Power transformations of absolute returns and long memory estimation

Violetta Dalla*
Department of Economic Sciences
National and Kapodistrian University of Athens
Athens, Greece

This Draft: June 30, 2014

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 volatility 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. Although the true long memory parameter of powers of absolute returns is the same across 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 various power transformations of absolute returns could be due to finite-sample behavior. The local Whittle estimates of various powers of absolute returns for the S&P500 index and the DM/USD exchange rate are examined.

JEL Classification: C14, C22

Keywords: Taylor effect, long memory stochastic volatility, local Whittle estimation

*I would like to thank L. Giraitis for helpful discussions and suggestions. Tel: +30 210 3689 816. E-mail: vidalla@econ.uoa.gr.
1 Introduction

In the empirical literature, asset returns $r_t$ are commonly found to be approximately uncorrelated over time while their non-linear transformations, such as powers of absolute returns $|r_t|^p$ and their logarithms, show significant autocorrelation over many lags. The degree of the autocorrelations has been found to vary across the different non-linear transformations chosen. This was first noted by Taylor (1986), who found that for various financial series the sample autocorrelations are higher for the absolute returns $|r_t|$ than for the squared ones $r_t^2$. In a later study, Ding, Granger, and Engle (1993) examined the returns of the S&P500 index and observed that the sample autocorrelations of $|r_t|^p$ for various values of $p$ tend to be highest when $p = 1$. This observation was termed in Granger and Ding (1995) as the Taylor effect and is considered as one of the stylized facts of asset returns. Further studies on stock indices and exchange rates by Ding and Granger (1996) and Granger, Ding, and Spear (1997) found that the sample autocorrelations of $|r_t|^p$ are highest for $p = 1$ and in some cases for smaller values $p < 1$. In the aforementioned studies, the sample autocorrelations of $|r_t|$ were always found to be bigger than those of $r_t^2$, leading Malmsten and Teräsvirta (2004) to term this behavior as the Taylor effect.

Various authors have examined the Taylor effect under different specifications for the returns. Ding, Granger, and Engle (1993) used Monte-Carlo simulations to analyze the sample autocorrelations of $|r_t|^p$ for the generalized autoregressive conditionally heteroscedastic (GARCH) model of Bollerslev (1986) and the absolute value GARCH (AVGARCH) model of Taylor (1986) and Schwert (1990). For the absolute and squared returns, He and Teräsvirta (1999) derived the population autocorrelation function for the AVGARCH model and examined the sample autocorrelations of the GARCH model through Monte-Carlo simulations. He, Teräsvirta, and Malmsten (2002) evaluated the population autocorrelation function of $|r_t|^p$ under the exponential GARCH (EGARCH) model of Nelson (1991). Perez and Ruiz (2003) considered the long memory stochastic volatility (LMSV) model of Breidt, Crato, and de Lima (1998) and Harvey (1998) and examined the population autocorrelations of $|r_t|^p$ derived by Harvey (1998) as well as their sample autocorrelations through Monte-Carlo experiments. Malmsten and Teräsvirta (2004) discussed the results in He, Teräsvirta, and Malmsten (2002) and also considered the autoregressive stochastic volatility (ARSV) model of Taylor (1986), under which the population autocorrelation function of $|r_t|^p$ were derived by Ghysels, Harvey, and Renault (1996). Mora-Galán, Pérez, and Ruiz (2004) examined the sample autocorrelations of $|r_t|^p$ under the ARSV model by the means of Monte-Carlo simulations. In most cases, these models were found to reproduce the Taylor effect either in the population autocorrelation function or in its sample counterpart, although sometimes for values of the kurtosis of $r_t$ that may not be realistic.

Nevertheless, with the exception of Perez and Ruiz (2003) and Mora-Galán, Pérez, and Ruiz (2004), it is either the population autocorrelations or the sample autocorrelations that were examined. One would then wonder whether any of the observed differences in the sample autocorrelations of $|r_t|^p$ are driven by differences in the finite-sample behavior of the sample autocorrelation due
to the power transformation $p$ chosen. For that reason, Perez and Ruiz (2003) and Mora-Galán, Pérez, and Ruiz (2004) examined the relative Monte-Carlo biases of the sample autocorrelations of $|r_t|^p$ and found that, in some cases, the relative biases are affected by the choice of the power $p$. However, no theoretical arguments were provided by the authors probably due to the complication that the population autocorrelation function of $|r_t|^p$ differ across the powers $p$ in a rather complicated manner. More recently, Teräsvirta and Zhao (2011) examined a plethora of financial returns using weighted sample autocorrelations, that are robust to outliers and the degree of kurtosis, and found no clear evidence of the Taylor effect, even though the Taylor effect was present when using the usual sample autocorrelations.

