Fitting the Heston Stochastic Volatility Model to Chinese Stocks

Ahmet Goncu (Corresponding author)

Dept. of Mathematical Sciences, Xi'an Jiaotong Liverpool University

Renai Road 111, Suzhou, 215123, China

Tel: 86-512-8816-1521 E-mail: Ahmet.Goncu@xjtlu.edu.cn

Hao Yang

Dept. of Mathematical Sciences, Xi'an Jiaotong Liverpool University

Renai Road 111, Suzhou, 215123, China

Received: May 2, 2014 Accepted: June 16, 2014 Published: June 24, 2014

doi:10.5296/ifb.v1i1.5571 URL: http://dx.doi.org/10.5296/ifb.v1i1.5571

Abstract

In this article we investigate the goodness-of-fit of the Heston stochastic volatility model for the Shanghai composite index and five Chinese stocks from different industries with the highest trading volume. We have jointly estimated the parameters of the Heston stochastic volatility for the daily, weekly and monthly timescales model by employing a kernel density of the empirical returns to minimize the mean-squared deviations between the theoretical and empirical return distributions. We find that the Heston model is able to characterize the empirical distribution of Chinese stock returns at the daily, weekly and monthly timescales.

Keywords: Heston stochastic volatility model, Goodness-of-fit test, Chinese stocks, Kernel density

1. Introduction

Empirical distribution of stock returns are often characterized by leptokurtosis, which conflicts with the fundamental assumption of normality of log-returns in the Black–Scholes model. Empirical characteristics of stock returns are often better described by models that allow for fat tails and high peaks. The use of stochastic processes that allow for a wider range of shapes for return distributions has been considered in the literature, where it is also well documented that stochastic volatility models or models based on jump-diffusion and pure jump processes significantly improve the Black-Scholes framework. Among others, some examples of these models are provided by Merton (1976), Madan and Seneta (1990), Heston (1993), Barndorff-Nielsen (1997, 1998) and Kou (2002).

In this article we investigate the goodness-of-fit of the Heston (1993) model in the Chinese stock market, which can be characterized by frequent extreme returns and fatter tails compared to developed stock markets. In this respect, the dataset we consider also tests the robustness of the Heston model. To the best of our knowledge there has been no study that investigates the goodness-of-fit of the Heston stochastic volatility model for Chinese stocks. Furthermore, we propose the use of a non-parametric kernel density in the estimation of the Heston model, which reduces the estimation error of model parameters by smoothing the empirical density of log-returns.

The Heston model is widely used in modelling stock prices and the pricing of financial derivatives due to three major advantages:

1) The constant volatility assumption of the Black–Scholes model is not satisfied, and the implied volatilities of option prices exhibit volatility smiles. As the seminal work of Bakshi et al. (1997) has shown, stochastic volatility should be incorporated for pricing and internal consistency, and stochastic volatility modelling often yields the best performance for hedging;

2) The existence of closed-form option pricing formulas. In the Heston model, closed-form solutions for vanilla options are given by the fast Fourier transform method of Carr and Madan (1998), which leads to computationally efficient pricing;

3) The probability distribution for log-returns under the Heston model is given in closed form by Dragulescu and Yakovenko (2002), which leads to efficient estimation of model parameters from historical stock returns.

The goodness-of-fit of the Heston model to the historical data of log-returns has been tested in studies by Dragulescu and Yakovenko (2002), Silva and Yakovenko (2003) and Daniel et al. (2005), which together show mixed results regarding the performance of the Heston model. Dragulescu and Yakovenko (2002) have derived the closed form of the probability distribution of log-returns in the Heston model to show that the Heston model provides a good fit to the DowJones index returns at different time intervals. However, as Daniel et al. (2005) has pointed out, Dragulescu and Yakovenko (2002) trimmed the dataset by removing extreme price movements. Without trimming, Daniel et al. (2005) have shown that the Heston model does not provide a good fit to the Dow Jones stock market index. Meanwhile, Silva and Yakovenko (2003) have verified the goodness-of-fit of the Heston model by explaining the NASDAQ, DowJones and S&P 500 indices and by documenting different results for different dataset periods. Our results show that without trimming the dataset the Heston model fits the empirical distribution of index and stock returns in the Chinese stock market well, especially for the daily log-returns. Given the dataset's frequent extreme events, sometimes the optimization routine used in the parameter estimation may fail to converge and alternative initial values might be needed.

