = 1$. As the LW estimator is based on the periodogram of the data, the corollary also suggests that the choice of the power p of absolute returns can affect the finite-sample behaviour of estimation based on second-order long-run dependence.

4 Monte-Carlo simulations

In this section, we present our results on Monte-Carlo simulations conducted to examine the effect of the power p on the finite sample behaviour of the LW estimator $\widehat{\alpha}_p$ under the LMSV model (1) for the returns. We carry out $2,000$ replications of sample size $n = 8192$. We employ the [4] algorithm and generate $\{h_t\}$ as a standard Gaussian $ARFIMA(0, \frac{\alpha_h}{2}, 0)$ process with $\alpha_h = 0.4, 0.8$. The process $\{u_t\}$ is generated independently of $\{h_t\}$ and is drawn as a sequence of i.i.d. standard Gaussian variables. We set $\sigma = \sigma_h = 1$. We choose the powers $p = 0.125, 0.25, ..., 2$ in the $p-$th power transformation of the absolute simulated returns. We take the bandwidth parameter to be $m = [n^{0.6}]$. We calculate the Monte-Carlo bias, standard deviation and RMSE. The Monte-Carlo bias and RMSE are presented in Figures 3-6 in Appendix B.

Concentrating on the effect of the power p, it is clear from the figures that the finite-sample bias and RMSE tends to increase in absolute value with p. The magnitude of the finite-sample bias is smallest at either at $p = 0.25$ or $p = 0.375$. The differences across the biases for $p = 0.125, .., 0.5$ are small as evident from the figures. The finite-sample bias of $\widehat{\alpha}_p$ when $p = 2$ is approximately twice as large of that when $p = 1$. It is also interesting to notice the difference in the biases for the two memory parameters $\alpha_h = 0.4, 0.8$; the finite-sample bias increases faster (in absolute value) with p when $\alpha_h = 0.8$ than when $\alpha_h = 0.4$.

The results of the Monte-Carlo simulations support our theoretical findings. Under the LMSV model (1) for the returns $\{r_t\}$, the finite-sample bias of the LW estimator applied to the transformation $\{|r_t|^p\}$ is affected by the choice of the power p and this effect depends on the long memory parameter α_h.

5 Conclusions

This paper considers the LMSV model for asset returns and examines the effect of the power p on the finite-sample behaviour of the LW estimator applied to powers of absolute returns $\{|r_t|^p\}$. We find that the finite-sample bias of the LW estimator of $\{|r_t|^p\}$ is affected by the choice of the power p. The Monte-Carlo experiment conducted is in line with our theoretical findings and suggests that the finite-sample bias of the LW estimator of $\{|r_t|^p\}$ is smallest when $p = 0.25$ or $p = 0.375$ and is increasing for bigger powers p.

There are two main conclusions to be drawn from our results. Firstly, for the estimation of the long memory parameter in the LMSV model for the returns, absolute returns are more appropriate than squared returns. This should not come as a surprise, as it is not the first time that absolute returns have been found to outperform squared returns. In their empirical study, [10] found that measures of the volatility based on absolute returns outperform in terms of predictability the equivalent measures based on squared returns; [9] showed that volatility measures based on absolute returns have more desirable properties that those based on squared returns. Secondly, the finite-sample behaviour of statistics applied to powers of returns and which are based on second-order moments is likely to be affected by the choice of the power. The latter conclusion would suggest that the Taylor effect observed in empirical applications can be also driven by the finite-sample properties of the sample autocorrelation function used to identify the dependence in powers of absolute returns.

Appendix A

Proof of Theorem 1. For any $p > 0$, we have under the LMSV model (1) that

$$\begin{aligned}
|r_t|^p &= \sigma^p E\left(|u_t|^p\right) \exp\left(p\sigma_h h_t\right) \\
&\quad + \sigma^p \left(|u_t|^p - E\left(|u_t|^p\right)\right) \exp\left(p\sigma_h h_t\right) \\
&: = \mu(p) + y_t + z_t,
\end{aligned}$$

where

$$\begin{aligned}
y_t &= \sigma^p E\left(|u_t|^p\right) \exp\left(p\sigma_h h_t\right) - \mu_p, \\
z_t &= \sigma^p \left(|u_t|^p - E\left(|u_t|^p\right)\right) \exp\left(p\sigma_h h_t\right),
\end{aligned}$$

and

$$\mu_p = \sigma^p E\left(|u_t|^p\right) E\left(\exp\left(p\sigma_h h_t\right)\right).$$

To show equations (4) and (5), we will apply Theorem 2 of [3]. Assumption B in [3] was shown by the same authors to hold for $\{|r_t|^p\}$. So, we need to establish assumption $T(\alpha_0, \beta)$ in [3] on the spectral density $f_{|r|^p}(.)$ of $\{|r_t|^p\}$.