Further indications that the power $p$ affects the finite-sample properties of estimators of dependence can be found in studies related to the asymptotic rate of decay of the autocorrelations of $|r_t|^p$. Under the LMSV model, Harvey (1998) established that the asymptotic rate of decay of the autocorrelations of $|r_t|^p$ is the same for all powers $p$. This common rate of decay is governed by the long memory parameter $\alpha$ in the LMSV model, so that powers of absolute returns have the same long memory parameter $\alpha$ irrespective of the power $p$. However, Monte-Carlo experiments with the LMSV model, performed by Wright (2002), Hurvich and Ray (2003), Deo and Hurvich (2003) and Dalla, Giraitis, and Hidalgo (2006), suggest that semiparametric estimators of the long memory parameter $\alpha$ of $|r_t|^p$ display higher degree of negative finite-sample bias when $p = 2$ than when $p = 1$. As these semiparametric estimators are periodogram-based, the question again arises as to whether when looking at $|r_t|^p$ 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 varying degree of dependence is also driven by the finite-sample behavior of the estimators used to identify the dependence in $|r_t|^p$.

The main purpose of this paper is to investigate the effect of the power $p$ of $|r_t|^p$ on the finite-sample behavior of estimation of long run dependence. We consider the LMSV model for the returns and examine estimation of the common long memory parameter $\alpha$ of $|r_t|^p$. We choose the local Whittle (LW) estimator of the long memory parameter introduced by Künsch (1987); for the $p$–th power of absolute returns under the LMSV model its consistency and asymptotic distribution were established by Dalla, Giraitis, and Hidalgo (2006) 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 $|r_t|^p$ 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 effect is quadratic in $p$. We also conduct a Monte-Carlo experiment as in Hurvich and Ray (2003) and Dalla, Giraitis, and Hidalgo (2006) extending the range of the powers $p$ and find that the finite-sample bias of the LW estimator of $|r_t|^p$ is severely affected by the choice of the power $p$. Finally, we discuss the LW estimates of various powers of absolute returns for two of the series examined in Ding and Granger (1996), the S&P500 index and the DM/USD exchange rate.
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. The Monte-Carlo study and the empirical application are contained in Sections 4 and 5, respectively, while Section 6 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 Breidt, Crato, and de Lima (1998) and Harvey (1998). We assume that the returns of an asset, denoted by \( r_t \), satisfy

\[
    r_t = u_t \sigma_t, \quad \sigma_t = \sigma \exp(\sigma_h h_t)
\]

where \( \sigma \) and \( \sigma_h \) are positive constants. Moreover, we assume 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 \( 0 < \alpha < 1 \), whose spectral density function \( f_{h}(\cdot) \) satisfies

\[
    f_{h}(\lambda) = \lambda^{-\alpha} \left( c_{0,h} + c_{1,h} \lambda^2 + o(\lambda^2) \right), \quad \lambda \to 0+,
\]

and autocorrelation function \( \rho_{h}(\cdot) \) has the property

\[
    \rho_{h}(\tau) \sim c_{h} \tau^{-1+\alpha}, \quad \tau \to +\infty.
\]

Assumptions A.1-A.3 are as those in Breidt, Crato, and de Lima (1998) and Harvey (1998). Under these assumptions, the latter authors derived various properties for the returns and their nonlinear transformations. They also considered that \( h_t \) is a stationary \( ARFIMA(p, \alpha/2, q) \) model, see Adenstedt (1974) and Granger and Joyeux (1980). Our assumption A.4 is satisfied by stationary \( ARFIMA(p, \alpha/2, q) \) models and is employed in Dalla, Giraitis, and Hidalgo (2006) to show the consistency of the LW estimator applied to powers of absolute returns under the LMSV model (1).

Following Harvey (1998), under assumptions A.1 and A.3 we have that the autocorrelation function \( \rho_{p}(\cdot) \) of \( |r_t|^p \) satisfies

\[
    \rho_{p}(\tau) = \frac{\exp\left( p^2 \sigma_h^2 \rho_{h}(\tau) \right) - 1}{\kappa_p \exp\left( p^2 \sigma_h^2 \right) - 1}, \quad \tau \geq 1,
\]

where
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. Hence,

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

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

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

As noted in Harvey (1998), it is clear from (2) that it is not possible to make statements about which power $p$ maximizes the autocorrelation function of $|r_t|^p$. Nevertheless, 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 $\alpha$. So, we focus on the estimation of this common long memory parameter $\alpha$ to examine the effect that the power $p$ has in estimating long run dependence in $|r_t|^p$.