This article is organized as follows. In the next section we briefly present the Heston stochastic volatility model and the probability distribution of log-returns. In Section III we present the dataset, while in Section IV we discuss the parameter estimation via distance minimization. Section V presents the goodness-of-fit tests and finally Section VI offers our conclusions as well as recommendations for future work.

2. The Heston Model and the Probability Distribution of Log-Returns

In the Heston (1993) model, stock prices follow the stochastic differential equation

$$dS_t = \mu S_t dt + \sigma_t S_t dW_t^{(1)} \tag{1}$$

where μ is the drift, σ_t is the volatility and $W_t^{(1)}$ is a standard Wiener process. Log-returns are given by $r_t = \log\left(\frac{S_t}{S_0}\right)$. Following Dragulescu and Yakovenko (2002), centred log-returns is given by $x_t = r_t - \mu t$ while the dynamics of the centred log-returns is given by

$$dx_t = -\frac{\nu_t}{2}dt + \sqrt{\nu_t}\, dW_t^{(1)}. \tag{2}$$

This equation denotes the variance by $\nu_t = \sigma_t^2$, which obeys the following Ornstein-Uhlenbeck process

$$d\nu_t = -\gamma(\nu_t - \theta)dt + \kappa\sqrt{\nu_t}\, dW_t^{(2)}, \tag{3}$$

where θ is the long-term mean of ν, γ is the rate of mean reversion, $W_t^{(2)}$ is a standard Wiener process and κ is the variance noise (vol-vol) parameter. In general, the Wiener processes in Equations 1 and 3 may be correlated and can be written as

$$dW_t^{(2)} = \rho dW_t^{(1)} + \sqrt{1-\rho^2}\, dZ_t, \tag{4}$$

where Z_t is a Wiener process independent of $W_t^{(1)}$ and $\rho \in [-1,1]$ is the correlation coefficient.

Dragulescu and Yakovenko (2002) have solved the forward Kolmogorov equation that governs the time evolution of the joint probability $P_t(x, \nu \mid \nu_i)$ given the initial value of the variance v_i to obtain the following probability distribution of centred log-returns given a time lag t:

$$P_t(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dp_x e^{ip_x x + F_t(p_x)} \tag{5}$$

with

$$F_t(p_x) = \frac{\gamma\theta}{\kappa^2}\Gamma t - \frac{2\gamma\theta}{\kappa^2}\ln\left(\cosh\frac{\Omega t}{2} + \frac{\Omega^2 - \Gamma^2 + 2\gamma\Gamma}{2\gamma\Omega}\sinh\frac{\Omega t}{2}\right), \tag{6}$$

where $\Omega = \sqrt{\Gamma^2 + \kappa^2(p_x^2 - ip_x)}$, $\Gamma = \gamma + i\rho k p_x$ where ρ is the correlation coefficient between two Wiener processes $W_t^{(1)}$ and $W_t^{(2)}$ and γ, θ, κ and μ are the parameters of the Heston model.

3. Data

To verify the goodness-of-fit of the Heston model, we have utilised daily closing prices of the Shanghai composite index and five stocks representing different industries with the highest trading volume. Namely, we consider the stock prices of China Minsheng Banking Corp. (Banking, 600016), Sinopec Group (Petroleum and Oil, 600028), Jiangsu Sunshine Corp. (Clothing and manufacturing, 600220), Shanghai Tongji Science & Technology Industrial Corp. (Construction, 600846), Chengdu B-ray Media Corp. (Media, 600880). In Table 1 we supply descriptive statistics for the daily log-returns for both the composite index and five stocks with their Shanghai Stock Exchange codes.

