A Realized Stochastic Volatility Model With Box–Cox Transformation

Tingguo ZHENG
Wang Yanan Institute for Studies in Economics, Xiamen University, Fujian 361005, China (*zhengtg@gmail.com*)

Tao SONG
Institute of Economics, School of Economics, Xiamen University, Fujian 361005, China (*songtao@xmu.edu.cn*)

This article presents a new class of realized stochastic volatility model based on realized volatilities and returns jointly. We generalize the traditionally used logarithm transformation of realized volatility to the Box–Cox transformation, a more flexible parametric family of transformations. A two-step maximum likelihood estimation procedure is introduced to estimate this model on the basis of Koopman and Scharth (2013). Simulation results show that the two-step estimator performs well, and the misspecified log transformation may lead to inaccurate parameter estimation and certain excessive skewness and kurtosis. Finally, an empirical investigation on realized volatility measures and daily returns is carried out for several stock indices.

KEY WORDS: Model misspecification; Realized volatility; State space model; Stock return; Two-step estimation.

1. INTRODUCTION

In financial economics and mathematical finance, stochastic volatility (SV) is one of the main concepts used to deal with time-varying volatility and codependence found in financial markets. As an alternative to GARCH class of models, SV class of models has received substantial attention. Many SV models have been used by financial econometricians to model time-varying volatilities for financial time series, such as stock prices, exchange rates, and interest rates (see Taylor 1986b; Harvey, Ruiz, and Shephard 1994; Jacquier, Polson, and Rossi 1994; Kim, Shephard, and Chib 1998; Chib, Nardari, and Shephard 2002, and others).

Most stochastic volatility models are based on financial asset returns and their conditional volatility processes typically contain variables based on squared or absolute returns. However, as pointed out by Alizadeh, Brandt, and Diebold (2002), Chou (2005), Brandt and Diebold (2006), and other authors, the returns-based SV models are inaccurate and inefficient, because they are based on the closing prices of the reference period, and thus fail to use the information contents inside the reference period. In addition, the estimation of SV models has proved quite difficult. It is well known that simple estimators such as the moment-based methods and the quasi-maximum likelihood (QML) approach are highly inefficient, and most other simulation-based estimators are computationally intensive (see Broto and Ruiz 2004 for a survey and performance comparison of some of these estimation techniques).

Recently, realized volatility or realized variance, introduced by Andersen and Bollerslev (1998), Andersen et al. (2001), and Barndorff-Nielsen and Shephard (2001), has been used for modeling stochastic volatility. One of the commonly used terms is the sum of squared high-frequency returns over a certain interval such as a day, and provides a consistent estimator of the latent volatility under an ideal market assumption. See Andersen, Bollerslev, and Diebold (2010) for a survey of realized volatility. When high-frequency data are available, realized volatility seems to be an appropriate measure of volatility. Barndorff-Nielsen and Shephard (2002) studied the use of realized volatility in estimating stochastic volatility, and showed that model-based methods may be particularly helpful in estimating historical records of actual volatility. However, there are two severe issues in measuring daily realized volatility from high frequency return data: the presence of nontrading hours and the presence of the market microstructure noise in transaction prices, which may lead to downward bias and upward bias in estimating the latent volatility, respectively (see Hansen and Lunde 2005 and Hasbrouck 2007 for details, and McAleer and Medeiros 2008 for a review of the realized volatility and effects of the microstructure noise). From this point of view, some studies such as those by Takahashi, Omori, and Watanabe (2009) and Koopman and Scharth (2013) model realized volatility and daily returns simultaneously based on the well-known (log-normal) stochastic volatility model, assuming that the realized volatility includes the microstructure noise but still contains a great deal of information on the latent volatility. On the other hand, daily returns contain less noise but do not include the sufficient information on the latent volatility. They showed that this model can estimate realized volatility biases and parameters simultaneously.

For modeling the positive-valued random variables of realized volatility, the use of transformations is very common and may be helpful when the original model does not satisfy the usual assumption. A natural choice for transforming the realized volatility data is the logarithm transformation. It was shown by Barndorff-Nielsen and Shephard (2005) that the finite sample distribution of log transformation of realized volatility is closer

Journal of Business & Economic Statistics
October 2014, Vol. 32, No. 4
DOI: 10.1080/07350015.2014.918544
Color versions of one or more of the figures in the article can be found online at *www.tandfonline.com/r/jbes*.

to the asymptotic standard normal distribution than that of a nontransformed version of realized volatility. Most empirical studies focus on the log transformation in volatility estimation and forecasting, for example, Andersen et al. (2001), Takahashi, Omori, and Watanabe (2009), and Koopman and Scharth (2013). However, a more flexible power transformation should be the Box–Cox transformation. This transformation controlled by the transformation parameter contains the logarithm version and the raw version as special cases. Moreover, a Box–Cox transformation for realized volatility may be closer to normal than the logarithm transformation, given an appropriate transformation parameter. The simulation evidence presented by Gonçalves and Meddahi (2009) shows that log transformation does not completely eliminate skewness in finite samples, and Gonçalves and Meddahi (2011) further suggested an appropriate Box–Cox transformation to reduce the finite sample skewness more effectively.

In this article, we aim to generalize the realized stochastic volatility model to a more general version with the Box–Cox transformation, hereafter referred to as the realized BCSV model. Although the Box–Cox transformation has been used for the returns-based SV model in past studies such as Yu, Yang, and Zhang (2006) and Zhang and King (2008), our extension of the model differs significantly from theirs, as we introduce multiple realized volatility measures to help in estimating the stochastic volatility. In addition, although Gonçalves and Meddahi (2011) also suggested the Box–Cox transformation for realized volatility and studied the accuracy based on the theory of Edgeworth expansions, they did not develop any econometric method to analyze this class of nonlinear transformation for volatility modeling and forecasting.

To estimate the realized BCSV model, we present a two-step maximum likelihood estimation procedure based on the study of Koopman and Scharth (2013). Due to the additional complexity of having the Box–Cox transformation in the model, it is not easy to estimate the realized BCSV model directly. However, for the joint analysis of returns and realized measures, the maximum likelihood estimation can be divided into two parts: for one part we put the realized measures into a transformed state space form and its likelihood function can be carried out via Kalman filter and smoother; and for the other part we implement maximum likelihood estimation by evaluating the expectation of the product of return densities conditional on realized measures. This two-step estimator is consistent (shown by Koopman and Scharth 2013) and the estimation procedure is much simpler than the Bayesian estimation method using Markov chain Monte Carlo (MCMC) technique developed by Takahashi, Omori, and Watanabe (2009).

The rest of this article is organized as follows. Section 2 contains detailed description of the realized BCSV model considered. In Section 3, we present the two-step estimator for the proposed model. The Monte Carlo simulation experiment is conducted in Section 4 to evaluate the finite sample performance of the two-step estimator and the impacts of model misspecification on parameter estimation and excessive skewness and kurtosis. Empirical applications to stock returns and realized measures of several stock indices are presented in Section 5. The final section presents conclusions.

2. THE REALIZED BCSV MODEL

2.1 Realized Volatility and Transformation

The realized volatility is an alternative volatility proxy that has gained much attention recently; see Andersen et al. (2001, 2003), and Barndorff-Nielsen and Shephard (2002, 2004). It is now used as volatility estimate calculated as the sum of intraday squared returns at short intervals.

Let p_t be the daily log price of a financial asset at day t and n be the sampling frequency within each period $[t-1, t]$, there are n continuous returns between $t-1$ and t,

$$r_{t,i} = p_{t-1+i/n} - p_{t-1+(i-1)/n},$$

for $i = 1, \ldots, n$. Then the realized volatility

$$\mathrm{RV}_t = \sum_{i=1}^{n} r_{t,i}^2, \tag{1}$$

is an estimate of the integrated volatility $\int_{t-1}^{t} \sigma^2(t)dt$ (Barndorff-Nielsen and Shephard 2002).

In the ideal situation that there were no market microstructure noise and the asset were always and continuously traded, the realized volatility would provide an asymptotically consistent estimator of the integrated volatility. For the real-world case, different estimates of realized volatility are available, such as the realized kernel estimator of Barndorff-Nielsen et al. (2008), the preaveraging-based realized variance estimator of Jacod et al. (2009), the subsampled realized variance estimator of Zhang, Mykland, and Aït-Sahalia (2005), and the subsampled median-based realized variance estimator of Andersen, Dobrev, and Schaumburg (2012). These different estimation methods provide us with multiple observable realized volatility measures (or referred to as realized measures).

The Box–Cox transformation is a useful family of power transformations proposed by Box and Cox (1964). Let

$$\mathrm{RV}_t^{(\lambda)} = \begin{cases} \left(\mathrm{RV}_t^{\lambda} - 1\right)/\lambda, & \text{if } \lambda \neq 0; \\ \log \mathrm{RV}_t, & \text{if } \lambda = 0. \end{cases} \tag{2}$$

It contains the log transformation (when $\lambda = 0$) and the raw statistic (when $\lambda = 1$) as special cases. Frequently, this Box–Cox transformation is used to transform the underlying variable to a normally distributed variable. Of course, in practice not all data could be power-transformed to normal.

Barndorff-Nielsen and Shephard (2002) argued that the central limit theorem approximation for the log of the standard realized variance estimator has a good finite sample performance in practical settings, making the log transformation a natural choice. However, Gonçalves and Meddahi (2011) showed on the basis of a Monte Carlo simulation study that specific Box–Cox transformation improves the accuracy of asymptotic approximations for realized estimators, and hence performs better than other transformation such as the raw (when $\lambda = 1$) and the log transformation (when $\lambda = 0$) in eliminating the bias and the skewness in finite samples.

In this article, we wish to determine an appropriate λ so that $\mathrm{RV}_t^{(\lambda)}$ is more approximated to a normal and has less skewness than the logarithm transformation.

2.2 Stochastic Volatility Dynamics for Transformed Realized Measures

Suppose we have K different realized volatility measures for a certain financial asset. For notational convenience we let $x_t = (\mathrm{RV}_{1t}, \ldots, \mathrm{RV}_{Kt})'$, $t = 1, \ldots, T$, where RV_{jt} is the jth realized measure at day t. Let $x_t^{(\lambda)} = (\mathrm{RV}_{1t}^{(\lambda_1)}, \ldots, \mathrm{RV}_{Kt}^{(\lambda_K)})'$ be the Box–Cox transformation of x_t, where λ_j is the transformation parameter for the jth realized measure ($j = 1, \ldots, K$).