For the estimation of the long memory parameter, we use the LW estimator, see Künsch (1987) and Robinson (1995). Given a generic set of data $x_1, \ldots, x_n$, the LW estimator $\hat{\alpha}_x$ is defined as

$$
\hat{\alpha}_x = \arg \min_{\alpha \in [-1,1]} U_n (\alpha),
$$

of the objective function

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

where $\lambda_j = \frac{2\pi j}{n}$, $j = 1, \ldots, n$ 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

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

Here, we consider the LW estimator applied to $|r_t|^p$, which we denote by $\hat{\alpha}_p$. In the LMSV model with assumptions A.1-A.4, Dalla, Giraitis, and Hidalgo (2006) showed the consistency of the LW estimator $\hat{\alpha}_p$. Under additional assumptions, they also established the asymptotic distribution
of $\hat{\alpha}_p$, which was found to be independent of the power $p$. However, the Monte-Carlo experiments in Hurvich and Ray (2003) and Dalla, Giraitis, and Hidalgo (2006) for the cases $p = 1$ and $p = 2$ suggest that the finite-sample behavior and particularly the finite-sample bias of the LW estimator $\hat{\alpha}_p$ 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 behavior of periodogram-based estimators applied to $|r_t|^p$.

3 Power transformations and LW estimation

In this section we provide our theoretical results on the finite-sample behavior of the LW estimator $\hat{\alpha}_p$ applied to $|r_t|^p$. Recall from the discussion in Section 2 above that for all powers $p$, the true long memory parameter of $|r_t|^p$ is $\alpha$. We use the notation

$$B_\beta = (2\pi)^{\beta/2}(\beta + 1)^{2}\quad \text{and} \quad Q_{m,p} = \frac{1}{m}\sum_{j=1}^{m} \left( \log \left( \frac{j}{m} \right) + 1 \right) \left( c_1(p)c_{0,h} \right)^{-1} \lambda_j^{\alpha} I_{\alpha}(\lambda_j),$$

where $c_1(p)$ is given in (12) below. The proof of the following proposition is found in Appendix A.

**Proposition 1** Suppose that $r_t$ follows the LMSV model (1) and that assumptions A.1-A.4 are satisfied.

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

$$\hat{\alpha}_p - \alpha = -\left( \frac{m}{n} \right)^{1-\alpha} \frac{\sigma_h^2 c_{0,h} C_\alpha B_{1-\alpha}}{2} p^2 - (Q_{m,p} - E(Q_{m,p}))(1 + o_P(1)) + o_P \left( m^{-1/2} + \left( \frac{m}{n} \right)^{1-\alpha} \right),$$

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

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

$$\hat{\alpha}_p - \alpha = -\left( \frac{m}{n} \right)^{\alpha} \frac{B_{\alpha}}{2\pi c_{0,h}} C(p, \alpha) - (Q_{m,p} - E(Q_{m,p}))(1 + o_P(1)) + o_P \left( m^{-1/2} + \left( \frac{m}{n} \right)^{\alpha} \right),$$

where

$$C(p, \alpha) = C_1(p, \alpha) + C_2(p) = \sum_{k=2}^{\infty} \frac{\sigma_h^{2(k-1)}S_k(\alpha)}{k!} = \sum_{l \in \mathbb{Z}} \rho_h^k(l), \quad \text{for } k = 2, 3, ...$$
For any power $p > 0$, equations (4) and (5) provide the deviation of the estimator $\hat{\alpha}_p$ from its true value $\alpha$. Notice that in both equations (4) and (5), the first term is non-stochastic and the second term has zero mean. For relatively large fixed $n$, the last term in (4) and (5) will become negligible. Therefore, for any $p > 0$, the dominant term in the finite-sample bias of the estimator $\hat{\alpha}_p$ is given by

$$-(\frac{m}{n})^{1-\alpha} \frac{\sigma_h^2}{2} C_{\alpha} B_1 - \alpha p^2, \quad \text{if } \alpha > \frac{1}{2},$$

or by

$$-(\frac{m}{n})^\alpha \frac{B_\alpha}{2\pi \sigma_h} C(p, \alpha), \quad \text{if } \alpha < \frac{1}{2}. \quad (7)$$

It is clear from (7) and (8) that the $p$-th power transformation of absolute returns affects the finite-sample bias of the LW estimator. If $\alpha > \frac{1}{2}$, the effect of $p$ is quadratic. If $\alpha < \frac{1}{2}$, the effect of $p$ comes through the quantity $C(p, \alpha)$. Notice that $C(p, \alpha)$ depends on the power $p$ and the long memory parameter $\alpha$ through the autocorrelation function of $h_t$ in the sums $S_k(\alpha)$. 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 $\alpha < \frac{1}{2}$, we have that $S_k(\alpha)$ is a positive finite quantity for all $k = 2, 3, \ldots$. So, $C_1(p, \alpha)$ is a strictly increasing function for $p > 0$. On the other hand, $C_2(p)$ depends on $p$ but not on $\alpha$. One can easily show that it is strictly increasing in $p$ when $p > \frac{1}{\sigma_h}$. For $p \leq \frac{1}{\sigma_h}$, $C_2(p)$ can be increasing and/or decreasing in $p$ depending on the value of $\sigma_h$, see Figures 1 and 2. Therefore, the function $C(p, \alpha)$ is strictly increasing in $p$ for all $p > \frac{1}{\sigma_h}$. For $p \leq \frac{1}{\sigma_h}$, the shape of $C(p, \alpha)$ depends on the value of $\sigma_h$ and whether $C_1(p, \alpha)$ dominates $C_2(p)$ or not. However, even if $\sigma_h$ were known, there are practical obstacles in the calculation of the derivative of $C_1(p, \alpha)$, since $C_1(p, \alpha)$ depends also on the autocorrelation function of $h_t$ for which no parametric model has been specified.