Under assumption A.3, we have that $\{y_t\}$ and $\{z_t\}$ are uncorrelated from each other and that $\{z_t\}$ is a sequence of zero mean uncorrelated random variables with variance σ_z^2. Hence, for all $\lambda \in (-\pi, \pi]$,

$$\begin{aligned} f_{|r|^p}(\lambda) &= f_y(\lambda) + f_z(\lambda) \\ &= f_y(\lambda) + \frac{\sigma_z^2}{2\pi}, \end{aligned} \tag{9}$$

where

$$\begin{aligned} \sigma_z^2 &= Var\left(\sigma^p\left(|u_t|^p - E\left(|u_t|^p\right)\right)\exp\left(p\sigma_h h_t\right)\right) \\ &= \sigma^{2p}\left(\frac{2^p}{\sqrt{\pi}}\Gamma\left(p+\frac{1}{2}\right) - \frac{2^p}{\pi}\Gamma^2\left(\frac{p+1}{2}\right)\right) \\ &\quad \times \exp\left(2p^2\sigma_h^2\right), \end{aligned} \tag{10}$$

using assumption A.3 and equations (20) and (21) in Lemma 1.

Next, we examine the spectral density $f_y(.)$. We have that $y_t = \sigma^p E\left(|u_t|^p\right)\exp\left(p\sigma_h h_t\right) - \mu_p := G_p(h_t)$. Since $E\left(G_p(h_t)\right) = 0$ and $E^2\left(G_p(h_t)\right) < \infty$, $\{y_t\}$ admits the Hermite expansion[2]

$$y_t = \sum_{k=1}^{\infty} \frac{c_k(p)}{k!} H_k(h_t),$$

where $H_k(x)$ is the $k-$th Hermite polynomial defined as

$$H_k(x) = (-1)^k e^{\frac{x^2}{2}} \frac{d^k(e^{-\frac{x^2}{2}})}{dx^k}, \quad x \in \mathbb{R}, \tag{11}$$

and $c_k(p)$ is the $k-$th Hermite coefficient given by

$$c_k(p) = E\left(G_p(h_t) H_k(h_t)\right). \tag{12}$$

Notice that for all $p > 0$, we have from Lemma 1 that $c_k(p) \neq 0$ for all $k = 1, 2, \ldots$

Following the steps of [3] pp. 228-229, we have that the spectral density of $\{y_t\}$ satisfies

$$f_y(\lambda) = \sum_{k=1}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(*k)}(\lambda), \tag{13}$$

where $f_h^{(*k)}(.)$ is the $k-$th order convolution of the spectral density of $\{h_t\}$ for which we have under assumption A.4 that, for $k \geq 2$:

i. If $k(1-\alpha_h) < 1$,

$$\begin{aligned} f_h^{(*k)}(\lambda) &= c_{0,h}^k C_k \lambda^{-1+k(1-\alpha_h)} \\ &\quad + o\left(\lambda^{-1+k(1-\alpha_h)}\right), \quad \text{as } \lambda \longrightarrow 0+, \end{aligned} \tag{14}$$

where $C_k = \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} |1 - l_1 - \ldots - l_{k-1}|^{-\alpha_h} |l_1|^{-\alpha_h} \times \ldots \times |l_{k-1}|^{-\alpha_h} dl_1 \ldots dl_{k-1}$.