We model the transformed realized measures as a static factor model with a common latent volatility process. The measurement equation is then given by

$$\mathrm{RV}_{j,t}^{(\lambda_j)} = \tau_j + \beta_j \theta_t + u_{j,t}, \quad u_t = (u_{1,t}, \ldots, u_{K,t})' \sim N(0, \Sigma_u), \tag{3}$$

for $j = 1, \ldots, K$, and $t = 1, \ldots, T$, where τ_j and β_j are unknown parameters. In (3), θ_t can be considered as the unobserved underlying volatility while $\mathrm{RV}_{j,t}^{(\lambda_j)}$ are all biased and noisy observations of θ_t. If $\tau_j \neq 0$ and $\beta_j \neq 1$, $j = 1, \ldots, K$, we implicitly take the transformed realized volatility as possibly biased estimates of daily volatility. In practice, it often restricts $\beta_j = 1$ for $j = 1, \ldots, K$ for convenience; see Takahashi, Omori, and Watanabe (2009), and Koopman and Scharth (2013). The bias may come from several sources such as nonlinear transformation of RV_t because of Jensen's inequality, the market microstructure noise and nontrading hours. In addition, the measurement disturbances u_t is assumed to be an independence sequence of normally distributed random vectors with zero mean and covariance matrix Σ_u. We decompose the covariance matrix to a $K \times 1$ vector of standard deviations, $\sigma_u = (\sigma_{u1}, \ldots, \sigma_{uK})'$, and a $K \times K$ correlation matrix, $R = \{\rho_{u,ij}\}_{i,j=1}^{K}$ with $\rho_{u,ij} = \rho_{u,ji}$ and $\rho_{u,ii} = 1$ for $i, j = 1, \ldots K$, according to the decomposition: $\Sigma_u = \mathrm{diag}\{\sigma_{u1}, \ldots, \sigma_{uK}\} \cdot R \cdot \mathrm{diag}\{\sigma_{u1}, \ldots, \sigma_{uK}\}$, where diag() is a function that inserts a vector into the diagonal of a matrix.

For the latent volatility θ_t in (3), we assume that it consists of multiple dynamic volatility process. Let θ_t a sum of $m \times 1$ vector containing the components $h_{1t}, \ldots, h_{mt}$, the volatility process is then defined as

$$\theta_t = \mu + \sum_{i=1}^{m} h_{it}, \quad h_{i,t+1} = \phi_i h_{i,t} + \eta_{it}, \quad \eta_{it} \sim N\big(0, \sigma_{\eta i}^2\big), \tag{4}$$

with the initial value $h_{i1} \sim N(0, \sigma_{\eta i}^2/(1 - \phi_i^2))$ known, where η_t represents a sequence of independent and identically normally distributed random variables and is independent of the measurement errors u_t. The constant term μ denotes the volatility mean, and the autoregressive coefficients $\{\phi_1, \ldots, \phi_m\}$ denote the volatility persistence for each component. To ensure stationarity and noninterchangeablility for h_{it}, we impose the constraints $|\phi_i| < 1$, $i = 1, \ldots, m$, and $\phi_1 > \cdots > \phi_m$. This specification can capture both short and long run components for volatility.

The model given by (3) and (4) can be easily rewritten in a conventional state space form as follows:

$$x_t^{(\lambda)} = c + Z\alpha_t + G\epsilon_t, \tag{5}$$

$$\alpha_{t+1} = d + T\alpha_t + H\epsilon_t, \tag{6}$$

for $t = 1, \ldots, T$, where $\epsilon_t \sim N(0, I_{K+m})$ with I_{K+m} being a $(K+m) \times (K+m)$ identity matrix, $x_t^{(\lambda)} = (\mathrm{RV}_{1,t}^{(\lambda_1)}, \ldots, \mathrm{RV}_{K,t}^{(\lambda_K)})'$, $\alpha_t = (h_{1t}, \ldots, h_{mt})'$, $GG' = \Sigma_u$, $c = (\tau_1 + \beta_1\mu, \ldots, \tau_K + \beta_K\mu)'$, $d = \mathbf{0}_{m\times 1}$, $HH' = \mathrm{diag}\{\sigma_{\eta 1}^2, \ldots, \sigma_{\eta,m}^2\}$,

$$Z = \begin{pmatrix} \beta_1 & \beta_1 & \cdots & \beta_1 \\ \beta_2 & \beta_2 & & \beta_2 \\ \vdots & & \ddots & \vdots \\ \beta_K & \beta_K & \cdots & \beta_K \end{pmatrix}, \quad \text{and}$$

$$T = \begin{pmatrix} \phi_1 & 0 & \cdots & 0 \\ 0 & \phi_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \phi_m \end{pmatrix}.$$

In this state space form, the latent volatility can be calculated by $\theta_t = \mu + \mathbf{1}'\alpha_t$, where $\mathbf{1}$ is a column vector of one. Since the measurement disturbances u_t and the transition disturbances η_t are assumed to be mutually independent, we have $GH' = \mathbf{0}$ and $HG' = \mathbf{0}$. In addition, the initial state vector $\alpha_1 \sim N(a_1, P_1)$ is also known, with a_1 is the vector of zero and the variance matrix $P_1 = \mathrm{diag}(\sigma_{\eta 1}^2/(1 - \phi_1^2), \ldots, \sigma_{\eta m}^2/(1 - \phi_m^2))$.

Unfortunately, the model given by (3) and (4) for the realized measures does not enable us to identify all parameters, since the bias parameters and the volatility mean, that is, τ_j, β_j, and μ, $j = 1, \ldots, N$, are not identified. Hence following Takahashi, Omori, and Watanabe (2009) and Koopman and Scharth (2013), we will include daily return in our model to ensure identifiability of all parameters in next section.

2.3 Simultaneous Modeling of Daily Returns and Realized Volatilities

Let p_t be the log closing price of an asset at day t and y_t be the daily continuously compounded return, defined as $y_t = p_t - p_{t-1}$. We assume

$$y_t = \sigma_t \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{iid}N(0, 1), \tag{7}$$

where σ_t is the daily volatility, and ε_t is an independent and identically distributed normal random variable with zero mean and unit variance. To simplify the exposition, we assume a zero mean for the returns process throughout this article. Our model can be easily extended to include nonzero mean structures.

In Equation (7), we also assume that the true volatility of asset price σ_t^2 is the inverse Box–Cox transformation for the latent volatility θ_t, that is,

$$\sigma_t^2 = \begin{cases} (1 + \delta\theta_t)^{1/\delta}, & \text{if } \delta \neq 0; \\ \exp(\theta_t), & \text{if } \delta = 0. \end{cases} \tag{8}$$

where δ is a new transformation parameter that is not related to the transformation parameters $(\lambda_1, \ldots, \lambda_K)$ for realized measures. This inverse transformation includes the standard exponential transformation function for θ_t when $\delta = 0$. Again, the dynamics of θ_t is assumed to be the same multiple component volatility process as (4). Note that the relationship between the volatility σ_t^2 and the latent volatility θ_t could also be equivalently represented as the following Box–Cox transformation function,

$$\theta_t = \begin{cases} (\sigma_t^{2\delta} - 1)/\delta, & \text{if } \delta \neq 0; \\ \log \sigma_t^2, & \text{if } \delta = 0. \end{cases} \tag{9}$$

To consider the possible asymmetric effect between daily returns and volatility components, we do not assume ε_t to be independent of η_t in (4), and the correlation coefficients are ρ_i, $i = 1, \ldots, m$. A negative dependence between returns and volatility components is often referred to as the leverage effect (Yu 2005). If $\rho_i < 0$, a fall in the stock price/return leads to an increase of volatility. Denote $z_t = (\varepsilon_t - \rho' \text{diag}\{\sigma_{1\eta}^{-1}, \ldots, \sigma_{m\eta}^{-1}\}\eta_t)/\sqrt{1 - \rho'\rho}$, where $\rho = (\rho_1, \ldots, \rho_m)'$ is a $m \times 1$ vector. Then the daily return can be rewritten as

$$y_t = \sigma_t \rho' \text{diag}\{\sigma_{1\eta}^{-1}, \ldots, \sigma_{m\eta}^{-1}\}\eta_t + \sqrt{1 - \rho'\rho} \cdot \sigma_t z_t. \tag{7'}$$

Here z_t is an iid standard normal random variable, and the disturbance terms z_t and η_t are mutually independent.

To summarize, the Equations (3), (4), (7), and (9) provide a complete formulation of the realized BCSV model for the daily returns and realized measures, and the bias parameters in Equation (3) can be now identified. The model suggested by Yu, Yang, and Zhang (2006) is a reduced model of our system, which only used (4), (7), and (9) based on only the daily returns without the realized measures. By adding some realized measures, our model carries more information for identifying the latent volatility or stochastic volatility. If all transformation parameters are equal to zero, that is, $\lambda_1 = \cdots = \lambda_K = \delta = 0$, we have the realized SV model with the log transformation specification proposed by Takahashi, Omori, and Watanabe (2009) and Koopman and Scharth (2013). For easy comparison, we will refer it to as the realized LNSV model.

3. ESTIMATION METHODS

3.1 The Likelihood of the Realized BCSV Model

Let ψ be a fixed and unknown parameter vector containing all model parameters. We partition the parameter vector into four subvectors: λ includes the Box–Cox transformation parameters used for transforming the realized measures, ψ_{ssf} includes all the parameters in (6), ψ_{rm} includes the parameters in (5), and ψ_{sv} includes the parameters in (7). Specifically, $\psi = (\lambda', \psi_{sv}', \psi_{ssf}', \psi_{rm}')'$, where

$$\lambda = \{\lambda_1, \ldots, \lambda_K\}, \psi_{sv} = \{\rho', \delta\}, \psi_{ssf} = \{\phi', \sigma_\eta'\}, \text{ and}$$
$$\psi_{\text{rm}} = \{\mu, \tau', \beta', \sigma_u', \text{vecl}(R)\},$$

and vecl(R) contains all the lower-diagonal elements of the correlation matrix as a vector, that is, $\rho_{u,ij}$ for $1 \leq j < i \leq K$.

Define the observation series of daily returns $y_{1:T} = (y_1, \ldots, y_T)'$, the observation series of realized measures $x_{1:T} = (x_1', \ldots, x_T')'$, and the series of unobservable state variables $\alpha_{1:T} = (\alpha_1', \ldots, \alpha_T')'$. The likelihood function of the realized BCSV model is then given by

$$L(y_{1:T}, x_{1:T}; \psi) = p(x_{1:T}; \psi) p(y_{1:T} | x_{1:T}; \psi), \tag{10}$$

where $p(x_{1:T}; \psi)$ is the likelihood of the linear state space model given by (5) and (6) associated with the realized measures and $p(y_{1:T} | x_{1:T}; \psi)$ is the conditional density associated with the daily returns. In the following, we obtain these two parts of likelihood function (10) separately.