It is also interesting to notice the effect of the kurtosis of the returns on the finite-sample bias of the LW estimator. In the LMSV model the kurtosis is controlled through the parameter $\sigma_h$, with higher values of kurtosis corresponding to higher values of $\sigma_h$. If $\alpha > \frac{1}{2}$, it is evident from (7) that for fixed power $p$ the finite-sample bias of the estimator $\hat{\alpha}_p$ increases in absolute terms with $\sigma_h$. When $\alpha < \frac{1}{2}$, the function $C(p, \alpha)$ in (8) is strictly increasing in $\sigma_h$ when $\sigma_h > \frac{1}{p}$. As above, for $\sigma_h \leq \frac{1}{p}$ we cannot conclude what the effect will be.

**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 $p$ chosen. In particular, when $\alpha > \frac{1}{2}$ or when $\alpha < \frac{1}{2}$ and $p > \frac{1}{\sigma_h}$, the dominant term in the finite-sample bias of the LW estimator $\hat{\alpha}_p$ increases (in absolute value) as $p$ increases.

---

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

The results help explain the outcomes of the Monte-Carlo experiments conducted by Hurvich and Ray (2003) and Dalla, Giraitis, and Hidalgo (2006), which found that the LW estimator $\hat{\alpha}_p$ 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 results also suggest that the choice of the power $p$ of absolute returns can affect the finite-sample behavior of estimation of long-run dependence.

4 Monte-Carlo simulations

In this section, we present Monte-Carlo simulations conducted to examine the effect of the power $p$ on the finite sample behavior of the LW estimator $\hat{\alpha}_p$ under the LMSV model (1) for the returns. We carry out 5,000 replications of sample size $n = 8192$. We employ the Davies and Harte (1987) algorithm and generate $h_t$ as a standard Gaussian ARFIMA$(0, \alpha/2, 0)$ process with $\alpha = 0.4, 0.8$. 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 $\sigma = 1$ and choose $\sigma_h = 0.5, 0.8$ implying a kurtosis of approximately 8.2 and 36.5, respectively, which are the sample kurtosis of the returns examined in Section 5 below. 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 of the form $m = [n^\gamma]$ with $\gamma = 0.5, 0.6, 0.7$. We calculate the Monte-Carlo bias, standard deviation and RMSE. The Monte-Carlo bias and RMSE are presented in Figures 3-14 in Appendix B.

It is evident that in all the cases, the finite-sample bias and RMSE of $\hat{\alpha}_p$ tend to increase in absolute terms with $m$, and the bias is negative. Concentrating on the effect of the power $p$, it is clear from the figures that the finite-sample bias is affected by the choice of power $p$. The magnitude of the finite-sample bias is smallest at $p = 0.75$ or 0.875 when $\sigma_h = 0.5$ and at $p = 0.375$ or 0.5 when $\sigma_h = 0.8$. As expected, the bias is stronger when $\sigma_h = 0.8$ and so the overall effect of $p$ in this case is more dramatic. It is also interesting to notice the difference in the biases for the two memory parameters $\alpha = 0.4, 0.8$; the finite-sample bias increases faster in absolute value with $p$ when $\alpha = 0.8$ than when $\alpha = 0.4$.

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 $\alpha$ and the kurtosis of the returns $r_t$.

5 Empirical application

We consider two of the series examined in Ding and Granger (1996), the S&P500 index and the DM/USD foreign exchange rate. We calculate daily returns for the S&P500 index from 04/01/1950 to 30/05/2008 with 14,695 observations, while the daily returns for the DM/USD exchange rate are
for the period 05/01/1971 to 31/12/1998 giving a sample of 7,020 observations. We plot the sample autocorrelation functions of the powers of absolute returns $|r_t|^p$ of the S&P500 index in Figure 15 and in Figure 16 for the DM/USD exchange rate using $p = 0.25, 0.5, ..., 2$. We evaluate the powers of absolute returns $|r_t|^p$ using $p = 0.125, 0.13, ..., 2$ and apply the LW estimator $\hat{\alpha}_p$ using bandwidth $m = [n^{0.5}]$. Moreover, we calculate the lower and upper bound of the 95% confidence interval of the LW estimates for $p = 0.125, 0.13, ..., 2$ using the asymptotic distribution of the estimator derived by Dalla, Giraitis, and Hidalgo (2006). The results are presented in Figures 17 and 18 for the S&P500 index and the DM/USD foreign exchange rate, respectively.

