A new generalized volatility proxy via the stochastic volatility model

Jong-Min Kim, Hojin Jung & Li Qin

To cite this article: Jong-Min Kim, Hojin Jung & Li Qin (2017) A new generalized volatility proxy via the stochastic volatility model, Applied Economics, 49:23, 2259-2268, DOI: 10.1080/00036846.2016.1237751

To link to this article: http://dx.doi.org/10.1080/00036846.2016.1237751

Published online: 18 Nov 2016.
A new generalized volatility proxy via the stochastic volatility model

Jong-Min Kim, Hojin Jung and Li Qin

ABSTRACT
This article proposes power transformation of absolute returns as a new proxy of latent volatility in the stochastic model. We generalize absolute returns as a proxy for volatility in that we place no restriction on the power of absolute returns. An empirical investigation on the bias, mean square error and relative bias is carried out for the proposed proxy. Simulation results show that the new estimator exhibiting negligible bias appears to be more efficient than the unbiased estimator with high variance.

KEYWORDS
Volatility; stochastic volatility; relative bias; mean square error

JEL CLASSIFICATION
C32; C52

I. Introduction

The volatility process is essential in order to measure financial asset returns. Two main streams of volatility models are the generalized autoregressive conditional heteroscedasticity (GARCH) model and the stochastic volatility (hereafter SV) model in the literature. However, these two models are very different because of their assumptions on the volatility states and on the error term. The former assumes deterministic volatility states and a single error term while the latter assumes stochastic volatility states and two error processes which can make the model more flexible than the GARCH model. Since volatility is latent, proxies are required when estimating volatility models. The SV model proposed by Taylor (1982) has received substantial attention for modeling volatility of financial time series in the last three decades. The most common proxy variables for volatility in financial markets are squared returns and absolute returns. The previous literature (see Taylor 1986; Triacca 2007; Giles 2008) uses the basic SV model to examine properties of the squared returns and absolute returns. The previous literature (see Taylor 1986; Triacca 2007; Giles 2008) uses the basic SV model to examine properties of the squared returns and absolute returns as an implicit estimator of the true unobserved volatility in financial markets.¹

In this article, we propose power transformation of absolute returns as a proxy of volatility. While we undertake a similar analysis of Giles (2008), we generalize his results to a case in which the power of absolute returns is allowed to vary freely. This study evaluates the accuracy and precision by using the standard metrics considered in the literature. Our empirical results show that our proposed estimator is the most efficient estimator compared to the unbiased estimator in terms of mean square error (hereafter MSE). This also infers that the new estimator improves proxies.

Commonly used volatility proxies are the absolute returns and the squared returns in the financial economics literature. Previous empirical studies found that the proxy of absolute returns has superior forecasting power than that of squared returns, see, for example, Ederington and Guan (2005) and Ghysels, Santa-Clara, and Valkanov (2006). Given this stylized empirical fact, we show in this study that power transformations of absolute returns are generally superior to the traditional use of absolute returns.

The layout of the article is as follows. In Section II we shed light on some properties of the proposed power transformation of absolute returns. Section III is devoted to empirical analysis. Concluding remarks are presented in Section IV.

¹Kalman filter proposed by Kalman (1960) through a state space representation is a common tool for the SV model. Moreover, empirical studies of Andersen, Benzoni, and Lund (2002) and Eraker, Johannes, and Polson (2003) improve model fit of the SV model by incorporating price jumps. Recently, Harvey (2013) proposes the dynamic conditional score model which is close to SV specification, where the volatility is treated as an unobserved stochastic process.© 2016 Informa UK Limited, trading as Taylor & Francis Group

CONTACT Hojin Jung 2hojin.jung@gmail.com School of Economics, Henan University, Kaifeng, Henan 475001, China

Supplemental data for this article can be accessed here.

http://dx.doi.org/10.1080/00036846.2016.1237751
II. A new measure of volatility in the SV model

Giles (2008) uses absolute returns as the proxy for unobserved volatility. We propose a power transformation method of absolute returns based on his analysis: $|r_t|^k$ as an estimator of $\sigma_t^k$, where $k > 0$, including two special values of $k$ of 1 and 2. Let $p_t$ be the spot price of a financial asset at time $t$ and its one-period return is defined as $r_t = \log_e \left( \frac{p_t}{p_{t-1}} \right)$, then the latent volatility is defined as $\sigma_t^2 = \text{var}(r_t|I_{t-1})$, where $I_{t-1}$ is the information set observed at time $t - 1$. In this study, we adapt a basic SV model (Taylor 2005, 278–83) and use Giles (2008) notation for comparability.

$$r_t = \mu + \sigma_t z_t, \quad z_t \sim iid \ N(0,1)$$

$$\log_e(\sigma_t^2) = y_0 + y_1 \log_e(\sigma_{t-1}^2) + u_t, \quad u_t \sim iid \ N(0, \sigma_u^2), \quad |y_1| < 1,$$

where $z_t$ and $u_t$ are independent processes.