3.1.1 The Kalman Filter and Likelihood Function for Realized Measures. For the state space given by (5) and (6), standard Kalman filter can be used to provide a recursive algorithm for computing the minimum mean squared error estimator of α_t conditional on $x_1^{(\lambda)}, \ldots, x_{t-1}^{(\lambda)}$ or $x_{1:t-1}^{(\lambda)}$, that is,

$$\alpha_{t|t-1} = E(\alpha_t | x_{1:t-1}^{\lambda}) = E(\alpha_t | x_1^{(\lambda)}, \ldots, x_{t-1}^{(\lambda)})$$

and its mean squared error (MSE)

$$P_{t|t-1} = E[(\alpha_{t|t-1} - \alpha_t)(\alpha_{t|t-1} - \alpha_t)' | x_{1:t-1}^{\lambda}].$$

More specifically, the Kalman filter is the set of recursions

$$\begin{aligned} v_t &= x_t^{(\lambda)} - d - Z\alpha_{t|t-1}, \quad & F_t &= Z P_{t|t-1} Z' + GG', \\ K_t &= (T P_{t|t-1} Z' + HG') F_t^{-1}, \quad & L_t &= T - K_t Z, \\ \alpha_{t+1|t} &= c + T\alpha_{t|t-1} + K_t v_t, \quad & P_{t+1|t} &= T P_{t|t-1} T' \\ & & &\quad + HH' - K_t F_t K_t'. \end{aligned} \tag{11}$$

The initial conditions $\alpha_{1|0} = a_1$ and $P_{1|0} = P_1$ are assumed to be unconditional mean and variance matrix.

The filter innovations (one step ahead prediction errors) are denoted by $v_t = x_t^{(\lambda)} - E(x_t^{(\lambda)} | x_{1:t-1}^{(\lambda)})$ and their variance matrix by

$$F_t = \text{var}(v_t) = \text{var}\{x_t^{(\lambda)} - E(x_t^{(\lambda)} | x_{1:t-1}^{(\lambda)})\}.$$

These two quantities form the necessary ingredients for the likelihood. Under the normality assumption, the conditional density of x_t can be written as

$$p(x_t | x_{1:t-1}; \lambda, \psi_{ssf}, \psi_{rm}) = \frac{1}{\kappa_t} \cdot \frac{1}{(2\pi)^{K/2} |F_t|^{1/2}} \cdot \exp\left\{-\frac{1}{2} v_t' F_t^{-1} v_t\right\} \cdot \prod_{j=1}^{K} x_{j,t}^{\lambda_j - 1}, \tag{12}$$

where ψ_{rm} and ψ_{ssf} are the parameter vectors including all parameters in the state space representation given by (5) and (6), $x_{1:t-1}$ denotes the available information of x up to $t - 1$. The term κ_t in (12) is a normalizing constant and depends on the fitted values of x_t and F_t. Note that the distribution of x_t in (12) is a truncated normal distribution for any $\lambda_j \neq 0$. See Freeman and Modarres (2006) for details about the multivariate extension for the Box–Cox transformation. However, in most situations $\kappa \approx 1$ for applications, for example, for univariate case see Box and Cox (1964) and Chen (1995), and for multivariate case see Riani (2004, 2009), hence the term $1/\kappa_t$ in (13) can be ignored.

By approximating κ_t with one, the likelihood function associated with the realized measures consequently becomes

$$p(x_{1:T};\lambda,\psi_{ssf},\psi_{rm}) \approx (2\pi)^{-TK/2}\cdot\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\ln|F_t| -\frac{1}{2}\sum_{t=1}^{T}v_t'F_t^{-1}v_t\right\}\cdot\prod_{t=1}^{T}\prod_{j=1}^{K}x_{j,t}^{\lambda_j-1}. \quad (13)$$

Therefore, based on the standard Kalman filter recursion, we can directly get the first part of likelihood function $p(x_{1:T};\psi) = p(x_{1:T};\lambda,\psi_{ssf},\psi_{rm})$ in (10).

3.1.2 The Likelihood of Returns Based on Selected Information. The second part of likelihood function $p(y_{1:T}|x_{1:T};\psi)$ is more complicated because the conditional volatility variable σ_t (or equivalently, θ_t) appears in both the conditional return Equation (7) and the latent volatility Equation (4) of the model. In fact, it is a high dimensional integral of the signal α, that is,

$$p(y_{1:T}|x_{1:T};\psi) = \int\prod_{t=1}^{T}p(y_t|x_t,\alpha_t,\alpha_{t+1};\psi_{sv}) \times p(\alpha_t|\alpha_{t-1};\psi)d\alpha_1,\ldots\alpha_T. \quad (14)$$

Moreover, due to the presence of endogeneity between daily return y_t and the realized measures x_{jt} for $j = 1,\ldots,K$ shown in Koopman and Scharth (2013), the distribution $p(y_t|x_t,\alpha_t,\alpha_{t+1};\psi_{sv})$ is currently not available. Let $\theta_t = \mu + \mathbf{1}'\alpha_t$ and $\eta_t = \alpha_{t+1} - T\alpha_t$, we can further simplify the conditional return density to another form, that is, $p(y_t|x_t,\alpha_t,\alpha_{t+1};\psi_{sv}) = p(y_t|x_t,\theta_t,\eta_t;\psi_{sv})$. Therefore, a complete analysis of likelihood function $p(y_{1:T}|x_{1:T};\psi)$ is infeasible unless we know the exact density of $p(y_t|x_t,\theta_t,\eta_t;\psi_{sv})$.

To evaluate the likelihood in (14), the approximate approach suggested by Koopman and Scharth (2013) is adopted here. First, by using selected information, the likelihood function can be approximated as $p(y_{1:T}|x_{1:T};\psi) \approx \prod_{t=1}^{T}p(y_t|x_{1:T};\psi)$ because the high-frequency information set for calculating the realized volatility at day t includes y_t, while returns introduce new information about the signal only via leverage effects and bias correction. To remove the endogeneity problem, we further approximate the likelihood $p(y_{1:T}|x_{1:T};\psi) \approx \prod_{t=1}^{T}p(y_t|x_{1:T}^{-t};\psi)$ by deleting the information that y_t share with the other realized measures at day t, where $x_{1:T}^{-t} = \{x_1,\ldots,x_{t-1},x_{t+1},\ldots,x_T\}$ and

$$p\big(y_t|x_{1:T}^{-t};\psi\big) = \int p(y_t|\theta_t,\eta_t;\psi_{sv}) \times p\big(\theta_t,\eta_t|x_{1:T}^{-t};\lambda,\psi_{ssf},\psi_{rm}\big)d(\theta_t,\eta_t). \quad (15)$$

The Equation (15) implies that the conditional density of $p(y_t|x_{1:T}^{-t};\psi)$ is effectively the expectation of the density $p(y_t|\theta_t,\eta_t;\psi_{sv})$ with respect to the density $p(\theta_t,\eta_t|x_{1:T}^{-t};\lambda,\psi_{ssf},\psi_{rm})$. Therefore, the evaluation of the integral (15) by two-dimensional numerical or Monte Carlo integration would be straightforward. We compute the mean and variance of the Gaussian density $p(\theta_t,\eta_t|x_{1:T}^{-t};\lambda,\psi_{ssf},\psi_{rm})$ by the deletion smoothing algorithm of de Jong (1989) applied to the models (3) and (4). For more details about the deletion smoothing algorithm see the Appendix in Koopman and Scharth (2013).

3.2 Two-Step Estimation Procedure

Based on the above likelihood analysis, this section presents the two-step maximum likelihood estimation procedure for our proposed model. In the first step, we implement the estimation of the state space model given by (5) and (6) with the transformed realized measures by running the Kalman filter (11) and maximizing the likelihood function $p(x_{1:T};\lambda,\psi_{ssf},\psi_{rm})$ with respect to λ, ψ_{ssf}, and ψ_{rm} in (13). For identification, the optimization is subject to the coefficients τ_j and β_j for $j = 1,\ldots,K$ in ψ_{rm}, which we refer to as ψ_{bias}, and we also restrict the parameters $\beta_j = 1$ $(j = 1,\ldots,K)$ for the realized measures. We take the particular value $\psi_{\text{bias}}^* = \{c_1^*,\ldots,c_K^*\}$, where $c_j^* = \tau_j + \mu$, $j = 1,\ldots,K$, and denote the resulting maximum likelihood estimates as $\widehat{\lambda}^*$, $\widehat{\psi}_{ssf}^*$, $\widehat{\psi}_{rm}^*$, and $\widehat{\psi}_{\text{bias}}^*$. It should be also noted that the first-step estimation is robust to the misspecification of the return Equation (7).

In the second step, the main purpose of maximum likelihood estimation is to implement the estimation of the parameters ψ_{sv} in (7) and the bias parameters ψ_{bias}, where we denote $\mu = \hat{c}_j^* - \tau_j$. Given the parameters estimates $\widehat{\lambda}^*$, $\widehat{\psi}_{ssf}^*$, and $\widehat{\psi}_{rm}^*$ from the first step, we estimate the remaining part of the parameter vector by maximizing the approximate density

$$p^*(y_{1:T}|x_{1:T};\psi) = \prod_{t=1}^{T}\int p(y_t|\theta_t,\eta_t;\psi_{sv})p(\theta_t,\eta_t|x^{-t}; \times \psi_{\text{bias}},\widehat{\lambda}^*,\widehat{\psi}_{ssf}^*,\widehat{\psi}_{rm}^*)d(\theta_t,\eta_t). \quad (16)$$

The optimization is with respect to ψ_{sv} and ψ_{bias} only in the second step. This integral can be evaluated by two-dimensional numerical or Monte Carlo integration via the deletion smoothing algorithm of de Jong (1989).

Now, we summarize the two-step estimation procedure based on the preceding discussion:

Step 1. Take the parameters $c_j^* = \tau_j + \mu$, $j = 1,\ldots,K$, then estimate parameters of the model given by (5) and (6) via the maximum likelihood estimation procedure based on the Kalman filter with the transformed realized measures in (2), and obtain parameter estimates $\widehat{\lambda}^*$, $\widehat{\psi}_{ssf}^*$, and $\widehat{\psi}_{rm}^*$.