Table 1. Descriptive statistics of the daily log-returns of the Chinese stocks and Shanghai composite index

	Mean	SD	Skewness	Kurtosis	Min.	Max.
Shanghai Index	0.00016	0.01556	-0.0812	7.75	-0.0926	0.0940
600016	-0.00020	0.02787	-3.4554	49.64	-0.4657	0.0962
600028	0.00016	0.02330	-0.1875	10.12	-0.2365	0.0967
600220	-0.00065	0.03443	-5.3728	100.35	-0.6747	0.0968
600846	-0.00039	0.03517	-1.8830	32.16	-0.5205	0.2801
600880	0.000033	0.03097	-4.7979	92.64	-0.6513	0.0957

Since for each stock the dates of the initial public offerings differ, we have therefore considered the longest available historical dataset for each stock. For the Shanghai composite index, we consider the period from 1 January 1998 to 6 May 2013. Daily, weekly and monthly log-returns are given by the following time series. Daily log-returns are calculated as

$R_t^{Daily} = \ln(S_{t+1}/S_t)$ for $t = 1, 2, ..., N$. Weekly log-returns we calculate $R_t^{Weekly} = \ln(S_{t+4}/S_t)$ for $t = 1, 6, 11, 16, ..., floor(N/5)$ and monthly log-returns are calculated as $R_t^{Monthly} = \ln(S_{t+21}/S_t)$ for $t = 1, 23, 55, ..., floor(N/22)$. No intersecting observations have been used in the above calculations, and in translation weekly and monthly returns are treated as stock price paths. The advantage of calculating the returns as above is discussed in Daniel et al. (2005).

Daniel et al. (2005) has pointed out that Dragulescu and Yakovenko (2002) trimmed the dataset by, for example, discarding observations that exceeded $\pm 4\%$ in daily returns. Since trimming the dataset reduces the fairness of the empirical analysis presented in Dragulescu and Yakovenko (2002), we do not trim the dataset by removing extreme values in the dataset.

4. Parameter Estimation

Parameter estimation can be done utilizing historical log-returns or options data or both. For example in the study by Eraker (2004) likelihood based inference is used to estimate stochastic volatility models jointly using options and stock price data. Although, option prices reflect information regarding the risk neutral density, in China there are no traded equity options in an organized exchange. Under the Heston model stock returns at different timescales are assumed to have the same parameters. Therefore, minimization of the distance between theoretical and empirical distributions is a suitable approach to fit a single set of parameters for different timescales jointly.

Empirical distribution of log-returns is often calculated by partitioning the space of log-returns. As provided by Dragulescu and Yakovenko (2002), the domain of log-returns can be partitioned into equally spaced bins of width Δr to allow simple counting of the number of log-returns belonging to each bin. Relative frequencies are then obtained by dividing these frequencies by the total sample size to yield the empirical distribution $P_t^*(x)$.

To obtain model parameters, Dragulescu and Yakovenko (2002) minimized the following objective function with respect to the parameters of the Heston model

$$\sum_{x,t} |\ln(P_t^*(x)) - \ln(P_t^{Heston}(x))|^2 \tag{7}$$

where the sum is taken over all available x and t. Since we work with a smaller dataset of log-returns, we have considered the time intervals $t = 1, 5, 22$ which correspond to daily, weekly and monthly log-returns. To improve the convergence of the optimization routine used, we smoothed the empirical distribution of log-returns and revised the objective function in Equation 7 as

$$\max_{\theta,\kappa,\sigma}\left\{\sum_{x,t} |K(P_t^*(x)) - P_t^{Heston}(x)|^2\right\} s.t. \quad \kappa > 0, \quad \sigma > 0, \tag{8}$$

where $K\left(P_t^*(x)\right)$ is the kernel density (Note 1) calculated from the empirical distribution of log-returns for each time interval $t = 1,5,22$. From Equation 8, a single set of parameters fitting the Heston model to daily, weekly and monthly log-returns jointly is obtained. Tables 2 and 3 supplies the estimated parameters of the Heston model for the daily and joint estimation at the daily, weekly and monthly intervals. In Figure 1 we plot the fitted Heston probability distribution function versus the empirical distribution for daily, weekly and monthly log-returns. Figure 1 thus shows that the Heston model fits the daily log-returns particularly well. In Figures 2, 3, and 4 we plot the fit of the Heston model for the highly traded stocks at the Shanghai Stock Exchange. From these figures it is clear that the Heston model can characterize the behaviour of empirical log-returns for at least the daily time interval.