The unconditional distribution of $\log_e(\sigma_t^2)$ is normal with mean $\frac{y_0}{1-y_1}$ and variance $\frac{\sigma_u^2}{1-y_1}$, given the stationarity condition in Equation (1). If $Y$ is a $N(m, \nu)$ random variable, then $X = \exp(Y)$ has a log-normal distribution with central moments given by

$$\mathbb{E}(X^k) = \exp \left( km + \frac{k^2 \nu}{2} \right).$$

The model given by Equations (1) and (2) can be adapted for our setting. Let $Y$ be equal to $\log_e(\sigma_t^2)$, then we have $X = \sigma_t^2$. Thus, Equation (2) can be easily rewritten as follows:

$$\mathbb{E}(\sigma_t^{2k}) = \exp \left( km + \frac{k^2 \nu}{2} \right).$$

By substituting the value of $\frac{k}{2}$, instead of $k$, in Equation (3), we obtain the following general form of the SV model:

$$\mathbb{E}(\sigma_t^k) = \exp \left( \frac{km}{2} + \frac{k^2 \nu}{8} \right),$$

where $k > 0$. This specification is the main novelty of this article. Using the properties of the integral representation of the gamma function, Giles (2008) shows that the absolute moments of a standard normal variate, $Z$, are

$$\mathbb{E}(|Z|^k) = \frac{2^{\frac{k}{2}} \Gamma \left( \frac{k+1}{2} \right)}{\sqrt{\pi} \Gamma \left( \frac{k}{2} \right)}; p > k.$$  

The results allow us to obtain the bias and MSE of $|r_t|^k$. Note that here consistent with the previous literature (see Giles 2008; Triacca 2007), we set $\mu = 0$.

**Case 1.** $|r_t|^k$ as an estimator of $\sigma_t^k$

We start by considering that $|r_t|^k$ is an estimator of $\sigma_t^k$. When $z_t$ is normally distributed, with Equation (5), the MSE is equal to

$$\text{MSE}(|r_t|^k) = \mathbb{E} \left[ (|r_t|^k - \sigma_t^k)^2 \right] = \mathbb{E} \left[ (|Z_t|^k \sigma_t^k - \sigma_t^k)^2 \right] = \mathbb{E} \left[ (|Z_t|^k - 1)^2 \sigma_t^{2k} \right]$$

$$= \mathbb{E} \left[ (|Z_t|^k - 1)^2 \right] \mathbb{E} \left[ \sigma_t^{2k} \right] = \mathbb{E} \left[ |Z_t|^{2k} - 2|Z_t|^k + 1 \right] \mathbb{E} \left[ \sigma_t^{2k} \right]$$

$$= \left[ 2^k \Gamma \left( \frac{k+1}{2} \right) - \frac{2^{\frac{k+2}{2}} \Gamma \left( \frac{k+1}{2} \right)}{\sqrt{\pi}} + 1 \right] \mathbb{E} \left[ \sigma_t^{2k} \right].$$

Equation (7)
We define the relative bias (RB) as the bias divided by its true value. The empirical RB is then given by

\[
RB(|r_t|^k) = \frac{\text{Bias}(|r_t|^k)}{\text{E}(\sigma_t^2)} = \frac{\text{E}(|r_t|^k - \sigma_t^2)}{\text{E}(\sigma_t^2)} = \text{E}(|Z_t|^k - 1) = \left[\frac{2^{\frac{k+1}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right)\right] - 1.
\]