Step 2. Using a maximum likelihood method via the deletion smoothing algorithm and the numerical or Monte Carlo integration with the estimated parameters from the first step, estimate (16) along with (7) and (9), and obtain the estimates $\widehat{\psi}_{sv}$ and $\widehat{\psi}_{\text{bias}}$.

3.3 Volatility Estimation With Inverse Transformation

After parameters estimation, we are interested in the volatility estimates σ_t^ν, for $t = 1,\ldots,T$. For the controlling parameter ν, we have the conditional variance σ_t^2 when $\nu = 2$, the conditional standard deviation σ_t when $\nu = 1$, and the standardized return y_t/σ_t when $\nu = -1$, and so on. Given the parameter estimates $\widehat{\psi}$ from the two-step maximum likelihood estimation procedure, there might be several methods to compute the volatilities. In our study, we compute the volatility σ_t^ν based on the Gaussian density $p(\alpha_t|\cdot)$ after running the Kalman filter and smoother algorithm. However, due to the inverse Box–Cox transformation, the estimation of the volatility is not straightforward.

Now we take the smoothed estimate of the volatility σ_t^ν as example. Given the smoothed estimate $\alpha_{t|T}$ and its mean squared error $P_{t|T}$ from the Gaussian density $p(\alpha_t|\mathcal{F}_T;\widehat{\psi})$, we immediately have the smoothed estimate of the Gaussian density $p(\theta_t|\mathcal{F}_T;\widehat{\psi})$ with the mean $\theta_{t|T} = E(\theta_t|\mathcal{F}_T) = \mu + \mathbf{1}'\alpha_{t|T}$ and the variance $V_{t|T} = \mathrm{var}(\theta_t|\mathcal{F}_T) = \mathbf{1}'P_{t|T}\mathbf{1}$, where $\mathcal{F}_T = \{x_1,\ldots,x_T\}$ denotes the whole sample information from the realized measures. With the mean $\theta_{t|T}$ and its variance $V_{t|T}$ known, our purpose is next to derive the moments of σ_t^ν, especially the mean $\sigma_{t|T}^\nu = E(\sigma_t^\nu|\mathcal{F}_T)$. Carroll and Ruppert (1981) and Taylor (1985) argued that one way to address the issue is to apply the inverse Box–Cox transformation and make inferences in the original scale of the data. For the log transformation case that $\delta = 0$, the exact expressions of the n-order moments of σ_t^ν is simply given by $E(\sigma_t^{n\nu}|\mathcal{F}_T) = \exp(s\theta_{t|T} + s^2V_{t|T}/2)$, where the number $s = n\nu/2$; see Johnson, Kotz, and Balakrishnan (1994). However, for the Box–Cox transformation case that $\delta \neq 0$, the n-order moment of σ_t^ν is highly nonlinear, and it is often represented by the integral of a binomial sum with respect to the density of $p(\theta_t|\mathcal{F}_T;\widehat{\psi})$, that is,

$$E\left(\sigma_t^{n\nu}|\mathcal{F}_T\right) = \int (1+\delta\theta_t)^{s/\delta} p(\theta_t|\mathcal{F}_T;\widehat{\psi})d\theta_t$$
$$= \int \sum_{k=0}^{\infty} \binom{s/\delta}{k} (\delta\theta_t)^k p(\theta_t|\mathcal{F}_T;\widehat{\psi})d\theta_t,$$

where the number also satisfies $s = n\nu/2$. It is not hard to prove that the exact expressions for mean and variance are almost a sum of infinite terms. In fact, the mean and variance do have a closed form solution only in particular case, namely $\delta = 0$, and $\delta = 1/p$, $p = 1, 2, 3, \ldots$. For further details, the reader is referred to Freeman and Modarres (2006) and Proietti and Riani (2009). To deal with the estimation problem, some approximate and computational solutions are available to obtain the estimate of the mean of σ_t^2, such as the approximate methods proposed by Taylor (1986a) and Guerrero (1993), and the Monte Carlo method suggested by Proietti and Riani (2009).

To obtain the estimate of the mean of σ_t^ν, we use the Monte Carlo method suggested by Proietti and Riani (2009). For the Monte Carlo evaluation, we draw repeated samples from the conditional distribution of $p(\theta_t|\mathcal{F}_T;\widehat{\psi})$ with its mean $\theta_{t|T}$ and the variance $V_{t|T}$, and then compute the Monte Carlo integration

$$\widehat{\sigma}_{t|T}^\nu = \frac{1}{G}\sum_{g=1}^{G}\left(1+\delta\theta_t^{(g)}\right)^{\nu/2\delta},$$

where $\theta_t^{(g)}$ is draw from $p(\theta_t|\mathcal{F}_T;\widehat{\psi})$ and it needs to satisfy that $\delta\theta_t^{(g)} \geq -1$, for $g = 1,\ldots,G$. We take the value of $G = 1000$ in our empirical application in Section 5.

Similarly, we can employ the Monte Carlo method to compute the prediction estimates of the volatility σ_t^ν or the realized measures x_t by the inverse transformation in the same way.

4. SIMULATION EXPERIMENTS

In this section, we investigate the performance of the two-step estimation method based on the deletion smoothing scheme. We also present Monte Carlo evidence regarding the impact of model misspecification when true process is the realized BCSV model or the realized LNSV model, and show some limited evidence of how the log transformation specification performs in practice if the true process is the realized BCSV model.

For this purpose, we generate 200 sets of data by a simplified version of the system given by the realized BCSV model with $\lambda_1 = \lambda_2 = \delta \neq 0$ (Case 1), and the realized LNSV model with $\lambda_1 = \lambda_2 = \delta = 0$ (Case 2). We also assume that there is only one component of volatility, that is, $m = 1$. Each sample is of length $T = 3000$. The data-generation processes can be formulated as follows: for the realized BCSV model (Case 1), $y_t = \sigma_t\varepsilon_t$, $\sigma_t = (1+\delta\theta_t)^{1/2\delta}$, $\varepsilon_t \sim N(0,1)$, $(x_{jt}^{\lambda_j} - 1)/\lambda_j = \tau_j + \theta_t + u_{jt}$, $u_{jt} \sim N(0,\sigma_{uj}^2)$, $\theta_t = \mu + h_t$, $h_{t+1} = \phi h_t + \eta_t$, $\eta_t \sim N(0,\sigma_\eta^2)$, $t = 1,\ldots,3000$; for the realized LNSV model (Case 2), $y_t = \sigma_t\varepsilon_t$, $\sigma_t = \exp(\theta_t)$, $\varepsilon_t \sim N(0,1)$, $\ln(x_{jt}) = \tau_j + \theta_t + u_{jt}$, $u_{jt} \sim N(0,\sigma_{uj}^2)$, $\theta_t = \mu + h_t$, $h_{t+1} = \phi h_t + \eta_t$, $\eta_t \sim N(0,\sigma_\eta^2)$, $t = 1,\ldots,3000$. In both cases, the disturbance series u_t have the correlation coefficient $\rho_{u,21}$, that is, $\mathrm{corr}(u_{1t},u_{2t}) = \rho_{u,21}$, and the disturbances ε_t and η_t have the correlation coefficient ρ_1 and are independent from u_t. For each sample, we estimate the model twice: (i) with the Box–Cox transformation specification, and (ii) with the log transformation specification. When estimating the model with both two specifications, the two-step estimation procedure introduced in Section 3.2 is employed.

The specific parameter values we assign are summarized below. In Case 1, the true process is the realized BCSV model, and the true parameters are assigned as

$$\rho_1 = -0.30,\ \lambda_1 = \lambda_2 = \delta = -0.05,\ \tau_1 = -0.10,$$
$$\tau_2 = -0.30,\ \beta_1 = \beta_2 = 1,$$
$$\sigma_{u1} = \sigma_{u2} = \sqrt{0.05},\ \rho_{u,21} = 0.80,\ \mu = 0.40,\ \phi = 0.98,$$
$$\sigma_\eta = \sqrt{0.05}.$$

In Case 2, the true process is the realized LNSV model, and the parameter values are

$$\rho_1 = -0.30,\ \lambda_1 = \lambda_2 = \delta = 0,\ \tau_1 = -0.10,$$
$$\tau_2 = -0.30,\ \beta_1 = \beta_2 = 1,$$
$$\sigma_{u1} = \sigma_{u2} = \sqrt{0.05},\ \rho_{u,21} = 0.80,\ \mu = 0.40,$$
$$\phi = 0.98,\ \sigma_\eta = \sqrt{0.05}.$$

Table 1 reports the results obtained from the two-step estimation procedure. For the case of the realized BCSV process ($\lambda_1 = \lambda_2 = \delta = -0.05$) or the realized LNSV process ($\lambda_1 = \lambda_2 = \delta = 0$), we restrict the parameters $\beta_1 = \beta_2 = 1$ and estimate the samples by both the Box–Cox transformation and log transformation specifications with one volatility component ($m = 1$). Several observations can be seen from the upper part of Table 1. First, the two-step MLE estimator exhibits very good performance. All means are very close to the true parameter values assigned for the DGP except for Case 1 when estimating with the log transformation specification. Moreover, these means and root square mean errors are nearly identical. The above result

Table 1. The results of the simulation experiments

	Case 1: Realized BCSV			Case 2: Realized LNSV		
	True	Box-Cox	Logarithm	True	Box-Cox	Logarithm
ρ_1	−0.30	−0.2951 (0.026)	−0.2943 (0.025)	−0.30	−0.2953 (0.025)	−0.2953 (0.026)
δ	−0.05	−0.0494 (0.026)	—	0.00	0.0004 (0.029)	—
λ_1	−0.05	−0.0505 (0.007)	—	0.00	−0.0005 (0.007)	—
λ_2	−0.05	−0.0506 (0.008)	—	0.00	−0.0006 (0.008)	—
τ_1	−0.10	−0.1060 (0.031)	−0.1067 (0.026)	−0.10	−0.1060 (0.031)	−0.1056 (0.026)
τ_2	−0.30	−0.3059 (0.031)	−0.3092 (0.027)	−0.10	−0.3059 (0.031)	−0.3054 (0.026)
σ_{u1}	0.2236	0.2234 (0.006)	0.2290 (0.006)	0.2236	0.2234 (0.006)	0.2234 (0.006)
σ_{u2}	0.2236	0.2238 (0.006)	0.2252 (0.006)	0.2236	0.2238 (0.006)	0.2238 (0.006)
$\rho_{u,21}$	0.80	0.7995 (0.010)	0.7985 (0.010)	0.80	0.7995 (0.010)	0.7994 (0.010)
μ	0.40	0.3972 (0.204)	0.4346 (0.204)	0.40	0.3973 (0.204)	0.3971 (0.200)
ϕ	0.98	0.9790 (0.004)	0.9790 (0.004)	0.98	0.9790 (0.004)	0.9790 (0.004)
σ_η	0.2236	0.2233 (0.007)	0.2265 (0.008)	0.2236	0.2233 (0.007)	0.2233 (0.007)
aSK(v_{1t})	0	−0.0043	0.0572	0	−0.0043	−0.0047
aSK(v_{2t})	0	−0.0071	0.0436	0	−0.0071	−0.0076
aK(v_{1t})	3	3.0957	3.1268	3	3.0958	3.0959
aK(v_{2t})	3	3.0839	3.1135	3	3.0840	3.0861