[2] For more details on Hermite expansions see [18] and [8].

ii. If $k(1-\alpha_h) = 1$,

$$f_h^{(*k)}(\lambda) \leq C\lambda^{-\delta}, \quad \text{as } \lambda \longrightarrow 0+,$$

for any $\delta > 0$.

iii. If $k(1-\alpha_h) > 1$,

$$f_h^{(*k)}(\lambda) \leq C, \quad \text{for all } \lambda \in (-\pi, \pi]. \tag{15}$$

From (9) and (13) we have that for all $\lambda \in (-\pi, \pi]$,

$$\begin{aligned} f_{|r|^p}(\lambda) &= f_y(\lambda) + \frac{\sigma_z^2}{2\pi} \\ &= c_1^2(p) f_h(\lambda) + \frac{c_2^2(p)}{2} f_h^{(*2)}(\lambda) \\ &\quad + \sum_{k=3}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(*k)}(\lambda) + \frac{\sigma_z^2}{2\pi}. \end{aligned}$$

a) Since $\alpha_h > \frac{1}{2}$, then we deduce from (14) that $f_h^{(*2)}(\lambda) = c_{0,h}^2 C_2 \lambda^{-1+2(1-\alpha_h)} + o\left(\lambda^{-1+2(1-\alpha_h)}\right)$ as $\lambda \longrightarrow 0+$, and from (15) that for all $k \geq 3$, $f_h^{(*k)}(\lambda) \leq C$ for all $\lambda \in (-\pi, \pi]$. Hence, from assumption A.4 we have that, as $\lambda \longrightarrow 0+$

$$\begin{aligned} f_{|r|^p}(\lambda) &= c_1^2(p)\lambda^{-\alpha_h}\left(c_{0,h} + c_{1,h}\lambda^2 + o(\lambda^2)\right) \\ &\quad + \frac{c_2^2(p)}{2} c_{0,h}^2 C_2 \lambda^{-(2\alpha_h - 1)} \\ &\quad + o\left(\lambda^{-(2\alpha_h-1)}\right) + C'(p) \\ &= \lambda^{-\alpha_h}\left(c_1^2(p) c_{0,h} + \frac{c_2^2(p)}{2} c_{0,h}^2 C_2 \lambda^{1-\alpha_h}\right. \\ &\quad \left. + o\left(\lambda^{1-\alpha_h}\right)\right), \end{aligned} \tag{16}$$

where $C'(p) = \sum\limits_{k=3}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(*k)}(\lambda) + \frac{\sigma_z^2}{2\pi}$, which satisfies that $C'_p = O\left(\lambda^{-(3\alpha_h - 2)}\right)$ as $\lambda \longrightarrow 0+$ using (14) and (15).

From (16) we have that assumption $T(\alpha_0, \beta)$ in [3] is satisfied. Therefore, we can apply Theorem 2 of [3] and Lemma 1.to obtain that

$$\begin{aligned} \widehat{\alpha}_p - \alpha_h &= -\left(\frac{m}{n}\right)^{\alpha_h} \frac{c_2^2(p) c_{0,h} C_2}{2c_1^2(p)} B_{1-\alpha_h} \\ &\quad -(Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) \\ &\quad + o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^{1-\alpha_h}\right) \\ &= -\left(\frac{m}{n}\right)^{\alpha_h} p^2 \frac{\sigma_h^2 c_{0,h} C_2 B_{1-\alpha_h}}{2} \\ &\quad -(Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) \\ &\quad + o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^{1-\alpha_h}\right), \end{aligned}$$

as required.

b) Since $\alpha_h < \frac{1}{2}$, then we deduce from (15) that for all $k \geq 2$, $f_h^{(*k)}(\lambda) \leq C$ for all $\lambda \in (-\pi, \pi]$. Hence, from assumption A.4 we have that, as $\lambda \longrightarrow 0+$

$$\begin{aligned} f_{|r|^p}(\lambda) &= c_1^2(p)\,\lambda^{-\alpha_h}\left(c_{0,h} + c_{1,h}\lambda^2 + o(\lambda^2)\right) + c(p) \\ &= \lambda^{-\alpha_h}\left(c_1^2(p)\,c_{0,h} + c(p)\lambda^{\alpha_h} + o\left(\lambda^{\alpha_h}\right)\right) \quad (17) \end{aligned}$$

where $c(p) = \sum_{k=2}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(*k)}(0) + \frac{\sigma_z^2}{2\pi}$.