We also derive the Student-\(t\) MSE of \(|r_t|^k\) by using Equation (6) when \(z_t\) is a standardized Student-\(t\) variate, that is,

\[
\text{MSE}(|r_t|^k) = \text{E}[(|r_t|^k - \sigma_t^2)^2] = \text{E}[(|T_t|^k \sigma_t^k - \sigma_t^2)^2] = \text{E}[(|T_t|^k - 1)^2 \sigma_t^{2k}]
\]

\[= \text{E}[(|T_t|^k - 1)^2] \text{E}[\sigma_t^{2k}] = \text{E}[T_t^{2k} - 2|T_t|^k + 1] \text{E}[\sigma_t^{2k}]
\]

\[= \left[\text{E}(|T_t|^{2k}) - 2 \text{E}(|T_t|^k) + 1\right] \text{E}[\sigma_t^{2k}]
\]

\[= \left[p^k \Gamma(k + \frac{1}{2}) \Gamma\left(\frac{k}{2} - k\right) - \frac{2p^k \Gamma\left(k + \frac{1}{2}\right) \Gamma\left(k - \frac{k}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{k}{2}\right)} + 1\right] \exp\left(\frac{ky_0}{1 - \gamma_1} + \frac{k^2 \sigma_u^2}{2(1 - \gamma_1^2)}\right).
\]

Case 2. \(|r_t|^k\) as an estimator of \(\sigma_t^2\)

Considering \(|r_t|^k\) as an estimator of \(\sigma_t^2\), our purpose now is to investigate the statistical properties of \(|r_t|^k\). In the same way as in Case 1, using Equation (5) and the normality assumption, we can derive the estimator errors as follows:

\[
RB(|r_t|^k) = \frac{\text{Bias}(|r_t|^k)}{\text{E}(\sigma_t^2)} = \frac{\text{E}(|r_t|^k - \sigma_t^2)}{\text{E}(\sigma_t^2)} = \left[\frac{2^{\frac{k}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right)\right] \exp\left(\frac{ky_0}{1 - \gamma_1} + \frac{k^2 \sigma_u^2}{2(1 - \gamma_1^2)}\right) - 1.
\]

We also write the RB of \(|r_t|^k\) as

\[
\text{MSE}(|r_t|^k) = \text{E}[(|r_t|^k - \sigma_t^2)^2] = \text{E}[(|Z_t|^k \sigma_t^k - \sigma_t^2)^2]
\]

\[= \text{E}(|Z_t|^{2k}) \text{E}[\sigma_t^{2k}] - 2 \text{E}(|Z_t|^k) \text{E}[\sigma_t^{2k}] + \text{E}[\sigma_t^4]
\]

\[= \left[\frac{2^{\frac{k}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right)\right] \exp\left(\frac{ky_0}{1 - \gamma_1} + \frac{k^2 \sigma_u^2}{2(1 - \gamma_1^2)}\right) + \exp\left(\frac{2y_0}{1 - \gamma_1} + \frac{2\sigma_u^2}{2(1 - \gamma_1^2)}\right)
\]

\[= \left[\frac{2^{\frac{k}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right)\right] \exp\left(\frac{ky_0}{2(1 - \gamma_1)} + \frac{k^2 \sigma_u^2}{8(1 - \gamma_1^2)}\right) \exp\left(\frac{y_0}{1 - \gamma_1} + \frac{\sigma_u^2}{2(1 - \gamma_1^2)}\right).
\]
The following equations are derived for the two metrics of the statistical estimator’s accuracy and precision under the Student-\( t \) distribution, using Equation (6).

\[
MSE(|r_i|^k) = \mathbb{E}\left[ (|r_i|^k - \sigma_i^2)^2 \right] = \mathbb{E}\left[ (|T_i|^k - \sigma_i^2)^2 \right] \\
= \mathbb{E}\left[ |T_i|^{2k} \right] \mathbb{E}\left[ \sigma_i^{2k} \right] - 2\mathbb{E}\left[ |T_i|^k \right] \mathbb{E}\left[ \sigma_i^{k+2} \right] + \mathbb{E}\left[ \sigma_i^4 \right] \\
= \left[ \frac{p^k \Gamma(k + \frac{1}{2}) \Gamma(\frac{p-k}{2})}{\sqrt{\pi} \Gamma(\frac{k}{2})} \right] \exp\left( \frac{ky_0}{1 - y_1} + \frac{k^2 \sigma_u^2}{2(1 - y_1^2)} \right) + \exp\left( \frac{2y_0}{1 - y_1} + \frac{2\sigma_u^2}{2(1 - y_1^2)} \right),
\]

\[
RB(|r_i|^k) = \frac{\text{Bias}(|r_i|^k)}{\text{Bias}(\sigma_i)} = \frac{\mathbb{E}(|r_i|^k - \sigma_i^2)}{\mathbb{E}(\sigma_i)} = \frac{\mathbb{E}(|r_i|^k)}{\mathbb{E}(\sigma_i)} - 1 \\
= \frac{\left[ \frac{p^k \Gamma(k + \frac{1}{2}) \Gamma(\frac{p-k}{2})}{\sqrt{\pi} \Gamma(\frac{k}{2})} \right] \exp\left( \frac{ky_0}{2(1 - y_1)} + \frac{k^2 \sigma_u^2}{8(1 - y_1^2)} \right)}{\exp\left( \frac{y_0}{1 - y_1} + \frac{\sigma_u^2}{2(1 - y_1^2)} \right)} - 1.
\]