Table 2. Estimated parameters for the Heston model for daily log-returns

Parameters	θ	κ	σ	ρ
Shanghai Index	0.0002515	0.0348	0.00467	-0.1202
600016	0.0004907	0.1705	0.01037	-0.0843
600028	0.0004577	0.7498	0.02803	0.19033
600220	0.0007333	24.916	0.79008	-0.0445
600846	0.0009454	23.545	0.79049	-0.0547
600880	0.0006919	0.1353	0.01412	0.1773

Table 3. Estimated parameters for the Heston model for the joint estimationat daily, weekly and monthly timescales

Parameters	θ	κ	σ	ρ
Shanghai Index	0.0002775	0.0883	0.00852	-0.0699
600016	0.000479	0.0752	0.00649	-0.1154
600028	0.0004212	0.2554	0.01278	0.27948
600220	0.0006060	0.1511	0.01045	-0.1249
600846	0.0006893	2.8013	0.03614	-0.1273
600880	0.0006892	0.1413	0.01441	0.17889

If Equation 8 is maximized for only the daily log-returns (i.e., t = 1), then all stocks considered yield convergence in the optimization routine. Reducing the number of time intervals in Equation 8 thus naturally improves the convergence and fit of the Heston model.

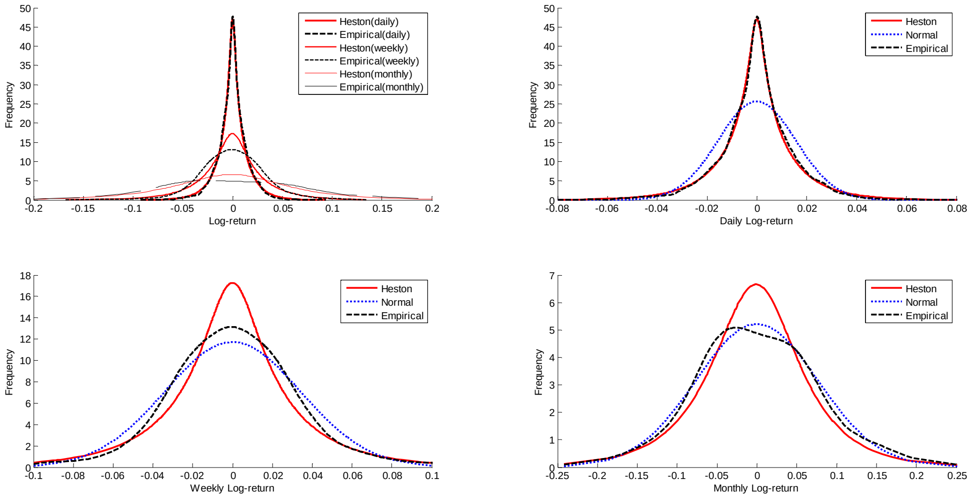

Figure 1. Fitted probability density function of the Heston model versus the empirical distribution of log-returns for the Shanghai composite index

Note. First subplot presents the jointly fitted Heston density for the daily, weekly, and monthly timescales, whereas other subplots shows the fit of the Heston model for daily, weekly and monthly returns separately.