From (17) we have that assumption $T(\alpha_0, \beta)$ in [3] is satisfied. Therefore, we can apply Theorem 2 of [3] and obtain that

$$\begin{aligned} \widehat{\alpha}_p - \alpha_h &= -\left(\frac{m}{n}\right)^{\alpha_h} \frac{c(p)}{c_1^2(p)\,c_{0,h}} B_{a_h} \\ &\quad -(Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) \\ &\quad +o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^{\alpha_h}\right) \\ &= -\left(\frac{m}{n}\right)^{\alpha_h} C(p)\frac{B_{a_h}}{2\pi c_{0,h}} \\ &\quad -(Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) \\ &\quad +o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^{\alpha_h}\right). \end{aligned}$$

where we denote $C(p) = 2\pi \frac{c(p)}{c_1^2(p)}$. We have that

$$\begin{aligned} C(p) &= \frac{2\pi \sum_{k=2}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(*k)}(0) + \sigma_z^2}{c_1^2(p)} \\ &= \frac{\sum_{k=2}^{\infty} \frac{c_k^2(p)}{k!} S_k + \sigma_z^2}{c_1^2(p)}, \end{aligned} \tag{18}$$

where for $k = 2, 3, ...$ we denote $S_k = \sum_{l \in \mathbb{Z}} \rho_h^k(l)$ and use assumption A.4 and the properties of the Hermite polynomials (11) to deduce that $f_h^{(*k)}(0) = \frac{1}{k!} f_{H_k(h)}^{(*k)}(0) = \frac{1}{2\pi} \sum_{l \in \mathbb{Z}} \rho_h^k(l)$. Lemma 1 and equation (10) imply that (18) becomes

$$\begin{aligned} C(p) &= \sum_{k=2}^{\infty} \frac{p^{2(k-1)}\sigma_h^{2(k-1)}}{k!} S_k \\ &\quad + \left(\frac{\sqrt{\pi}\Gamma\left(p + \frac{1}{2}\right)}{\Gamma^2\left(\frac{p+1}{2}\right)} - 1\right) \frac{\exp\left(p^2\sigma_h^2\right)}{p^2\sigma_h^2}, \end{aligned}$$

to complete the proof of the theorem. ■

Lemma 1 *Suppose that assumptions A.1-A.3 hold.*

a) For any $p > 0$ and $k = 1, 2, ..$ we have that the Hermite coefficients in (12) satisfy

$$c_k(p) = \sigma^p E\left(|u_t|^p\right) p^k \sigma_h^k \exp\left(\frac{p^2\sigma_h^2}{2}\right). \tag{19}$$

b) For any $p > 0$ we have that

$$E\left(|u_t|^p\right) = \frac{2^{\frac{p}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right). \tag{20}$$

Proof. a) We have that for all $k \geq 1$,

$$\begin{aligned} c_k(p) &= E\left(G_p(h_t)\,H_k(h_t)\right) \\ &= E\left(\sigma^p E\left(|u_t|^p\right) \exp\left(p\sigma_h h_t\right) H_k(h_t)\right) \\ &\quad -\mu_p E\left(H_k(h_t)\right) \\ &= \sigma^p E\left(|u_t|^p\right) p^k \sigma_h^k \exp\left(\frac{p^2\sigma_h^2}{2}\right), \end{aligned}$$

since we have that

$$\begin{aligned} E\left(\exp\left(p\sigma_h h_t\right)\right) &= \int_{-\infty}^{\infty} \exp\left(p\sigma_h x\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\ &= \exp\left(\frac{p^2\sigma_h^2}{2}\right), \end{aligned} \tag{21}$$

and for all $k = 1, 2, ..$ that

$$\begin{aligned} E\left(\exp\left(p\sigma_h h_t\right) H_k(h_t)\right) &= \int_{-\infty}^{\infty} \exp\left(p\sigma_h x\right) H_k(x) \\ &\quad \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\ &= \exp\left(\frac{p^2\sigma_h^2}{2}\right) \int_{-\infty}^{\infty} H_k(x) \frac{1}{\sqrt{2\pi}} \\ &\quad \times \exp\left(-\frac{\left(x - p\sigma_h\right)^2}{2}\right) dx \\ &= p^k \sigma_h^k \exp\left(\frac{p^2\sigma_h^2}{2}\right). \end{aligned}$$

b) Also, we have that

$$\begin{aligned} E\left(|u_t|^p\right) &= \int_{-\infty}^{\infty} |x|^p \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\ &= \frac{2^{\frac{p}{2}}}{\sqrt{\pi}} \int_0^{\infty} x^{\frac{p-1}{2}} \exp(-x)\,dx \\ &= \frac{2^{\frac{p}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right), \end{aligned}$$

from the definition of the gamma function. ■

Appendix B

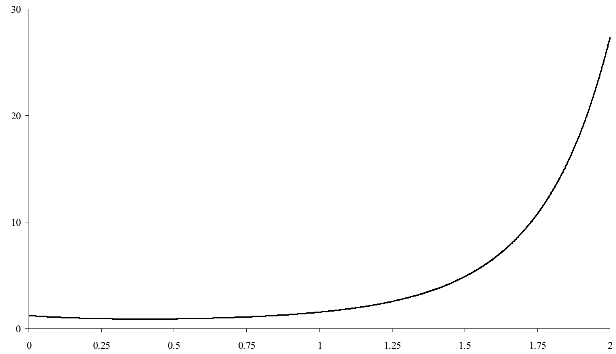

Figure 1: The graph of $C_2(p)$ in $[0, 2]$ with $\sigma_h^2 = 1$.