**Case 3.** \( |r_i|^k \) as an estimator of \( \sigma_i \)

We turn to the last case, in which \( |r_i|^k \) is considered as an estimator of \( \sigma_i \). As we stated in the two previous cases, we obtain the MSE and RB for the different distributions of \( z_t \). Using Equation (5), the normal MSE and RB are also directly obtained:

\[
MSE(|r_i|^k) = \mathbb{E}\left[ (|r_i|^k - \sigma_i^2)^2 \right] = \mathbb{E}\left[ (|Z_i|^k - \sigma_i^2)^2 \right] \\
= \mathbb{E}\left[ |Z_i|^{2k} \right] \mathbb{E}\left[ \sigma_i^{2k} \right] - 2\mathbb{E}\left[ |Z_i|^k \right] \mathbb{E}\left[ \sigma_i^{k+1} \right] + \mathbb{E}\left[ \sigma_i^4 \right] \\
= \left[ \frac{2^k \Gamma(k + \frac{1}{2})}{\sqrt{\pi}} \right] \exp\left( \frac{ky_0}{1 - y_1} + \frac{k^2 \sigma_u^2}{2(1 - y_1^2)} \right) + \exp\left( \frac{y_0}{1 - y_1} + \frac{\sigma_u^2}{2(1 - y_1^2)} \right),
\]

\[
RB(|r_i|^k) = \frac{\text{Bias}(|r_i|^k)}{\text{Bias}(\sigma_i)} = \frac{\mathbb{E}(|r_i|^k - \sigma_i^2)}{\mathbb{E}(\sigma_i)} = \frac{\mathbb{E}(|r_i|^k)}{\mathbb{E}(\sigma_i)} - 1 \\
= \frac{\left[ \frac{2^k \Gamma(k + \frac{1}{2})}{\sqrt{\pi}} \right] \exp\left( \frac{ky_0}{2(1 - y_1)} + \frac{k^2 \sigma_u^2}{8(1 - y_1^2)} \right)}{\exp\left( \frac{y_0}{2(1 - y_1)} + \frac{\sigma_u^2}{8(1 - y_1^2)} \right)} - 1.
\]
\[ \text{MSE}(|r_i|^k) = \mathbb{E} \left[ (|r_i|^k - \sigma_i)^2 \right] = \mathbb{E} \left[ (|r_i|^k^2 \sigma_i^2 - \sigma_i)^2 \right] \]
\[ = \mathbb{E} \left[ T_i^2 \sigma_i^2 \right] \mathbb{E} \left[ \sigma_i^2 \right] - 2 \mathbb{E} \left[ T_i^k \right] \mathbb{E} \left[ \sigma_i^{k+1} \right] + \mathbb{E} \left[ \sigma_i^2 \right] \]
\[ = \frac{\beta k^k \Gamma(k + \frac{1}{2}) \Gamma(k - \frac{1}{2})}{\sqrt{\pi} \Gamma(k)} \exp \left( \frac{ ky_0}{2(1 - \gamma)} + \frac{k^2 \sigma_i^2}{2(1 - \gamma)} \right) + \exp \left( \frac{ y_0}{1 - \gamma} + \frac{\sigma_i^2}{2(1 - \gamma)} \right) \]
\[ - \frac{2 \beta^2 \Gamma(k + \frac{1}{2}) \Gamma(k - \frac{1}{2})}{\sqrt{\pi} \Gamma(k)} \exp \left( \frac{ ky_0}{2(1 - \gamma)} + \frac{k^2 \sigma_i^2}{8(1 - \gamma)} \right) \exp \left( \frac{ y_0}{2(1 - \gamma)} + \frac{\sigma_i^2}{8(1 - \gamma)} \right), \]  

\[ \text{RB}(|r_i|^k) = \frac{\text{Bias}(|r_i|^k)}{\mathbb{E}(\sigma_i)} = \frac{\mathbb{E}(|r_i|^k - \sigma_i)}{\mathbb{E}(\sigma_i)} - 1 \]
\[ = \frac{\frac{k^k \Gamma(k + \frac{1}{2}) \Gamma(k - \frac{1}{2})}{\sqrt{\pi} \Gamma(k)} \exp \left( \frac{ ky_0}{2(1 - \gamma)} + \frac{k^2 \sigma_i^2}{8(1 - \gamma)} \right)}{\exp \left( \frac{ y_0}{2(1 - \gamma)} + \frac{\sigma_i^2}{8(1 - \gamma)} \right)} - 1. \]  