Figures 15 and 16 of the sample autocorrelations of $|r_t|^p$ are in-line with those in Ding and Granger (1996). The autocorrelations of $|r_t|^p$ tend to be highest when $p \approx 1$ in the case of the S&P500 index and when $p = 0.25$ for the DM/USD exchange rate. The other main difference between the two plots is that for the DM/USD exchange rate the sample autocorrelations of $|r_t|^p$ seem to be decreasing when $p$ increases, while this happens only for $p \geq 1$ in the case of the S&P500 index.

The latter behavior is somewhat similar to what we observe in the plots of the LW estimates $\hat{\alpha}_p$. It is clear that the LW estimates $\hat{\alpha}_p$ are decreasing when $p$ increases in the case of the DM/USD foreign exchange rate. On the other hand, the LW estimates $\hat{\alpha}_p$ for the S&P500 index achieve a maximum at $p = 0.375$, are slowly increasing when $p < 0.375$ and decrease rather rapidly for $p > 0.375$. It is also interesting to notice that for S&P500 index, the shape of the LW estimates resembles that in the Monte-Carlo simulations with $\sigma_h = 0.8$.

Finally, it is worth noting that the 95% confidence intervals do not intersect for very small $p$ and for $p = 2$. This may lead someone to conclude that the different power transformations produce different values for the long memory parameter, and therefore models, such as the LMSV, would be excluded. However, keeping in mind the theoretical results and the Monte-Carlo simulations, one cannot exclude that the graphs on the LW estimates $\hat{\alpha}_p$ are driven by the finite-sample behavior of the LW estimator so that the long memory parameter is the same for the various powers $p$ and a LMSV model may be appropriate.

6 Conclusions

This paper considers the LMSV model for asset returns and examines the effect of the power $p$ on the finite-sample behavior 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 in the context of LMSV models the finite-sample bias of the LW estimator of $|r_t|^p$ is smallest for low values of $p$. We apply the LW estimator to the powers of absolute returns for the S&P500 index and the DM/USD foreign exchange rate and find LW estimates that vary across the powers $p$ with a behavior similar to that in our theoretical findings and simulations. Therefore, one cannot exclude
the possibility that the long memory parameter is the same for the different powers \( p \), although the LW estimates suggest otherwise.

Although the focus of this paper is on powers of absolute returns \(|r_t|^p\) and the effect of the power \( p \), it is worth discussing the effect of logarithmic squared transformation that is commonly used, \( \ln r_t^2 \). It is well known that under the LMSV model, \( \ln r_t^2 \) is written as a signal plus noise model, where the signal is a linear process and the noise is i.i.d., see for example Hurvich, Moulines, and Soulier (2005). Here, we show that \(|r_t|^p\) can also be written as a signal plus noise model, but the noise is no longer i.i.d. but a non-linear process possibly with long memory. Therefore, one might expect the LW estimator to have better finite-sample properties when looking at \( \ln r_t^2 \) than at \(|r_t|^p\). This has been found in the Monte-Carlo experiments for the case \( p = 2 \), see Dalla, Giraitis, and Hidalgo (2006). However, the signal plus noise decomposition for \( \ln r_t^2 \) is driven by the exponential function in the volatility model and so it is hard to generalize that the logarithmic squared transformation should be preferred. Actually, Monte-Carlo simulations on \( \ln r_t^2 \) under the LMSV model (not presented here) suggest that there are powers \( p \) for which the finite-sample behavior of the LW estimator is slightly better for \(|r_t|^p\) than for \( \ln r_t^2 \), especially in the case of low excess kurtosis. Various bias corrections adjustments have been introduced for the LW estimator in the case of the \( \ln r_t^2 \) under the LMSV model, see Hurvich and Ray (2003) and Hurvich, Moulines, and Soulier (2005). These bias-corrected estimators could potentially be adapted for \(|r_t|^p\). However the Monte-Carlo results of Gonçalves da Silva and Robinson (2008) warrant against the use of these estimators as the finite-sample biases are reduced in the expense of high standard deviation and bimodal distribution.

There are two main conclusions to be drawn from our results. Firstly, for the estimation of the long memory parameter of powers of absolute returns under the LMSV model there is no unique power \( p \) that produces the smallest finite-sample bias. Nevertheless, there is a tendency for small values of \( p \) to be preferable, so that absolute returns are likely to be more appropriate than squared ones. 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, Ghysels, Santa-Clara, and Valkanov (2006) found that measures of the volatility based on absolute returns outperform in terms of predictability the equivalent measures based on squared returns. Moreover, Forsberg and Ghysels (2007) showed that volatility measures based on absolute returns have more desirable properties in the presence of sampling errors and jumps than those based on squared returns.