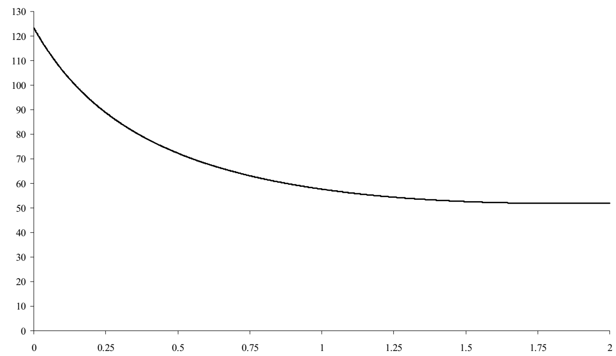

Figure 2: The graph of $C_2(p)$ in $[0, 2]$ with $\sigma_h^2 = 0.01$.

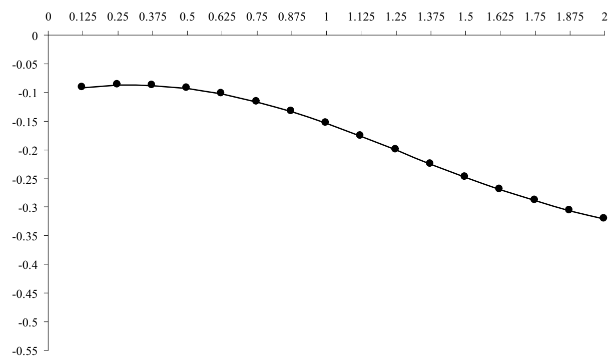

Figure 3: Bias of $\widehat{\alpha}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha_h = 0.4$, $n = 8192$ and $m = [n^{0.6}]$.

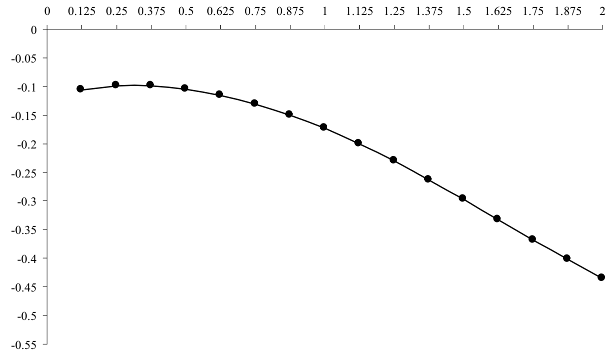

Figure 4: Bias of $\widehat{\alpha}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha_h = 0.8$, $n = 8192$ and $m = [n^{0.6}]$.

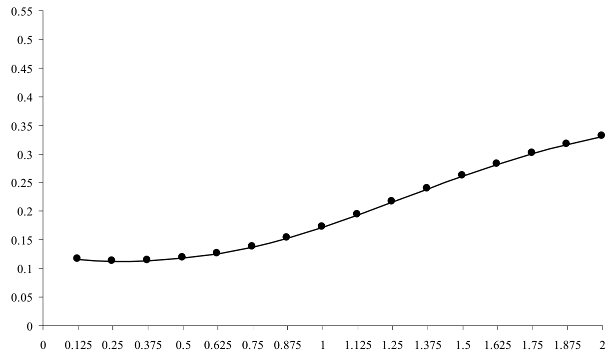

Figure 5: RMSE of $\widehat{\alpha}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha_h = 0.4$, $n = 8192$ and $m = [n^{0.6}]$.