III. Empirical analysis of a new proxy measure

In order to assess the efficiency of our proposed proxy along several dimensions, we utilize simulated data. In the simulation design, the datasets are generated from the parameter ranges deduced from the evidence compiled by Taylor (2005, 287–8), such as \( y_0 \in (0.95, 0.99), y_1 \in (-0.54, -0.1) \) and \( \sigma_u^2 \in (0.0018, 0.0478) \). Given the typical values of the parameters, there are many possible combinations to construct datasets. In this study, we report empirical results from only the case in which we use the most conservative set of values for the parameters such as \( k \in \{0.3, 0.8, 1, 1.3, 1.8, 2\} \), \( y_0 \in \{0.95, 0.97, 0.99\} \), and \( \gamma_1 \in \{-0.54, -0.25, -0.1\} \) and \( \sigma_u^2 \in \{0.0018, 0.025, 0.0478\} \).

We report the normal RB and MSE of Case 1 by using Equations (7) and (8). Table 1 shows that the RB is as low as \(-0.065\) when \( k = 1.8 \), and in

| Table 1. RB, MSE and RE of \( |r_i|^k \) in Case 1 (normal distribution). |
|-------------------------------------------------|
| Relative bias | \( (y_0 = 0.95, y_1 = -0.54, \sigma_u^2 = 0.0018) \) | \( (y_0 = 0.97, y_1 = -0.25, \sigma_u^2 = 0.025) \) | \( (y_0 = 0.99, y_1 = -0.1, \sigma_u^2 = 0.0478) \) | Monte Carlo DGP |
|---|---|---|---|
| \( k \) | \( \text{RB} \) | \( \text{RB} \) | \( \text{RB} \) | \( \text{RB} \) |
| 0.3 | -0.1331 | -0.1331 | -0.1331 | -0.1331 |
| 0.8 | -0.2045 | -0.2045 | -0.2045 | -0.2045 |
| 1 | -0.2021 | -0.2021 | -0.2021 | -0.2021 |
| 1.3 | -0.1740 | -0.1740 | -0.1740 | -0.1740 |
| 1.8 | -0.0659 | -0.0659 | -0.0659 | -0.0659 |
| 2 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Mean square error | 0.0901 | 0.0945 | 0.0982 | 0.0936 |
| 0.3 | 0.4760 | 0.5416 | 0.5991 | 0.5287 |
| 1 | 0.7493 | 0.8812 | 1.0003 | 0.8555 |
| 1.3 | 1.4319 | 1.7700 | 2.0891 | 1.7053 |
| 1.8 | 4.3268 | 5.8182 | 7.3372 | 5.5414 |
| 2 | 7.1040 | 9.6656 | 12.3949 | 9.0753 |
| % of RE \( (\gamma_1) \) | 164.2% | 166.47% | 168.9% | 163.77% |

A power term \( (k) \) is raised to construct the series of \( |r_i|^k \) for an estimator of \( \sigma_u^2 \). In particular, the underlying data is generated by a combination of three parameters \( (y_0, y_1, \text{and } \sigma_u^2) \), which designate to increase their values across the specifications. The calculated RB and MSE are obtained by using Equations (7) and (8) under a normal distribution assumption. A value is highlighted in the upper panel if RB is less than the critical value of 10% and for its corresponding MSE in the lower panel.

\footnote{Note that Giles (2008) analysis is a special case in our model when we assume that \( k = 1 \).}

\footnote{The use of these ranges of the parameters is valid for the following reasons: (1) As discussed in Taylor (2005), the estimates of these parameters given by previous empirical studies are usually between the ranges. (2) This study chooses the same parameter ranges used in Giles (2008) for comparison with his study.}

\footnote{For simplicity, \( k \) takes such values, although in principle \( k \) can be extended to accommodate any values.}

\footnote{We find that the results shown in this study are not sensitive to other datasets constructed by different parameter values. The complete datasets used in this study and all of our empirical results are available in the Supplemental data.}
particular, they are zero across the columns when $k = 2$. If RB is less than 10%, then it is negligible in Cochran’s (1963) sense. Now we have one unbiased estimator and another biased, but acceptable, estimator. Then, it is reasonable to compare variance or relative efficiency (hereafter RE) of two estimators. The normal MSE and RE of $|r_t|^k$ are presented at the bottom of Table 1. The empirical RE is defined as the ratio of two different MSEs. When $k$ changes from 2 to 1.8, the empirical RE in percentage terms is much greater than 164% across specifications. Our proposed estimator with an acceptable level of bias could lead to a systematically significant increase in efficiency. This empirical result strongly supports that under normality, the negligible biased estimator of $|r_t|^k$ when $k = 1.8$ is more efficient than the unbiased estimator of $|r_t|^k$ when $k = 2$.