Secondly, the finite-sample behavior 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 and the kurtosis of the returns. The latter conclusion is in line with the results of Teräsvirta and Zhao (2011) and suggests that the Taylor effect can also be driven by the finite-sample properties of the sample autocorrelation function of the powers of absolute returns. This raises doubts as to whether the Taylor effect should be considered a stylized fact and warrants against the rejections of models solely based on the dependence of powers of absolute returns.
A Appendix

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

$$|r_t|^p = \sigma^p E(|u_t|^p) \exp(p \sigma_h h_t) + \sigma^p (|u_t|^p - E(|u_t|^p)) \exp(p \sigma_h h_t)$$

where

$$y_t = \sigma^p E(|u_t|^p) \exp(p \sigma_h h_t) - \mu_p,$$

$$z_t = \sigma^p (|u_t|^p - E(|u_t|^p)) \exp(p \sigma_h h_t),$$

and

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

To show equations (4) and (5), we will apply Theorem 2 of Dalla, Giraitis, and Hidalgo (2006). Assumption B in Dalla, Giraitis, and Hidalgo (2006) was shown by the same authors to hold for $|r_t|^p$. So, we need to establish Assumption $T(\alpha_0, \beta)$ in Dalla, Giraitis, and Hidalgo (2006) on the spectral density $f_{|r|^p} (\cdot)$ 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 $\sigma_z^2$. Hence, for all $\lambda \in (-\pi, \pi]$,

$$f_{|r|^p} (\lambda) = f_y (\lambda) + f_z (\lambda)$$

$$= f_y (\lambda) + \frac{\sigma_z^2}{2\pi},$$

where

$$\sigma_z^2 = \text{Var} (\sigma^p (|u_t|^p - E(|u_t|^p)) \exp(p \sigma_h h_t))$$

$$= \sigma^2 \left( E \left( |u_t|^{2p} \right) - E^2 (|u_t|^p) \right) E(\exp(2p \sigma_h h_t))$$

$$= \sigma^2 \left( \frac{2^p}{\sqrt{\pi}} \Gamma \left( p + \frac{1}{2} \right) - \frac{2^p}{\pi} \Gamma^2 \left( \frac{p + 1}{2} \right) \right) \exp(2p^2 \sigma_h^2),$$

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

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

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

\(^2\)For more details on Hermite expansions see Taqqu (1979) and Dobrushin and Major (1979).
where $H_k(x)$ is the $k$–th Hermite polynomial defined as

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

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

$$c_k(p) = E(G_p(h_t) H_k(h_t)). \quad (12)$$

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 Dalla, Giraitis, and Hidalgo (2006) pp. 228-229, we have that the spectral density of $y_t$ satisfies

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

where $f_h^{(sk)}(\cdot)$ 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 $k(1-\alpha) < 1$,

$$f_h^{(sk)}(\lambda) = c_{0,b}^k C_k \lambda^{-1+k(1-\alpha)} + o\left(\lambda^{-1+k(1-\alpha)}\right), \quad \text{as } \lambda \to 0+, \quad (14)$$

where $C_k = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} |l_1 - \ldots - l_{k-1}|^{-\alpha} |l_1|^{-\alpha} \ldots |l_{k-1}|^{-\alpha} dl_1 \ldots dl_{k-1}$.

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

$$f_h^{(sk)}(\lambda) \leq C \lambda^{-\delta}, \quad \text{as } \lambda \to 0+, \quad \text{for any } \delta > 0.$$

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

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

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

$$f_{\rho}\rho(\lambda) = f_y(\lambda) + \frac{\sigma^2_\rho}{2\pi}$$

$$= c_1^2(p) f_h(\lambda) + \frac{c_2^2(p)}{2} f_h^{(s2)}(\lambda) + \sum_{k=3}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(sk)}(\lambda) + \frac{\sigma^2_\rho}{2\pi}.$$

a) Since $\alpha > \frac{1}{2}$, then we deduce from (14) that $f_h^{(s2)}(\lambda) = c_{0,b}^2 C_\alpha \lambda^{1-2\alpha} + o(\lambda^{1-2\alpha})$ as $\lambda \to 0+$, and from (15) that for all $k \geq 3$, $f_h^{(sk)}(\lambda) \leq C$ for all $\lambda \in (-\pi, \pi]$. Hence, from assumption A.4 we
have that, as \( \lambda \to 0^+ \)

\[
f_{|r|^p}(\lambda) = c_1^2(p) \lambda^{-\alpha} \left( c_{0,h} + c_{1,h} \lambda^2 + o(\lambda^2) \right) + \frac{c_2^2(p)}{2} c_{0,h}^2 C_\alpha \lambda^{1-2\alpha} + o\left( \lambda^{1-2\alpha} \right) + C'(p)
\]

\[
= \lambda^{-\alpha} \left( c_1^2(p) c_{0,h} + c_1^2(p) c_{1,h} \lambda^2 + o(\lambda^2) + \frac{c_2^2(p)}{2} c_{0,h}^2 C_\alpha \lambda^{1-\alpha} + o(\lambda^{1-\alpha}) \right)
\]

\[
= \lambda^{-\alpha} \left( c_1^2(p) c_{0,h} + \frac{c_2^2(p)^2}{2} c_{0,h}^2 C_\alpha \lambda^{1-\alpha} + o(\lambda^{1-\alpha}) \right),
\]

where \( C'(p) = \sum_{k=3}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(sk)}(\lambda) + \frac{c_2^2}{2\pi} \), which satisfies that \( C'_p = O(\lambda^{2-3\alpha}) \) as \( \lambda \to 0^+ \) using (14) and (15).