NOTE: For each cell, the statistics given are based on 200 simulated samples, each consisting of a time series of length 3000. The quantities aSK(·) and aK(·) denote the average skewness and kurtosis of the innovations, respectively. The mean and root mean squared error (in parentheses) for each estimator are shown.

is consistent with the Monte Carlo result from the only LNSV model given by Koopman and Scharth (2013). Second, we find that the parameter estimates may get worse if the realized BCSV model is misspecified and estimated by the log transformation specification. In particular, the biases in the estimates of τ_2, σ_{u1}, σ_{u2}, μ, or σ_η are significantly unnegligible. This is not surprising because they are based on the wrong model specification. Third, for Case 2, we get nearly the same sample performance when modeling the sample with the Box–Cox transformation and log transformation specifications. The Monte Carlo experiment result shows that the means of the parameter estimates are almost the same and the mean of the estimated λ is very close to zero. The above results also imply that the Box–Cox transformation specification would be preferable regardless of the true process being the realized BCSV model or the realized LNSV model.

We have also examined whether the use of log transformation leads to excessive skewness and kurtosis and non-Gaussian in the residuals. By 1000 replications of Monte Carlo simulation, we run different specifications (Box–Cox or logarithm) on the generated data from the realized BCSV model or the realized LNSV model. The corresponding results of average skewness and kurtosis are reported in the bottom of Table 1. It is shown that the average skewness and kurtosis for the logarithm specification of the realized BCSV model obviously increase but not very large compared to other correct specifications. This implies that the misuse of log transformation for the data generated from the realized BCSV model may lead to the certain excessive skewness and kurtosis and so the non-Gaussianity in the residuals.

In addition, to assess the quality of a normality test of $\lambda = 0$, we simulate the test statistics for the Z-test (standardized statistic) with the null hypothesis $H_0 : \lambda = 0$ and concern with two types of error. We repeat another 1000 replications for Case 1 and Case 2, and get all estimates of parameters λ_1 and λ_2 and their corresponding standard errors. For Case 2 with $\lambda_1 = \lambda_2 = 0$, the simulated Type I error rates at the significance level (0.05) are 0.045 and 0.046 for λ_1 and λ_2, respectively. This result indicates that the sampling distribution of the Box–Cox transformation parameter is approximately standard normal. Similarly for Case 1, our simulation result shows that the powers of Z-test for λ_1 and λ_2 are almost one.

5. EMPIRICAL APPLICATION

5.1 Data Series

The empirical analysis is based on the daily data taken from the "Oxford-Man Institute's realized library version 0.2" (available at *http://realized.oxford-man.ox.ac.uk*). It contains daily returns and realized volatilities constructed with intraday prices. The sample period is from January 3, 2000, to May 15, 2012. We select five stock indices for our empirical analysis, including Standard & Poor 500 Index (SP500), FTSE 100 Index (FTSE), Nikkei 225 Index (Nikkei), Deutscher Aktien-Index (DAX), and Dow Jones Industrial Average Index (DJIA). Specifically, we use their daily returns and two realized volatility measures: realized variance estimators (RV5m, computed by the sum of squared 5-minute log returns) and realized kernel (RK, computed tick by tick data, after cleaning, using the methodology of Barndorff-Nielsen et al. 2008) estimators for each stock index. In addition, the missing values in the data are deleted for convenience.

Table 2 reports summary statistics for the daily returns and two realized measures for each stock index. We observe that the skewness and kurtosis of both realized measures show evident nonnormal distribution, while the distributions of log transformed realized measures are close to normal. This result is consistent with most empirical studies mentioned in the introduction. It is also noted that the autocorrelation functions (ACFs) are still very large even when the lag is 22 days or roughly one month, indicating the existence of strong volatility persistence. In addition, we can clearly see the discrepancies between the

Table 2. Descriptive statistics of returns and realized kernels for stock indices

	SP500 RET	FTSE RET	Nikkei RET	DAX RET	DJIA RET
T	3082	3099	2987	3130	3084
Mean	−0.001	−0.045	−0.049	−0.047	0.015
Std. Dev.	1.314	1.041	1.218	1.434	1.257
Skewness	−0.132	−0.125	−0.405	−0.061	0.026
Kurtosis	9.597	6.616	13.642	7.445	10.363

	SP500		FTSE		Nikkei		DAX		DJIA	
	RV5m	RK	RV5m	RK	RV5m	RK	RV5m	RK	RV5m	RK
Mean	1.454	1.388	1.059	1.049	1.208	1.250	2.147	2.117	1.406	1.310
Std. Dev.	2.910	2.935	1.905	1.746	1.793	2.030	3.487	3.607	3.016	2.895
Skewness (raw)	10.014	13.413	9.373	6.961	8.535	8.805	9.689	7.483	11.734	13.676
Kurtosis (raw)	185.13	335.65	153.85	77.260	109.32	109.75	72.888	88.433	242.74	331.47
ACF(1)	0.779	0.806	0.849	0.858	0.785	0.771	0.845	0.851	0.747	0.796
ACF(5)	0.703	0.719	0.781	0.789	0.666	0.648	0.765	0.774	0.675	0.712
ACF(22)	0.556	0.565	0.663	0.664	0.507	0.498	0.635	0.639	0.531	0.557
Skewness (log)	0.489	0.504	0.319	0.268	0.258	0.327	0.283	0.338	0.546	0.612
Kurtosis (log)	3.338	3.435	2.968	2.925	3.627	3.782	2.955	3.045	3.523	3.712

NOTE: The top panel reports the descriptive statistics of returns (RET), and the bottom panel reports the descriptive statistics of realized measures including the 5-min realized variance (RV5m) and the realized kernel (RK). The skewness and kurtosis of both the raw and log transformed realized measures are computed. Std. Dev. is the daily standard deviation of daily returns or realized mesures. ACF(j) denotes the autocorrelation functions at j lag for the realized measures.

sample variance (or the square of standard deviation) of stock returns and the volatility mean of realized measures. This characteristic of stock market data is captured by the bias term τ in the realized stochastic volatility model.

5.2 Estimation Results

We model the returns and two realized measures by the realized BCSV model introduced in Section 2. On the basis of the Bayesian information criterion, we find that SP500, Nikkei, DAX, and DJIA require $m = 3$ AR(1) processes to model the latent volatility θ_t in (4), while FTSE requires $m = 2$. In addition, we model the realized measures by (3) with the restriction that $\beta_1 = \beta_2 = 1$, following Takahashi, Omori, and Watanabe (2009), and Koopman and Scharth (2013). All estimation results are obtained by the two-step estimation method of Section 3.2.

Table 3 reports the parameter estimates for each stock index with the realized BCSV specification. As shown in the table, all estimates of the transformation parameters λ_1 and λ_2 are relatively small but significantly differ from zero at 5% level,

Table 3. Parameter estimation results of the realized BCSV model

Parameter	SP500	FTSE	Nikkei	DAX	DJIA
ρ_1	−0.2766** (0.106)	−0.4163** (0.099)	−0.0724 (0.143)	−0.2777** (0.097)	−0.2283** (0.086)
ρ_2	−0.2353** (0.096)	−0.1880** (0.053)	−0.2777** (0.090)	−0.2839** (0.082)	−0.2696** (0.078)
ρ_3	−0.3795** (0.058)		−0.1741** (0.047)	−0.0738* (0.036)	−0.2958** (0.058)
δ	−0.0017 (0.043)	−0.0209 (0.039)	−0.1175** (0.045)	0.0559 (0.038)	−0.0111 (0.045)
λ_1	−0.0535** (0.009)	−0.0450** (0.008)	−0.0148 (0.011)	0.0244** (0.008)	−0.0389** (0.009)
λ_2	−0.0520** (0.008)	−0.0210** (0.008)	−0.0316** (0.011)	0.0139 (0.008)	−0.0488** (0.009)
τ_1	−0.1448** (0.037)	−0.0870** (0.037)	−0.0105 (0.035)	0.0361 (0.037)	−0.1122** (0.035)
τ_2	−0.1740** (0.037)	−0.0781* (0.037)	−0.0149 (0.035)	0.0167 (0.037)	−0.1489** (0.035)
σ_{u1}	0.4298** (0.015)	0.3786** (0.012)	0.3225** (0.022)	0.2913** (0.048)	0.4677** (0.014)
σ_{u2}	0.3839** (0.016)	0.3612** (0.012)	0.3525** (0.020)	0.2656** (0.052)	0.3831** (0.016)
$\rho_{u,21}$	0.9508** (0.004)	0.9270** (0.005)	0.8892** (0.014)	0.8676** (0.047)	0.9325** (0.004)
μ	−0.1436 (0.254)	−0.4431 (0.304)	−0.2205 (0.223)	0.2211 (0.295)	−0.2186 (0.232)
ϕ_1	0.9941** (0.004)	0.9948** (0.003)	0.9959** (0.003)	0.9949** (0.003)	0.9932** (0.004)
ϕ_2	0.9419** (0.053)	0.8179** (0.061)	0.9515** (0.020)	0.9378** (0.053)	0.9283** (0.050)
ϕ_3	0.6472** (0.150)		0.5769** (0.119)	0.3932** (0.149)	0.6127** (0.192)
$\sigma_{\eta 1}$	0.0856** (0.036)	0.0931** (0.019)	0.0523** (0.017)	0.0872** (0.032)	0.0895** (0.026)
$\sigma_{\eta 2}$	0.1463** (0.042)	0.2247** (0.018)	0.1326** (0.024)	0.1405** (0.026)	0.1496** (0.042)
$\sigma_{\eta 3}$	0.2249** (0.037)		0.2524** (0.030)	0.3047** (0.045)	0.1968** (0.041)

NOTE: The reported results are estimated by the realized BCSV model. The vector of realized measures include the realized variance (5-min) and the realized kernel estimators. Parameter estimates is carried out by the two-step method and is based on the deletion smoothing algorithm as presented in Section 3.2. The standard deviation errors of parameter estimates are reported in parentheses. "∗∗" and "∗" denote that the parameter estimate is significant according to the standard Z-test at 1% and 5% level of significance, respectively.