We also conduct a simulation exercise. The RB and MSE statistics in the last column are computed via stochastic simulations in a Monte Carlo setting with 1000 replications, where artificial series for returns are simulated according to the proposed data-generating process (DGP). The figures in the column are the average values of the statistics obtained by implementing the simulation. We find that the empirical results are consistent with the theoretical values delivered by Equations (7) and (8).

A closer inspection of how our proposed estimator performs in a fat tailed returns distribution is investigated in a further simulation. For that purpose, different degrees of freedom are induced, such as $p \in \{5, 7, 10\}$. Table 2 provides the RB and MSE of the new estimator $|r_t|^k$ in the Student-$t$ case by using Equations (9) and (10). Keeping degree of freedom $p$ as constant when power $k$ increases, bias shows a tendency of becoming negative from $k = 0.3$ to $k = 1.0$ and then positive from $k = 1.3$ to $k = 2$ across all specifications. This finding implies that the bias is close to zero around $k = 1.3$. The RB of $|r_t|^k$ is less than 10% when $k = 1.3$. It is interesting to note here that the MSE is much smaller when $k = 0.3$ than when $k$ takes other values. When we choose $|r_t|^{0.3}$ as an estimator instead of $|r_t|^{1.3}$, it provides significant improvement in precision, which is desirable, even at the expense of losing only a little accuracy. In accordance with our previous results, the proposed estimator $|r_t|^k$ is a better volatility proxy even when $z_t$ is a standardized Student-$t$ variate.

Note that the upper panel in Table 1 reports identical figures because the RB depends on the power ($k$) of absolute returns as shown in Equation (8). For a similar reason, we observe the same RB across the degrees of freedom ($p$) at ($k$) in Table 2. $k$ and $p$ are critical factors that determine the RBs in a fat tailed returns distribution, as shown in Equation (10). We also find from our simulation exercise that the empirical results obtained for the standardized Student-$t$ converge towards the results obtained for the standardized normal processes as the degrees of freedom ($p$) increase. For example, when $p = 100$, the estimated RBs are $-0.1316$, $-0.2000$, $-0.1961$, $-0.1650$, $-0.0496$ and $0.0204$, respectively. The complete simulation results in this study can be available upon request.

Figure 1 illustrates that holding $\gamma_0$, $\gamma_1$ and $\sigma_u^2$ constant, as $k$ increases, MSE also increases when $z_t$ is a standardized Student-$t$ variate. An interesting feature of the figure is that the RB and MSE are significantly larger when $p = 5$, which means that fatter tails in the returns distribution cause the estimator to become more biased and imprecise.

Since we have already discussed in detail the application of the proposed estimator to empirical analysis in Case 1, we now focus directly on the empirical results in Cases 2 and 3. Tables 3 and 5 report the RB and MSE obtained from Case 2 under different types of returns distributions. Under the normal specification, all RBs for any given value of $k$ are greater than 10%, except when $k = 2$. Note that Triacca (2007) considers $|r_t|^2$ as an unbiased estimator of $\sigma_t^2$ and our empirical findings support Triacca’s result. On the other hand, under the Student-$t$ specification, the sign and magnitude of our proposed estimator’s RB depend on the parameter values. Our empirical results also indicate that $|r_t|^2$ becomes a biased estimator. A comparison of the MSEs of $|r_t|^2$ and the others shows that it performs worse as a proxy of volatility in this more general Student-$t$ case. However, the RBs of $|r_t|^k$ are lower than 10% when $k = 1.8$ and $p = 10$, suggesting that $|r_t|^k$ is a better estimator than the unbiased estimator considered in Triacca (2007).

---

*We also find that for the typical parameter value ranges, as degrees of freedom increase the MSE of $|r_t|^k$ monotonically decreases.*
Table 2. RB and MSE of $|r_t|^k$ in Case 1 (t-distribution).