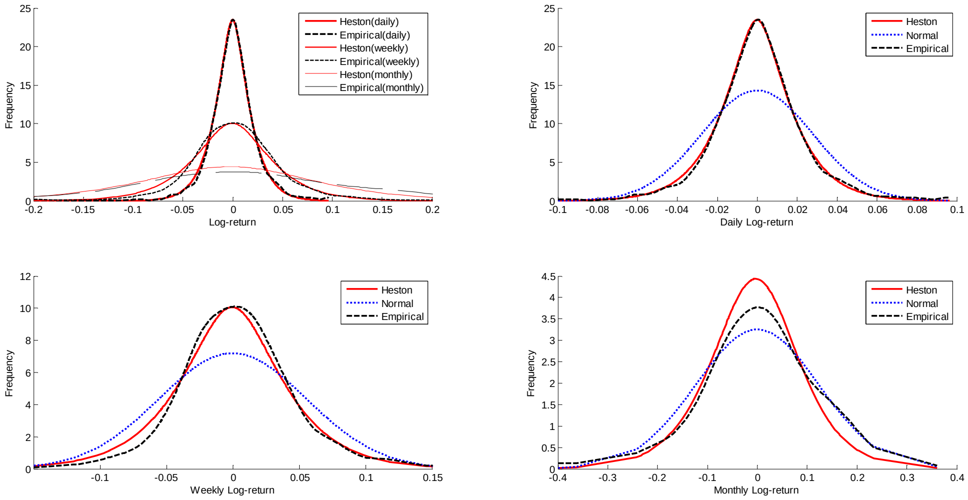

Figure 2. Fitted probability density function of the Heston model versus the empirical distribution of log-returns for stock 600016

Note. First subplot presents the jointly fitted Heston density for the daily, weekly, and monthly timescales, whereas other subplots shows the fit of the Heston model for daily, weekly and monthly returns separately.

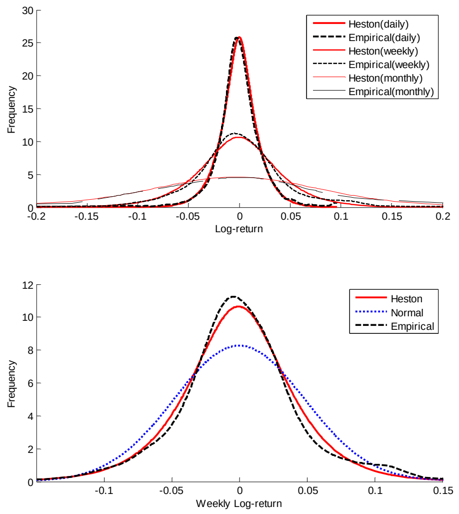

Figure 3. Fitted probability density function of the Heston model versus the empirical distribution of log-returns for stock 600028

Note. First subplot presents the jointly fitted Heston density for the daily, weekly, and monthly timescales, whereas other subplots shows the fit of the Heston model for daily, weekly and monthly returns separately.

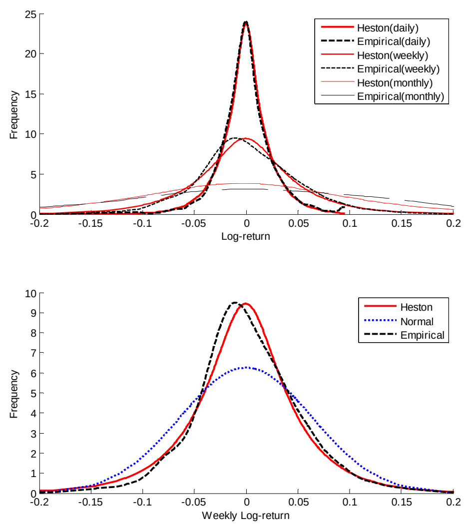

Figure 4. Fitted probability density function of the Heston model versus the empirical distribution of log-returns for stock 600880

Note. First subplot presents the jointly fitted Heston density for the daily, weekly, and monthly timescales, whereas other subplots shows the fit of the Heston model for daily, weekly and monthly returns separately.

5. Goodness-of-Fit Tests

Once the model parameters were estimated for the common set of daily, weekly and monthly log-returns, we tested the goodness-of-fit of the Heston model using via three statistical tests. In this section we comment briefly on the implemented tests and present our results.