From (16) we have that assumption Assumption \( T(\alpha_0, \beta) \) in Dalla, Giraitis, and Hidalgo (2006) is satisfied. Therefore, we can apply Theorem 2 of Dalla, Giraitis, and Hidalgo (2006) and Lemma 1 to obtain that

\[
\tilde{\alpha}_p - \alpha = - \left( \frac{m}{n} \right) \alpha c_2^2(p) c_{0,h} C_\alpha B_{1-\alpha} - (Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) + o_p \left( m^{-\frac{1}{2}} + \left( \frac{m}{n} \right)^{1-\alpha} \right)
\]

\[
= - \left( \frac{m}{n} \right)^\alpha \sigma_h^2 c_{0,h} C_\alpha B_{1-\alpha} - (Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) + o_p \left( m^{-\frac{1}{2}} + \left( \frac{m}{n} \right)^{1-\alpha} \right),
\]

as required.

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

\[
f_{|r|^p}(\lambda) = c_1^2(p) \lambda^{-\alpha} \left( c_{0,h} + c_{1,h} \lambda^2 + o(\lambda^2) \right) + C(p, \alpha)
\]

\[
= \lambda^{-\alpha} \left( c_1^2(p) c_{0,h} + c_1^2(p) c_{1,h} \lambda^2 + o(\lambda^2) + C(p, \alpha) \lambda^\alpha \right)
\]

\[
= \lambda^{-\alpha} \left( c_1^2(p) c_{0,h} + c(p, \alpha) \lambda^\alpha + o(\lambda^\alpha) \right),
\]

where \( c(p, \alpha) = \sum_{k=2}^{\infty} \frac{c_k^2(p)}{k!} f_h^{(sk)}(0) + \frac{c_2^2}{2\pi} \).

From (17) we have that assumption Assumption \( T(\alpha_0, \beta) \) in Dalla, Giraitis, and Hidalgo (2006) is satisfied. Therefore, we can apply Theorem 2 of Dalla, Giraitis, and Hidalgo (2006) and obtain
that
\[
\hat{\alpha}_p - \alpha = -\left(\frac{m}{n}\right)^\alpha \frac{c(p, \alpha)}{c_1^2(p)c_{0,h}} B_{a_h} - (Q_{m,p} - E(Q_{m,p}))(1 + o_p(1)) + o_p\left(m^{-\frac{1}{2}} + \left(\frac{m}{n}\right)^\alpha\right).
\]

where we denote \(C(p, \alpha) = 2\pi \frac{c(p, \alpha)}{c_1^2(p)}\). We have that
\[
C(p, \alpha) = 2\pi \frac{C(p, \alpha)}{c_1^2(p)} = 2\pi \sum_{k=2}^{\infty} \frac{c_2^2(p)}{c_1^2(p)} f_h^{(sk)}(0) + \sigma_z^2
\]
\[
= \sum_{k=2}^{\infty} \frac{c_2^2(p)}{c_1^2(p)} S_k(\alpha) + \sigma_z^2
\]
(18)

where for \(k = 2, 3, \ldots \) we denote \(S_k(\alpha) = \sum_{l \in \mathbb{Z}} \rho_h^k(l)\) and use assumption A.4 and the properties of the Hermite polynomials (11) to deduce that \(f_h^{(sk)}(0) = \frac{1}{k!} f_h^{(sk)}(0) = \frac{1}{2\pi} \sum_{l \in \mathbb{Z}} \rho_h^k(l)\). Lemma 1 and equation (10) imply that (18) becomes
\[
C(p, \alpha) = \sum_{k=2}^{\infty} \frac{p^{2(k-1)} \sigma_h^{2(k-1)}}{k!} S_k(\alpha) + \left(\frac{\sqrt{\pi} \Gamma \left(p + \frac{1}{2}\right)}{\Gamma \left(p + \frac{3}{2}\right)} - 1\right) \exp\left(\frac{p^2 \sigma_h^2}{2}\right),
\]
to complete the proof of the proposition.