<table>
<thead>
<tr>
<th>$k$</th>
<th>$(y_0 = 0.95, y_1 = -0.54, \sigma^2_u = 0.0018)$</th>
<th>$(y_0 = 0.97, y_1 = -0.25, \sigma^2_u = 0.025)$</th>
<th>$(y_0 = 0.99, y_1 = -0.1, \sigma^2_u = 0.0478)$</th>
<th>(Monte Carlo DGP)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p = 5$</td>
<td>$p = 7$</td>
<td>$p = 10$</td>
<td>$p = 5$</td>
</tr>
<tr>
<td>0.3</td>
<td>-0.0999</td>
<td>-0.1101</td>
<td>-0.1173</td>
<td>-0.0999</td>
</tr>
<tr>
<td>0.8</td>
<td>-0.0965</td>
<td>-0.1317</td>
<td>-0.1556</td>
<td>-0.0965</td>
</tr>
<tr>
<td>1</td>
<td>-0.0510</td>
<td>-0.1017</td>
<td>-0.1353</td>
<td>-0.0510</td>
</tr>
<tr>
<td>1.3</td>
<td>0.0659</td>
<td>-0.0184</td>
<td>-0.0721</td>
<td>0.0659</td>
</tr>
<tr>
<td>1.8</td>
<td>0.4322</td>
<td>0.2408</td>
<td>0.1287</td>
<td>0.4322</td>
</tr>
<tr>
<td>2</td>
<td>0.6667</td>
<td>0.4000</td>
<td>0.2500</td>
<td>0.6667</td>
</tr>
</tbody>
</table>

Mean square error

<table>
<thead>
<tr>
<th>$k$</th>
<th>$(y_0 = 0.95, y_1 = -0.54, \sigma^2_u = 0.0018)$</th>
<th>$(y_0 = 0.97, y_1 = -0.25, \sigma^2_u = 0.025)$</th>
<th>$(y_0 = 0.99, y_1 = -0.1, \sigma^2_u = 0.0478)$</th>
<th>(Monte Carlo DGP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.0987</td>
<td>0.0951</td>
<td>0.0931</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>0.7331</td>
<td>0.6225</td>
<td>0.5636</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.4262</td>
<td>1.1196</td>
<td>0.9660</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>4.0429</td>
<td>2.7331</td>
<td>2.1590</td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>29.8159</td>
<td>13.6978</td>
<td>8.9531</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>78.2371</td>
<td>27.6131</td>
<td>16.3953</td>
<td></td>
</tr>
</tbody>
</table>

A power term ($k$) is raised to construct the series of $|r_t|^k$ for an estimator of $\sigma_u$. In particular, the underlying data is generated by a combination of four parameters ($y_0$, $y_1$, $\sigma^2_u$, and $p$), which designate to increase their values across the specifications. The calculated RB and MSE are obtained by using Equations (9) and (10) under a Student-t distribution assumption. A value is highlighted in the upper panel if RB is less than the critical value of 10% and for its corresponding MSE in the lower panel.
Lastly, we also present empirical evidence regarding the performance of $|r_t|^k$ as a predictor of $\sigma_t$. Table 4 shows that the RB in absolute value is less than 10% when $k = 1.3$ with the normal distribution in Case 3. Note that $|r_t|$ is a biased estimator of $\sigma_t$ when $z_t$ is normally distributed. This is consistent with the findings in Giles (2008). With this particular parameter, $k = 1.3$, the proposed proxy of volatility exhibits good performance for a standardized Student-t distribution with 10 degrees of freedom. Its RB becomes less than 10% when considering the Student-t distribution with large degrees of freedom. This provides additional support that $|r_t|^k$ is a good proxy for measuring volatility. Overall, the accuracy and precision of $|r_t|^k$ as an estimator of volatility are remarkable when we compare RB and MSE across cases. When estimating volatility, $|r_t|^k$ has smaller MSE than the absolute return or the squared return often used in literature, at least for the typical parameter settings. McKenzie (1999) found that a power term of 1.25 is optimal based on the standard root MSE when modelling the exchange rate volatility.

Further insights into the differences across cases can be gained by comparing figures. Figures 2 and 3 display the RB and MSE obtained from Cases 2 and 3, respectively. They are consistent with the pattern we

---

Table 3. RB and MSE of $|r_t|^k$ in Case 2 (normal distribution).