The chi-square goodness-of-fit test is a discrete goodness-of-fit test in which the range of observations is divided into *k* bins. For this test we have considered the bins

$$\left(-\infty,-0.025,-0.01,-0.005,-0.001,0.001,0.005,0.01,0.025,\infty\right),$$

$$\begin{matrix}(-\infty,-0.05,-0.035,-0.025,-0.0175,-0.01,-0.005,-0.0015,\\ 0.0015,0.005,0.01,0.0175,0.025,0.035,0.05,\infty)\end{matrix},$$

$$\left(-\infty,-0.12,-0.08,-0.05,-0.02,-0.01,0.01,0.02,0.05,0.08,0.12,\infty\right),$$

for the daily, weekly and monthly time intervals, respectively. The above intervals were chosen so that the observed frequencies in each bin would be similar. The degree of freedom of the chi-square test equals $dof = \#bins - 1 - m$, where m is the number of parameters of the model being tested. The corresponding critical values at a 95% confidence level for the chi-square goodness-of-fit test of the Heston model at daily, weekly and monthly time intervals are given as 11.07, 19.67 and 14.07, respectively. As the benchmark case for the normal distribution, the critical values are given as 14.07, 21.03 and 15.51, for daily, weekly and monthly time intervals, respectively.

The Anderson–Darling (AD) (1952) goodness-of-fit test provides a good measure of distance between empirical and theoretical densities. Therefore, a smaller test suggests a better fit to the data, whereas the Kolmogorov–Smirnov (KS) test measures the maximal discrepancy between the expected and observed cumulative distributions of log-returns. To calculate the KS statistic, we used the empirical and theoretical cumulative distribution functions and computed the maximum discrepancy between them. Since the estimators and goodness-of-fit test statistics were calculated from the same dataset, both AD and KS test statistics are not sufficient to accept the tested model. Chi-square test does not present this problem, thus critical values are available to test the null hypothesis.

In Tables 4, 5 and 6 we present the goodness-of-fit test results for the daily, weekly and monthly log-returns, respectively. Normal distribution, which was used as a benchmark scenario, is rejected by the chi-square goodness-of-fit test for all stocks and the composite index.

Table 4. Chi-square, Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit test results for daily log-returns, where critical values at a 95% confidence level for the chi-square goodness-of-fit test of the Heston model and normal distribution are given as 11.07 and 14.07, respectively

	Chi-square		Anderson–Darling		Kolmogorov–Smirnov	
	Normal	Heston	Normal	Heston	Normal	Heston
Shanghai Index	407.60*	2.80	62.45	4.93	0.0791	0.0488
600016	466.93*	0.83	71.60	3.03	0.0974	0.0280
600028	297.50*	0.88	43.45	2.16	0.0803	0.0290
600220	704.31*	1.86	47.25	7.01	0.1138	0.0312
600846	590.48*	21.77*	52.98	19.15	0.0910	0.0254
600880	752.08*	1.66	94.89	1.90	0.0993	0.0187

Note. *Rejected at the 95% confidence level.

Table 4 shows that the Heston model cannot be rejected at a 95% confidence level for the daily log-returns, whereas the normal distribution can clearly be rejected. AD statistics indicate that the distance between the empirical and theoretical distributions is small. In Table 5, we present results for the weekly log-returns, for which the Heston model also provides a good fit except for the Shanghai composite index. Table 6 also shows that the Heston model has a good fit for monthly log-returns. Except for monthly log-returns, normal distribution can be consistently rejected for all stocks and the composite index.

Table 5. Chi-square, Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit test results for weekly log-returns, where critical values at a 95% confidence level for the chi-square goodness-of-fit test of the Heston model and normal distribution are given as 19.67 and 21.03, respectively

	Chi-square		Anderson–Darling		Kolmogorov–Smirnov	
	Normal	Heston	Normal	Heston	Normal	Heston
Shanghai Index	46.61*	31.77*	4.89	1.84	0.0561	0.0355
600016	97.56*	2.27	13.04	2.34	0.1111	0.0591
600028	54.95*	4.58	6.64	1.80	0.0792	0.0521
600220	162.95*	4.25	21.87	2.96	0.1320	0.0614
600846	140.09*	16.40	19.42	6.10	0.0955	0.0638
600880	116.60*	2.69	15.92	2.82	0.1091	0.0538

Note. *Rejected at the 95% confidence level.