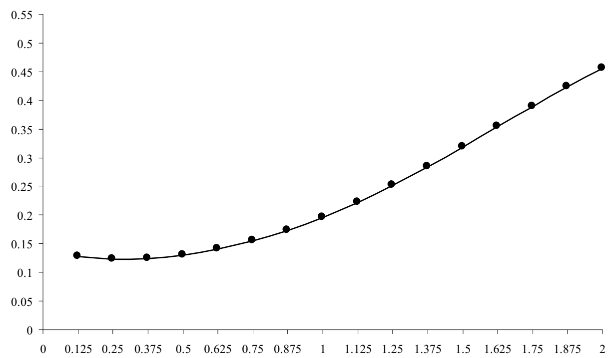

Figure 6: RMSE of $\widehat{\alpha}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha_h = 0.8$, $n = 8192$ and $m = [n^{0.6}]$.

References

[1] Adenstedt, R. K., "On large-sample estimation for the mean of a stationary random sequence," *Annals of Mathematical Statistics,* V2, N6, pp. 1095-1107, 1974.

[2] Breidt, F. J., Crato, N., de Lima, P., "The detection and estimation of long memory in stochastic volatility," *Journal of Econometrics,* V83, N2, pp. 325-348, 1998.

[3] Dalla, V., Giraitis, L., and Hidalgo, J., "Consistent estimation of the memory parameter for nonlinear time series," *Journal of Time Series Analysis,* V83, N2, pp. 325-348, 2006.

[4] Davies, R. B., Harte D. S., "Tests for Hurst effect," *Biometrika,* V74, N1, pp. 95-101, 1987.

[5] Deo, R., Hurvich C. M., "Estimation of long memory in volatility," *Theory and applications of long-range dependence,* Birkhäuser, Boston, pp. 251-301, 2003.

[6] Ding, Z., Granger, C. W. J., "Modelling volatility persistence of speculative returns: A new approach," *Journal of Econometrics,* V73, N1, pp. 185-215, 1996.

[7] Ding, Z., Granger, C. W. J., Engle, R. F., "A long memory property of stock market returns and a new model," *Journal of Empirical Finance,* V1, N1, pp. 83-106, 1993.

[8] Dobrushin, R. L., Major, P., "Non-central limit theorems for non-linear functionals of Gaussian fields," *Zeitschrift für Wahrsscheinlichkiet Theorie und Verwandte Gebiete,* V50, N1, pp. 27-52, 1979.

[9] Forsberg, L., Ghysels E., "Why do absolute returns predict volatility so well?," *Journal of Financial Econometrics,* V5, N1, pp. 31-67, 2007.

[10] Ghysels, E., Santa-Clara, P., Valkanov, R., "Predicting volatility: getting the most out of the return data sampled at different frequencies," *Journal of Econometrics,* V131, N1-2, pp. 59-95, 2006.

[11] Granger, C. W. J., Ding, Z., "Some properties of absolute return: An alternative measure of risk," *Annales d' Economic et de Statistique*, V40, pp. 67-92, 1995.

[12] Granger, C. W. J., Ding, Z., Spear, S., "Stylized facts on the temporal and distributional properties of daily data from speculative markets," Manuscript, Department of Economics, *San Diego,* 1995.

[13] Granger, C. W. J., Joyeux, R., "An introduction to long-memory time series and fractional differencing," *Journal of Time Series Analysis,* V1, N1, pp. 15-29, 1980.

[14] Harvey, A. C., "Long memory in stochastic volatility," *Forecasting volatility in financial markets,* Butterworth-Heinemann, Oxford, pp. 307-320, 1998.

[15] Hurvich C. M., Ray, B. K., "The local Whittle estimator of long-memory stochastic volatility," *Journal of Financial Econometrics,* V1, N3, pp. 445-470, 2003.

[16] Künsch, H., "Statistical aspects of self-similar process," *Proceedings of the 1st World Congress of the Bernoulli Society,* Science Press, Utrecht, pp. 67-74, 1987.

[17] Robinson, P. M., "Gaussian semiparametric estimation of long range dependence," *Annals of Statistics,* V23, N5, pp. 1630-1661, 1995.

[18] Taqqu, M. S., "Convergence of integrated processes of arbitrary Hermite rank," *Zeitschrift für Wahrsscheinlichkiet Theorie und Verwandte Gebiete,* V50, N1, pp. 53-83, 1979.

[19] Taylor, S. J., *Modelling Financial Time Series,* Wiley, New York, 1986.

[20] Wright, J. H., "Log-periodogram estimation of long memory volatility dependencies with conditionally heavy tailed returns," *Econometric Reviews,* V21, N4, pp. 397-417, 2002.

ISBN:978-988-17012-3-7