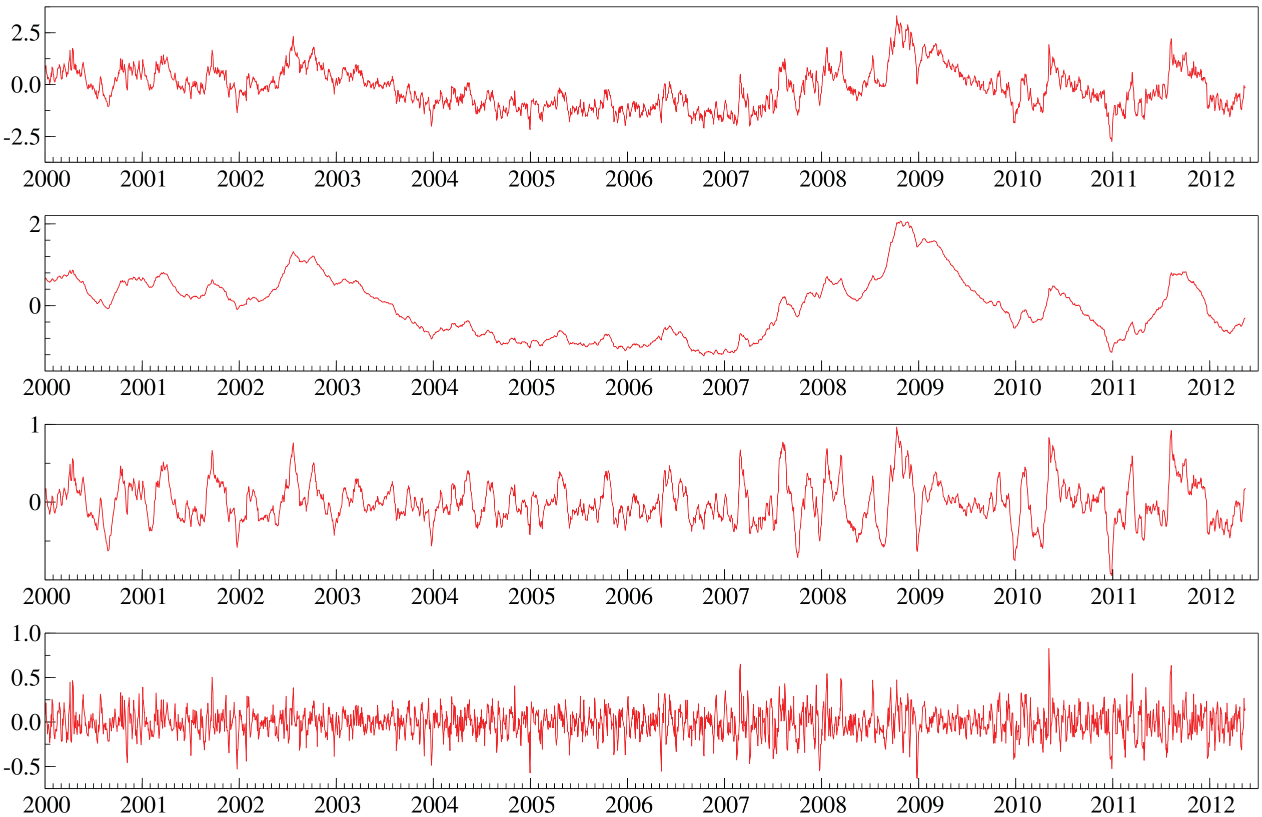

Figure 1. Estimated latent volatility signal and individual components for S&P 500. Note: The top of the four subfigures is the estimated latent volatility signal θ_t and the following are three volatility components h_{it}, $i = 1, 2, 3$.

except λ_1 for Nikkei and λ_2 for DAX. It indicates that the Box–Cox transformation specification for modeling these realized measures is indeed needed and is more appropriate than the logarithm transformation specification. However, for the transformation parameter δ, only the estimate for Nikkei is barely significant, showing that it is probably not necessary to have a Box–Cox transformation between the conditional volatility σ_t and the volatility signal θ_t.

For the bias parameters τ_1 and τ_2 in (3), the estimates for SP500, FTSE, and DJIA are significantly negative. The negative biased estimate reflects the bias between the transformed realized volatility and the latent stochastic volatility signal, and also explains their downwards discrepancies between the mean of realized measures and the variance of returns. However, the bias estimates for Nikkei and DAX are not significant, implying that there are no evident differences between the transformed realized measures and the volatility signal for these two stock indices.

The parameter estimates of autoregressive coefficients ϕ_i and standard deviations $\sigma_{\eta,i}$ suggest that three AR(1) components for θ_t work well for SP500, Nikkei, DAX, and DJIA and two components for FTSE. We find that the first AR components in these models are all near unit root processes with estimated AR coefficients larger than 0.99 and relatively small estimated standard deviations. The second AR components are persistent processes with estimated autoregressive coefficients between 0.82 and 0.95. Except for FTSE, the third AR components are of relatively short memory with AR coefficients estimated as low as 0.39 and reaching a maximum of 0.65. Using SP500 as example, Figure 1 shows the volatility signal and its three volatility components, where the former is the sum of the volatility components and the volatility mean μ. This figure shows that the first component determines the long-range trend of the volatility signals, and the third components capture the short-run processes, respectively.

From the estimated correlation coefficients ρ_i, it is seen that for all stock indices the volatility components have negative, highly significant correlations with the stock returns, indicating that leverage effects significantly impact the volatility components. The estimated leverage effects obtained here contrast with previous studies which have found leverage effects for both long-run and short-run components; see, for example, Koopman and Scharth (2013), and for transitory components only, see Engle and Lee (1999).

Figure 2 presents the smoothed volatility estimates of σ_t, and the estimated standardized returns calculated by $\hat{\varepsilon}_t = y_t\hat{\sigma}_t^{-1}$, where $\hat{\sigma}_t^{-1}$ is the estimated inverse of the standard deviation using the computation method in Section 3.3. The estimated nonparametric density estimates and the QQ plots for the standardized returns are also reported in the figure. The result shows that the standardized returns of SP500, FTSE, Nikkei, and DAX are approximately normal, which validates the normal hypothesis for the disturbance term ε_t in (7). The skewness and kurtosis of $\hat{\varepsilon}_t$ reported in Table 4 provide further evidence in support of this result. However, for DJIA, the standardized returns still have obviously heavy tail which can be seen in the fourth and last panels of DJIA in Figure 2. It might be useful to use the student-t distribution to model this nonnormal disturbance. The

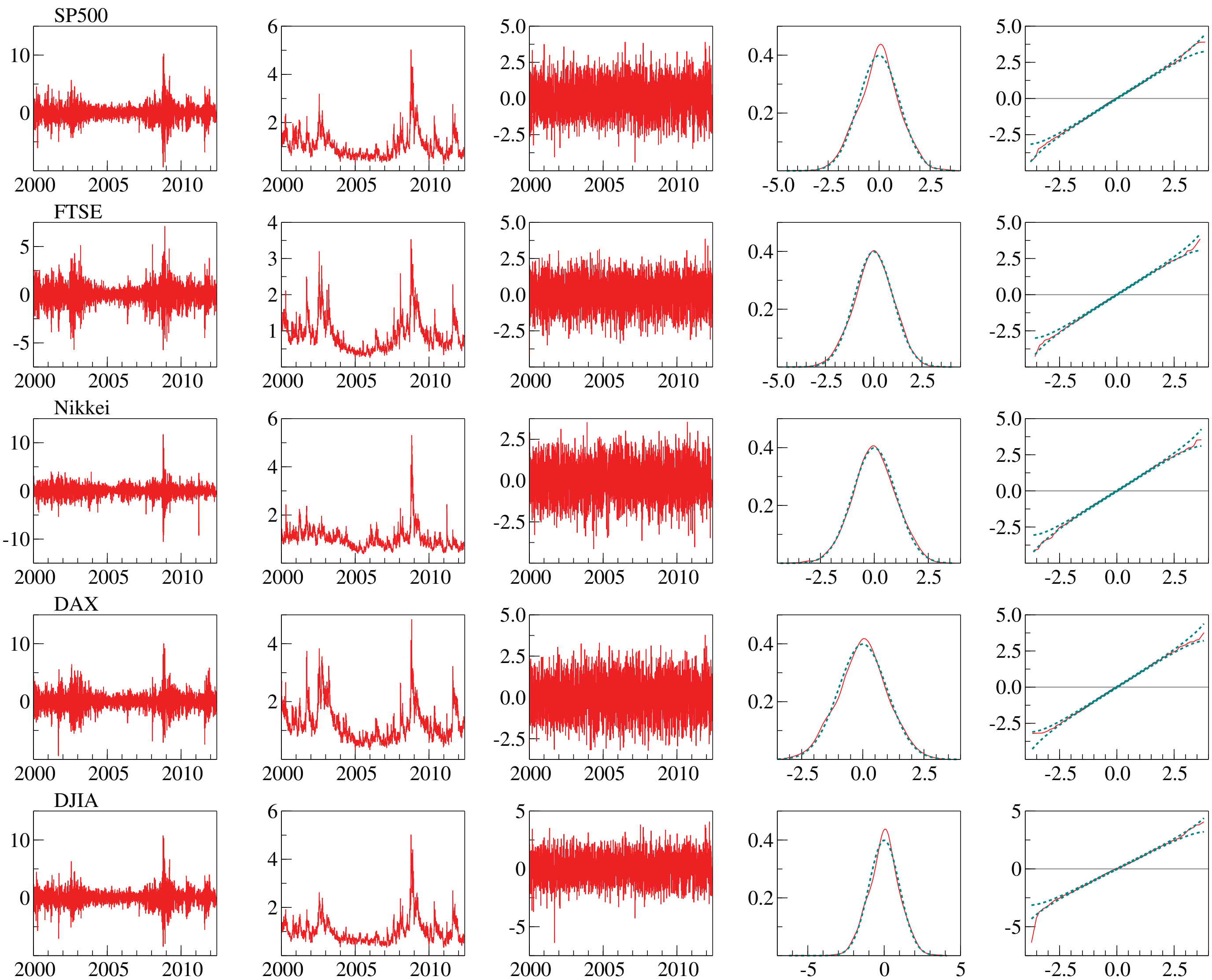

Figure 2. Volatility estimates and standardized returns. Note: From the left to the right, the panels plot for each stock index the daily returns, the estimated volatilities (σ_t), the estimated standardized returns, the nonparametric density estimates, and the QQ plots, respectively. In the fourth panel, the solid line denotes the density estimates and the dot line denotes the standard normal or $N(0, 1)$. In the last panel, the solid line denotes the QQ line and the dot lines are pointwise asymptotic 95% standard error bands, as derived in Engler and Nielsen (2009).

relevant estimation problem of the realized stochastic volatility model will be straightforward on the basis of Koopman and Scharth (2013).