Table 6. Chi-square, Anderson–Darling and Kolmogorov–Smirnov goodness-of-fit test results for monthly log-returns, where critical values at a 95% confidence level for the chi-square goodness-of-fit test of the Heston model and normal distribution are given as 14.07 and 15.51, respectively

	Chi-square		Anderson–Darling		Kolmogorov–Smirnov	
	Normal	Heston	Normal	Heston	Normal	Heston
Shanghai Index	12.73	6.64	0.41	0.90	0.0429	0.0609
600016	8.11	3.03	0.99	1.04	0.0775	0.0721
600028	13.87	0.90	1.49	1.09	0.0688	0.0724
600220	17.99	0.64	2.05	1.46	0.0979	0.0573
600846	14.59	1.21	2.00	3.23	0.0859	0.0544
600880	7.99	8.14	1.17	1.82	0.0697	0.0788

Note. *Rejected at the 95% confidence level.

6. Conclusion and Future Work

We investigated the goodness-of-fit of the Heston stochastic volatility model to the empirical distribution of stock and index returns in the Chinese stock market. This article shows that the Heston model provides a good fit for Chinese stocks, especially for their daily log-returns. It should be noted that we fit a single set of model parameters that is able to fit to the daily, weekly and monthly log-returns simultaneously. Overall, the goodness-of-fit test statistics provide evidence that the Heston model cannot be rejected, especially for daily log-returns. However, one drawback of using the Heston model is the difficulty in the convergence of parameter estimation, which may be attributed to frequent extreme movements in Chinese stock returns. To improve the parameter estimation optimization, we smoothed the empirical distribution of log-returns via a kernel density.

Future work should focus on what these results imply for risk management. Since the Heston model can capture the heavy-tailed empirical distributions of Chinese stocks, it might also perform well in estimating quantiles and risk measures, such as the value-at-risk.

References

Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain ‘goodness-of-fit’ criteria based on stochastic processes. *Annals of Mathematical Statistics*, *23*, 193–212. http://dx.doi.org/10.1214/aoms/1177729437

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, *52*, 2003–2049. http://dx.doi.org/10.1111/j.1540-6261.1997.tb02749.x

Barndorff-Nielsen, O. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, *24*, 1–13. http://dx.doi.org/10.1111/1467-9469.t01-1-00045

Barndorff-Nielsen, O. (1998). Processes of normal inverse Gaussian type. *Finance and*

Stochastics, *2*, 41–68. http://dx.doi.org/10.1007/s007800050032

Carr, P., & Madan, D. B. (1998). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, *2*, 61–73.

Daniel, G., Joseph, N. L., & Bree, D. S. (2005). Stochastic volatility and the goodness-of-fit of the Heston model. *Quantitative Finance*, *5*, 199–211. http://dx.doi.org/10.1080/14697680500148521

Dragulescu, A. A., & Yakovenko, V. M. (2002). Probability distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance*, *2*, 443–453. http://dx.doi.org/10.1080/14697688.2002.0000011

Eraker, B. (2004). Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *Journal of Finance*, *59*, 1367–1404. http://dx.doi.org/10.1111/j.1540-6261.2004.00666.x

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, *6*, 327–343. http://dx.doi.org/10.1093/rfs/6.2.327

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, *48*, 1086–1101. http://dx.doi.org/10.1287/mnsc.48.8.1086.166

Madan, D. B., & Seneta, E. (1990). On modelling and pricing weather derivatives. *The Journal of Business*, *63*, 511–524. http://dx.doi.org/10.1086/296519

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*, 125–144. http://dx.doi.org/10.1016/0304-405X(76)90022-2

Silva, A. C., & Yakovenko, V. M. (2003). Comparison between the probability distribution of returns in the Heston model and empirical data for stock indexes. *Physica A*, *324*, 303–310. http://dx.doi.org/10.1016/S0378-4371(02)01903-9

Note

Note 1. Kernel density smoothes the empirical distribution of log-returns; the MATLAB function 'ksdensity(.)' is implemented. Once a smoothed empirical probability distribution is obtained, we maximize the objective function in Equation 8 by employing the constrained maximization function 'fmincons(.)' in MATLAB.

Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).