<table>
<thead>
<tr>
<th>$k$</th>
<th>$(\gamma_0 = 0.97, \gamma_1 = -0.25, \sigma_0^2 = 0.025)$</th>
<th>$(\gamma_0 = 0.99, \gamma_1 = -0.1, \sigma_0^2 = 0.0478)$</th>
<th>Monte Carlo DGP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.8958</td>
<td>1.7112</td>
<td>2.7218</td>
</tr>
<tr>
<td>0.8</td>
<td>1.1177</td>
<td>1.8449</td>
<td>2.7338</td>
</tr>
<tr>
<td>1</td>
<td>1.2751</td>
<td>1.9848</td>
<td>2.8385</td>
</tr>
<tr>
<td>1.3</td>
<td>1.7631</td>
<td>2.5505</td>
<td>3.4568</td>
</tr>
<tr>
<td>1.8</td>
<td>4.3924</td>
<td>6.2933</td>
<td>8.3690</td>
</tr>
<tr>
<td>2</td>
<td>6.9208</td>
<td>10.2211</td>
<td>13.9541</td>
</tr>
</tbody>
</table>

A power term ($k$) is raised to construct the series of $|r_t|^k$ for an estimator of $\sigma_t^2$. In particular, the underlying data is generated by a combination of three parameters ($\gamma_0$, $\gamma_1$, and $\sigma_0^2$), which designate to increase their values across the specifications. The calculated RB and MSE are obtained by using Equations (11) and (12) under a normal distribution assumption. A value is highlighted in the upper panel if RB is less than the critical value of 10% and for its corresponding MSE in the lower panel.
Table 4. RB and MSE of $|r_t|^k$ in Case 3 (normal distribution).

<table>
<thead>
<tr>
<th>$k$</th>
<th>$(\gamma_0 = 0.95, y_1 = -0.54, \sigma_u^2 = 0.0018)$</th>
<th>$(\gamma_0 = 0.97, y_1 = -0.25, \sigma_u^2 = 0.025)$</th>
<th>$(\gamma_0 = 0.99, y_1 = -0.1, \sigma_u^2 = 0.0478)$</th>
<th>Monte Carlo DGP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>-0.3017</td>
<td>-0.3413</td>
<td>-0.3708</td>
<td>-0.3336</td>
</tr>
<tr>
<td>0.8</td>
<td>-0.2521</td>
<td>-0.2647</td>
<td>-0.2745</td>
<td>-0.2624</td>
</tr>
<tr>
<td>1</td>
<td>-0.2021</td>
<td>-0.2021</td>
<td>-0.2021</td>
<td>-0.2021</td>
</tr>
<tr>
<td>1.3</td>
<td>-0.0937</td>
<td>-0.0699</td>
<td>-0.0506</td>
<td>-0.0742</td>
</tr>
<tr>
<td>1.8</td>
<td>0.1964</td>
<td>0.2837</td>
<td>0.3571</td>
<td>0.2682</td>
</tr>
<tr>
<td>2</td>
<td>0.3626</td>
<td>0.4888</td>
<td>0.5970</td>
<td>0.4668</td>
</tr>
</tbody>
</table>

Mean square error

- 0.3: 0.2388
- 0.8: 0.5272
- 1: 0.7519
- 1.3: 1.3857
- 1.8: 4.4051
- 2: 7.1570

A power term ($k$) is raised to construct the series of $|r_t|^k$ for an estimator of $\sigma$. In particular, the underlying data is generated by a combination of three parameters ($\gamma_0$, $y_1$, and $\sigma_u^2$), which designate to increase their values across the specifications. The calculated RB and MSE are obtained by using Equations (15) and (16) under a normal distribution assumption. A value is highlighted in the upper panel if RB is less than the critical value of 10% and for its corresponding MSE in the lower panel.
observed in Figure 1. Given Figures 1–3, it is clear that the bias and MSE in Case 1 are relatively smaller than those of the other cases. This provides strong evidence that Case 1, $|r|^k$ as an estimator of $\sigma^k$, is more efficient than Case 2, $|r|^k$ as an estimator of $\sigma^2$, and Case 3, $|r|^k$ as an estimator of $\sigma$.

IV. Conclusion

This article proposes power transformation of absolute returns as a proxy for the unobservable volatility. The RB and MSE serve as yardsticks for measuring the performance of such a proxy in the various parameter conditions. We provide empirical evidence that the new proposed estimator outperforms two popular proxies such as the absolute and squared returns in terms of efficiency. The power transformation of absolute returns is a biased estimator with negligible bias and small variance. Moreover, it exhibits better fits in a heavy tailed returns distribution which is typical for data from financial markets.  

Disclosure statement

No potential conflict of interest was reported by the authors.

References


Note: Additional information for this article can be found in the Supplementary data at http://facultypages.morris.umn.edu/~jongmink/research/OnlineAppendix.pdf