5.3 Model Comparison

Table 4 presents an in-sample comparison between the Box–Cox transformation specification and the log transformation specification. From the table, the log-likelihood values of the realized BCSV model are markedly larger for all stock indices than those of the realized LNSV model. The result shows that we can reject the null of the log transformation specification if constructing the conventional likelihood ratio test on $\lambda_1 = \lambda_2 = \delta = 0$, with the computed Monte-Carlo p-values being less than 0.05 for all stock indices (0.001, 0.001, 0.003, 0.014, and 0.002 for SP500, FTSE, Nikkei, DAX, and DJIA, respectively, based on 1000 times of experiments). Taking SP500, for example, in Figure 3 we show the log-likelihood function as a function of the transformation parameter maximized over all other parameters. It can be seen that it is concave and the maximum is significantly higher than that setting λ to zero. Therefore, consistent with the estimation result mentioned before, the Box–Cox transformation specification is more preferable for these stock indices than the log transformation specification. Moreover, the results of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) in the table also show further evidence. Although the BIC values for the Nikkei and DAX using the realized BCSV model are not lower than those using the realized LNSV model, we can still get the smaller BIC values of the realized BCSV model than the pure realized LNSV model if restricting some of insignificant parameters in term of the parameter estimates in Table 3. For instance, when we restrict $\lambda_2 = 0$ and $\delta = 0$, the BIC value of the BCSV model is 12,661, smaller than that of the realized LNSV model.

Table 4. In-sample comparison of alternative models

	Realized BCSV					Realized LNSV				
	SP500	FTSE	Nikkei	DAX	DJIA	SP500	FTSE	Nikkei	DAX	DJIA
log.lik	−2927.5	−693.9	−3613.1	−6264.3	−3290.8	−2949.2	−731.9	−3625.0	−6273.0	−3310.2
AIC	5891.0	1417.8	7262.3	12565	6617.7	5928.4	1487.8	7279.9	12576	6650.4
BIC	5999.6	1507.8	7370.3	12673	6726.3	6018.9	1559.8	7370.0	12667	6740.9
Skewness (v_{1t})	0.2357	0.3742	0.5250	0.1530	0.2424	0.3509	0.4947	0.5570	0.2196	0.3343
Kurtosis (v_{1t})	3.9211	5.7552	4.5415	4.9045	4.1203	4.0146	6.0650	4.6361	4.9703	4.2263
Skewness (v_{2t})	0.1943	0.3299	0.4363	0.2549	0.2805	0.3159	0.3893	0.5123	0.2879	0.4039
Kurtosis (v_{2t})	4.1123	6.1815	4.2070	5.1552	4.7267	4.2925	6.2444	4.3840	5.1988	5.0255
Skewness (ε_t)	−0.0269	−0.1129	−0.0903	−0.0409	−0.0832	−0.0281	−0.1144	−0.0917	−0.0420	−0.0883
Kurtosis (ε_t)	3.1606	2.9989	3.2321	2.9492	3.4749	3.1579	3.0050	3.2601	2.9442	3.4840
RMSE (y_t^2)	4.2873	2.2822	4.5975	4.5737	4.0964	4.2251	2.2821	4.7802	4.5576	4.0416
MAE (y_t^2)	1.6483	1.0725	1.4730	2.0211	1.5135	1.6528	1.0693	1.4684	2.0442	1.5168
RMSE (RV5m)	1.9012	1.4253	1.2257	2.3133	2.0296	1.9064	1.4249	1.2242	2.3118	2.0246
MAE (RV5m)	0.6698	0.4589	0.4741	0.8282	0.6597	0.6611	0.4550	0.4735	0.8324	0.6570
RMSE (RK)	2.0300	1.2127	1.3978	2.4506	1.9659	2.0346	1.2131	1.4090	2.4532	1.9660
MAE (RK)	0.6124	0.4270	0.5241	0.8208	0.5517	0.6068	0.4259	0.5234	0.8218	0.5485

NOTE: log.lik denotes the estimated log-likelihood value of the realized stochastic volatility model, AIC and BIC denote the Akaike information criterion and Bayesian information criterion, and RMSE and MAE denote the root mean squared error and the mean absolute error of volatility estimates. We take squared returns, 5-min realized variance (RV5m), and realized kernel (RK) as possible volatility proxies.

We next consider the improvement we have made of our advantages of the Box–Cox transformation specification over the log transformation specification. This analysis is through a selection of diagnostic statistics such as the skewness and the kurtosis, based on the one-step ahead prediction residuals v_{jt}, $j = 1, 2$. By comparing the results of the realized BCSV model with those of the realized LNSV model, we find that the skewness and the kurtosis of v_t improve a little bit. This is consistent with the previous results from the Monte Carlo simulation in Section 4. However, the realized BCSV model still suffers the problem of excessive skewness and high kurtosis. We think the possible reason of this phenomenon may be caused by jumps and/or heteroscedasticity in the realized volatilities (see also Corsi et al. 2008). A direct generalization to the nonlinear and/or non-Gaussian problem may have some difficulty in estimation and so it will exceed our discussion. We will leave this direction for future study. In addition, no autocorrelations for the residuals v_{1t} and v_{2t} can be found from the autocorrelation functions (not reported here). It indicates that multiple autoregressive processes of volatility components for modeling the volatility persistence are adequate.

Table 4 compares in-sample forecast performance of the realized BCSV and LNSV models by comparing their root mean squared error (RMSE) and mean absolute error (MAE) for

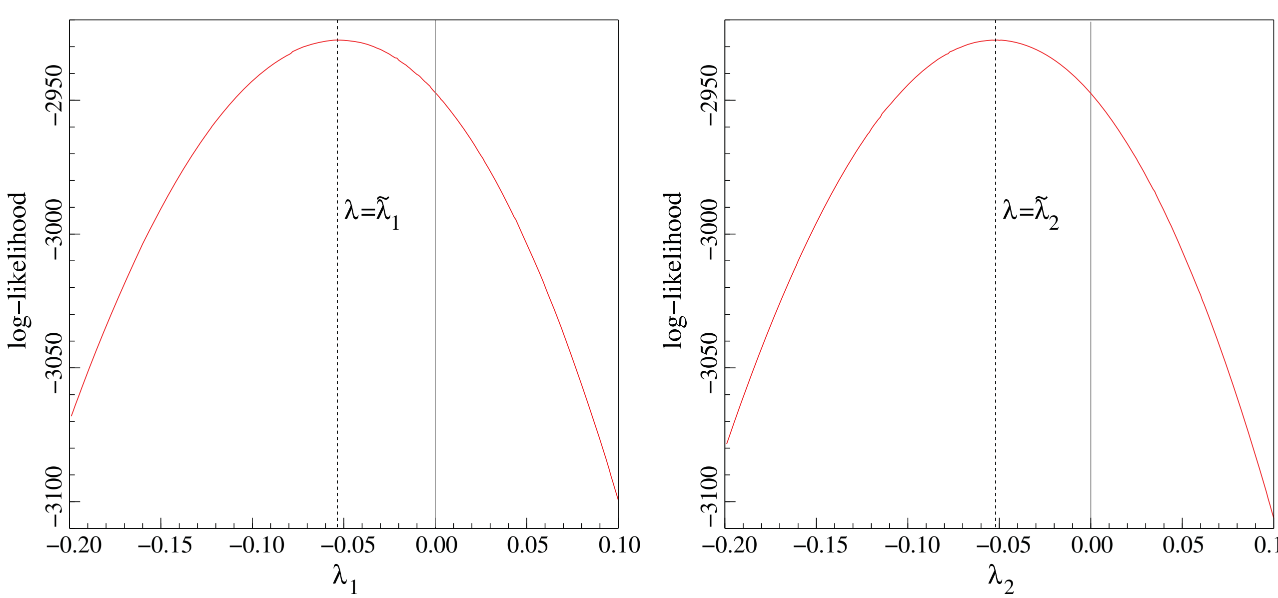

Figure 3. Estimated log-likelihood against transformation parameters: S&P 500. Note: The log-likelihoods are obtained from the two-step MLE estimation of the realized BCSV model. Each fixed transformation parameter ranges from −0.2 to 0.1. The dotted lines $\lambda = \tilde{\lambda}_1$ (left) and $\lambda = \tilde{\lambda}_2$ (right) are the corresponding points that arrive at maximum log-likelihood values for two parameters λ_1 and λ_2, respectively.

Table 5. Out-of-sample forecasting results

	Relative RMSE					Relative MAE				
	SP500	FTSE	Nikkei	DAX	DJIA	SP500	FTSE	Nikkei	DAX	DJIA
(a) y_t^2										
$h = 1$	1.0015	0.9946	1.0418	0.9921	0.9987	1.0024	0.9958	1.0004	0.9994	1.0007
$h = 5$	1.0019	0.9956	1.0420	0.9902	0.9993	1.0027	0.9968	1.0042	0.9977	1.0012
$h = 22$	1.0019	0.9968	1.0463	0.9870	0.9999	1.0025	0.9975	1.0262	0.9949	1.0011
(b) RV5m										
$h = 1$	1.0029	1.0037	0.9979	1.0134	1.0181	1.0007	1.0054	1.0030	1.0048	1.0068
$h = 5$	1.0020	1.0048	0.9959	1.0044	1.0127	0.9998	0.9989	1.0133	1.0009	1.0045
$h = 22$	1.0025	1.0075	0.9827	1.0006	1.0094	1.0029	1.0046	1.0294	0.9987	1.0064
(c) RK										
$h = 1$	1.0067	1.0136	0.9921	0.9920	1.0187	1.0083	1.0016	1.0048	0.9940	1.0070
$h = 5$	1.0058	1.0074	0.9947	0.9954	1.0112	1.0046	0.9996	1.0167	0.9968	1.0053
$h = 22$	1.0049	1.0098	0.9786	0.9964	1.0081	1.0060	1.0095	1.0286	0.9968	1.0079

NOTE: Relative RMSE and relative MAE denote the RMSE and MAE ratios of realized BCSV model to realized LNSV model, respectively. We take squared returns, 5-min realized variance (RV5m), and realized kernel (RK) as possible volatility proxies. h denotes the step length or the forecast horizon.

predicting conditional volatilities (σ_t^2). For this purpose, the prediction values is calculated via the Kalman filter prediction and the method of volatility estimation in Section 3.3. We take both squared returns and realized volatility measures as possible volatility proxies. First, considering the squared returns as the volatility proxy, the MAE result shows that for SP500, DAX, and DJIA, the realized BCSV model performs better than the realized LNSV model, while the RMSE result shows that only for Nikkei the realized BCSV model performs better than the realized LNSV model. Second considering the realized variance (RV5m) as the volatility proxy, the MAE and RMSE results only support the realized BCSV model for DAX and SP500 being more preferable. Third, considering the realized kernel (RK) as the volatility proxy, the RMSE result prefers the realized BCSV model for all five stocks, while the MAE result only supports the realized BCSV model for DAX.