\begin{lemma}
Suppose that assumptions A.1-A.3 hold.
\end{lemma}

\begin{enumerate}
\item For any \(p > 0\) and \(k = 1, 2, \ldots\) we have that the Hermite coefficients in (12) satisfy
\[
c_k(p) = \sigma^p E\left(|u_t|^p\right) p^k \sigma_h^k \exp\left(\frac{p^2 \sigma_h^2}{2}\right).
\]
(19)
\item For any \(p > 0\) we have that
\[
E\left(|u_t|^p\right) = \frac{2^p \sqrt{\pi}}{\Gamma \left(p + \frac{1}{2}\right)}.
\]
(20)
\end{enumerate}
Proof. a) We have that for all $k \geq 1$,

\[
c_k (p) = E (G_p (h_t) H_k (h_t)) = E (\sigma^p E (|u_t|^p) \exp (p\sigma h_t) H_k (h_t)) - \mu_p E (H_k (h_t)) = \sigma^p E (|u_t|^p) E (\exp (p\sigma h_t) H_k (h_t)) = \sigma^p E (|u_t|^p) \rho^k \sigma^k \exp \left( \frac{p^2 \sigma^2_h}{2} \right),
\]

since we have that

\[
E (\exp (p\sigma h_t)) = \int_{-\infty}^{\infty} \exp (p\sigma x) \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx = \exp \left( \frac{p^2 \sigma^2_h}{2} \right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{(x - p\sigma_h)^2}{2} \right) dx = \exp \left( \frac{p^2 \sigma^2_h}{2} \right),
\]

and for all $k = 1, 2, \ldots$ that

\[
E (\exp (p\sigma h_t) H_k (h_t)) = \int_{-\infty}^{\infty} \exp (p\sigma_h x) H_k (x) \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx = \exp \left( \frac{p^2 \sigma^2_h}{2} \right) \int_{-\infty}^{\infty} H_k (x) \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{(x - p\sigma_h)^2}{2} \right) dx = p^k \sigma^k \exp \left( \frac{p^2 \sigma^2_h}{2} \right).
\]

b) Also, we have that

\[
E (|u_t|^p) = \int_{-\infty}^{\infty} |x|^p \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) dx = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} x^p \exp \left( -\frac{x^2}{2} \right) dx = \frac{2^p}{\sqrt{\pi}} \int_{0}^{\infty} x^{p-1} \exp (-x) dx = \frac{2^p}{\sqrt{\pi}} \Gamma \left( \frac{p + 1}{2} \right),
\]

from the definition of the gamma function. ■
B Appendix

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

Figure 2: The graph of the function $C_2(p)$ in $[0, 2]$ with $\sigma_h = 0.1$. 
Figure 3: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, \ldots, 2$, when $\alpha = 0.4$, $\sigma_h = 0.5$, $n = 8192$ and $m = \lceil n^{0.5} \rceil$.

Figure 4: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, \ldots, 2$, when $\alpha = 0.4$, $\sigma_h = 0.5$, $n = 8192$ and $m = \lceil n^{0.6} \rceil$.

Figure 5: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, \ldots, 2$, when $\alpha = 0.4$, $\sigma_h = 0.5$, $n = 8192$ and $m = \lceil n^{0.7} \rceil$. 
Figure 6: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, ... , 2$, when $\alpha = 0.8$, $\sigma_h = 0.5$, $n = 8192$ and $m = [n^{0.5}]$. 

Figure 7: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, ... , 2$, when $\alpha = 0.8$, $\sigma_h = 0.5$, $n = 8192$ and $m = [n^{0.6}]$. 

Figure 8: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, ... , 2$, when $\alpha = 0.8$, $\sigma_h = 0.5$, $n = 8192$ and $m = [n^{0.7}]$. 

18
Figure 9: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, \ldots, 2$, when $\alpha = 0.4$, $\sigma_h = 0.8$, $n = 8192$ and $m = [n^{0.5}]$.

Figure 10: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, \ldots, 2$, when $\alpha = 0.4$, $\sigma_h = 0.8$, $n = 8192$ and $m = [n^{0.6}]$.

Figure 11: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, \ldots, 2$, when $\alpha = 0.4$, $\sigma_h = 0.8$, $n = 8192$ and $m = [n^{0.7}]$. 
Figure 12: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha = 0.8$, $\sigma_h = 0.8$, $n = 8192$ and $m = [n^{0.5}]$.

Figure 13: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha = 0.8$, $\sigma_h = 0.8$, $n = 8192$ and $m = [n^{0.6}]$.

Figure 14: Bias (left) and RMSE (right) of $\hat{a}_p$ for $p = 0.125, 0.25, ..., 2$, when $\alpha = 0.8$, $\sigma_h = 0.8$, $n = 8192$ and $m = [n^{0.7}]$. 
Figure 15: Sample autocorrelations at lags 1, ..., 200 of the powers $p = 0.25, ..., 2$ of absolute returns of the S&P500 index.

Figure 16: Sample autocorrelations at lags 1, ..., 200 of the powers $p = 0.25, ..., 2$ of absolute returns of the DM/USD exchange rate.
Figure 17: LW estimates (black line) with the upper and lower bound of their 95% confidence intervals (dotted black lines) for the powers $p = 0.125, 0.13, ..., 2$ of absolute returns of the S&P500 index.

Figure 18: LW estimates (black line) with the upper and lower bound of their 95% confidence intervals (dotted black lines) for the powers $p = 0.125, 0.13, ..., 2$ of absolute returns of the DM/USD exchange rate.
References