We also consider out-of-sample forecasting performance using the two metrics: relative RMSE and relative MAE, which are RMSE and MAE ratios of realized BCSV model to realized LNSV model, respectively. To achieve this goal, we use the sample starting from January 1, 2010, to May 15, 2012 to test forecasting performance. The results of out-of-sample forecast comparison are reported in Table 5 with the corresponding forecast horizon or step length (h) being 1, 5, and 22 days. Similarly, we can find that for some stock indices with certain volatility proxy, the realized BCSV model performs better than the realized LNSV model. For example, the RMSE and MAE results associated with returns as the volatility proxy for the FTSE and DAX show that the proposed realized BCSV model is preferable compared to the realized LNSV model; the RMSE results with the realized variance (RV5m) and the realized kernel (RK) as the volatility proxies show that the Nikkei prefers the realized BCSV model.

In short, the above results of in-sample forecasting and out-of-sample forecasting reveal that the realized BCSV model performs competitively in predicting the conditional volatility for some stock indices and volatility proxies compared to the realized LNSV model.

6. CONCLUDING REMARKS

In this article, we have introduced a class of realized stochastic volatility models based on the Box–Cox transformation which takes a more flexible structure than the logarithm transformation for the realized volatility measures. This class of realized BCSV model can be seen as an extension of the realized LNSV model proposed by Takahashi, Omori, and Watanabe (2009) and Koopman and Scharth (2013), and also an extension of the return-based BCSV model introduced by Yu, Yang, and Zhang (2006) and Zhang and King (2008). To estimate the realized BCSV model, we have further presented a two-step maximum likelihood estimation procedure based on the Kalman filter and the deletion smoothing algorithm.

The simulation results show good performance of the two-step maximum likelihood estimation for the realized BCSV model and reveal that model misspecification has significant impacts on parameter estimation and excessive skewness and kurtosis if the realized BCSV model is misspecified and estimated by the log transformation specification. The Monte Carlo evidence also suggest that the Box–Cox transformation specification would be preferable regardless of the true process. The empirical results with two realized volatility measures and daily returns for five stock indices show the feasibility of the realized BCSV model and demonstrate that the Box–Cox specification is more appropriate than the logarithm specification in modeling stochastic volatilities in terms of the in-sample results such as the values of log-likelihood, AIC and BIC. In addition, the results of sample forecasting performance show that the realized BCSV model performs competitively compared to the realized LNSV model in predicting the conditional volatility.

ACKNOWLEDGMENTS

The authors thank the Joint Editors, an Associate Editor and two anonymous referees, Marcel Scharth at the Department of Econometrics, VU University Amsterdam, and our colleague Ming Lin at the Wang Yanan Institute for Studies in Economics, Xiamen University, for helpful comments and suggestions. This research was supported by the National Natural Science Foundation of China (No. 71371160, 11101341), Program for New Century Excellent Talents in University (NCET-13-0509), and Program for New Century Excellent Talents in Fujian Province University (NCETFJ).

[Received February 2013. Revised April 2014.]

REFERENCES

Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002), "Range-Based Estimation of Stochastic Volatility Models," *The Journal of Finance*, 57, 1047–1091. [593]

Andersen, T. G., and Bollerslev, T. (1998), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885–905. [593]

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2010), "Parametric and Nonparametric Volatility Measurement," in *Handbook of Financial Econometrics*, eds. L. P. Hansen and Y. Aït-Sahalia, Amsterdam: North-Holland, pp. 67–138. [593]

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55. [593,594]

——— (2003), "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625. [594]

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012), "Jump-Robust Volatility Estimation Using Nearest Neighbor Truncation," *Journal of Econometrics*, 169, 75–93. [594]

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008), "Designing Realised Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536. [594,599]

Barndorff-Nielsen, O. E., and Shephard, N. (2001), "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society*, Series B, 63, 167–241. [593]

——— (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society*, Series B, 64, 253–280. [593,594]

——— (2004), "Power and Bipower Variation With Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–48. [594]

——— (2005), "How Accurate is the Asymptotic Approximation to the Distribution of Realised Volatility?," in *Identification and Inference for Econometric Models: A Festschrift for Thomas J. Rothenberg*, eds. D. W. K. Andrews and J. H. Stock, New York: Cambridge University Press, pp. 306–331. [593]

Box, G. E. P., and Cox, D. R. (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society*, Series B, 26, 211–252. [594,596]

Brandt, M., and Diebold, F. (2006), "A No-Arbitrage Approach to Range-Based Estimation of Return Covariance and Correlations," *Journal of Business*, 79, 61–74. [593]

Broto, C., and Ruiz, E. (2004), "Estimation Methods for Stochastic Volatility Models: A Survey," *Journal of Economic Surveys*, 18, 613–649. [593]

Carroll, R. J., and Ruppert, D. (1981), "Prediction and the Power Transformation Family," *Biometrika*, 68, 609–616. [598]

Chen, G. (1995), "Generalized Log-Normal Distributions With Reliability Application," *Computational Statistics & Data Analysis*, 19, 309–319. [596]

Chib, S., Nardari, F., and Shephard, N. (2002), "Markov Chain Monte Carlo Methods for Stochastic Volatility Models," *Journal of Econometrics*, 108, 281–316. [593]

Chou, R. (2005), "Forecasting Financial Volatilities With Extreme Values: The Conditional Autoregressive Range (CARR) Model," *Journal of Money, Credit and Banking*, 37, 561–582. [593]

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008), "The Volatility of Realized Volatility," *Econometric Reviews*, 27, 46–78. [603]

de Jong, P. (1989), "Smoothing and Interpolation With the State Space Model," *Journal of the American Statistical Association*, 84, 1085–1088. [597]

Engle, R. F., and Lee, G. J. (1999), "A Long-Run and Short-Run Component Model of Stock Return Volatility," in *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, eds. R. F. Engle and H. White, New York: Oxford University Press, pp. 475–497. [601]

Engler, E., and Nielsen, B. (2009), "The Empirical Process of Autoregressive Residuals," *Econometrics Journal*, 12, 367–381. [602]

Freeman, J., and Modarres, R. (2006), "Inverse Box-Cox: The Power Normal Distribution," *Statistics & Probability Letters*, 76, 764–772. [596,598]

Gonçalves, S., and Meddahi, N. (2009), "Bootstrapping Realized Volatility," *Econometrica*, 77, 283–306. [594]

——— (2011), "Box-Cox Transforms for Realized Volatility," *Journal of Econometrics*, 160, 129–144. [594]

Guerrero, V. M. (1993), "Time Series Analysis Supported by Power Transformations," *Journal of Forecasting*, 12, 37–48. [598]

Hansen, P. R., and Lunde, A. (2005), "A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data," *Journal of Financial Econometrics*, 3, 525–554. [593]

Harvey, A., Ruiz, E., and Shephard, N. (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247–264. [593]

Hasbrouck, J. (2007), *Empirical Market Microstructure: The Institutions, Economics and Econometrics of Securities Trading*, New York: Oxford University Press. [593]

Jacod, J., Li, Y., Mykland, P., Podolskij, M., and Vetter, M. (2009), "Microstructure Noise in the Continuous Case: The Pre-Averaging Approach," *Stochastic Processes and their Applications*, 119, 2249–2276. [594]

Jacquier, E., Polson, N., and Rossi, P. (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 12, 371–417. [593]

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994), *Continuous Univariate Distributions* (Vol. 1, 2nd ed.), New York: Wiley. [598]

Kim, S., Shephard, N., and Chib, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison With ARCH Models," *Review of Economic Studies*, 65, 361–393. [593]

Koopman, S. J., and Scharth, M. (2013), "The Analysis of Stochastic Volatility in the Presence of Daily Realised Measures," *Journal of Financial Econometrics*, 11, 76–115. [593,594,595,596,597,599,600,601,604]

McAleer, M., and Medeiros, M. C. (2008), "Realized Volatility: A Review," *Econometric Reviews*, 27, 10–45. [593]

Proietti, T., and Riani, M. (2009), "Transformations and Seasonal Adjustment," *Journal of Time Series Analysis*, 30, 47–69. [598]

Riani, M. (2004), "Robust Multivariate Transformations to Normality: Constructed Variables and Likelihood Ratio Tests," *Statistical Methods & Applications*, 13, 179–196. [596]

——— (2009), "Robust Transformations in Univariate and Multivariate Time Series," *Econometric Reviews*, 28, 262–278. [596]

Takahashi, M., Omori, Y., and Watanabe, T. (2009), "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously," *Computational Statistics & Data Analysis*, 53, 2404–2426. [593,594,595,596,600,604]

Taylor, J. M. (1985), "Power Transformations to Symmetry," *Biometrika*, 72, 145–152. [598]

——— (1986a), "The Retransformed Mean After a Fitted Power Transformation," *Journal of the American Statistical Association*, 81, 114–118. [598]

Taylor, S. J. (1986b), *Modeling Financial Time Series*, New York: Wiley. [593]

Yu, J. (2005), "On Leverage in a Stochastic Volatility Model," *Journal of Econometrics*, 127, 165–178. [596]

Yu, J., Yang, Z., and Zhang, X. (2006), "A Class of Nonlinear Stochastic Volatility Models and Its Implications for Pricing Currency Options," *Computational Statistics & Data Analysis*, 51, 2218–2231. [594,596,604]

Zhang, X., and King, M. L. (2008), "Box-Cox Stochastic Volatility Models With Heavy-Tails and Correlated Errors," *Journal of Empirical Finance*, 15, 549–566. [594,604]

Zhang, L., Mykland, P., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411. [594]