Herd behavior and volatility in financial markets[†]

Beum-Jo Park[1]

[1]Department of Economics, Dankook University

Abstract

Relaxing an unrealistic assumption of a representative percolation model, this paper demonstrates that herd behavior leads to a high increase in volatility but not trading volume, in contrast with information flows that give rise to increases in both volatility and trading volume. Although detecting herd behavior has posed a great challenge due to its empirical difficulty, this paper proposes a new methodology for detecting trading days with herding. Furthermore, this paper suggests a herd-behavior-stochastic-volatility model, which accounts for herding in financial markets. Strong evidence in favor of the model specification over the standard stochastic volatility models is based on empirical application with high frequency data in the Korean equity market, strongly supporting the intuition that herd behavior causes excess volatility. In addition, this research indicates that strong persistence in volatility, which is a prevalent feature in financial markets, is likely attributed to herd behavior rather than news.

Keywords: Herd-behavior-stochastic-volatility model, Markov Chain Monte Carlo algorithm, realized bipower variation, realized volatility, spline regression.

1. Introduction

High volatility is a prevailing property in financial markets. Its source has been a central question in financial literature. One can reasonably infer that return volatility may be induced by trades related to the arrival of information in markets (as early influential work on this issue, see Clark, 1973; Copeland, 1976; Jennings *et al.* 1981; Admati and Pfleiderer, 1988, among others). The information-flow paradigm, however, cannot provide insight into exactly why we frequently observe high volatility in financial markets even in the absence of any significant information or news including macroeconomic announcements. Therefore, many of the extant studies have tried to explain the phenomenon using the models of herd behavior that may be sufficient to induce high volatility without significant information.

In particular, recent studies in the field of econophysics have focused on microscopic dynamic models to explain herd behavior systemically. The most representative microscopic dynamic model is percolation model (Stauffer and Sornette, 1999; Cont and Bouchaud, 2000, etc.), which is based on the hypothesis of information cascades due to imperfect information. It is supposed that a trader is in an information cascade situation if, after imitating others, he makes an investment decision independent of his private information. According to the hypothesis, the information cascade is likely to arise quickly and can be shattered by even a small amount of public information. Thus, idiosyncratically, herd behavior makes financial markets unstable and can sometimes result in a panic situation (Eguíluz and Zimmermann, 2000).

Although herd behavior in financial markets has been relatively well documented, there have been few empirical studies on the influence of herding on volatility associated with trading volumes. The main reason for this may be the technical difficulty of detecting herd behavior. Thus, this study suggests a method for detecting herd behavior which makes financial markets turbulent even in the absence of news and increases volatility of returns excessively. This study also proposes a herd-behavior-stochastic-volatility model that allows us to investigate the dynamic relationship between herd behavior, volatility, and trading volumes. Consequently, this study contributes to the literature on herd behavior as well as volatility in several ways. First, this study focuses on a new finding that, even in the absence of news, financial markets can often have excessive volatility that, in general, is negatively related to trading volumes. Obviously, this finding cannot be explained by information-flow models, which induce a positive relationship. Although it has been overlooked by researchers so far, it is quite crucial because it can be easily observed in financial markets. Thus, this study theoretically clarifies the cause of the divergence from the expected relationship using the extended percolation model under a more realistic assumption. Second, this study presents a method for detecting days with herd behavior based on the theoretical result and the concepts of realized volatility and realized bipower variation, lately developed by Andersen *et al.* (2003) and Barndorff-Nielsen and Shephard (2004) respectively. Third, using the detection method we define an indicator function as 1 at a day with herding or 0 otherwise. Then, we can specify a herd-behavior- stochastic-volatility model whose volatility equation is switched by the indicator function. Of course, standard maximum likelihood methods cannot be directly applied to estimate the herd-behavior-stochastic-volatility model because the unobserved volatility and the parameter vector have to be estimated simultaneously and further the observation equation is nonlinear in the state variable. Therefore, the Markov chain Monte Carlo (MCMC) method (Jacquier *et al.* 1994) is an appropriate estimation method for the model. This study estimates the herd-behavior-stochastic-volatility model using high-frequency data on the stock market of South Korea, which is well known to exhibit severe herding (Chang *et al.* 2000 and Kim *et al.* 2004).

The outline of the rest of this paper is as follows. Section 2 explains percolation models associated with herd behavior and extends it more realistically. For the practical implementations of the results, Section 3 develops the herd-behavior- stochastic-volatility model with special emphasis on the ease of detecting herd behavior. Using data on trading volume and daily realized volatility obtained from high frequency data, Section 4 provides some empirical evidence on the effect of herding on conditional volatility persistence, analyzes the relationship between volatility and trading volume, and derives the feasibility of the model as compared with other stochastic models. The final section concludes with brief suggestions

for future research.

2. Herd behavior and volatility in the Cont-Bouchaud model

This study expands the Cont-Bouchaud model, which is representative of one such perco-lation models, to include the examination of the effect of herd behavior on both volatility and trading volume. It is assumed that there are N agents in a cluster, which is a group of agents acting together in a stock market, and the $i^{th}(1 \leq i \leq N)$ agent trades in an asset whose price at time t is p_t. The state of agent i is represented by $\xi^i = +1, -1, 0$ corre-sponding to buying $\xi^i = +1$, selling $\xi^i = -1$, or waiting $\xi^i = 0$. Agents can be isolated or connected through links that form a cluster sharing the same information, so the network of links evolves dynamically. All agents are inactive in the beginning stage ($\xi^i = 0$ for all i), but they are initiated through a chain reaction of other linked agents. That is, at time t, an agent starts to either buy the asset or sell the asset with probability $((a^+, a^-))$:

$$P(\xi^i = +1) = a^+, P(\xi^i = -1) = a^- \tag{2.1}$$

Then, agents in the same cluster imitate this action instantly and make a connection; several networks can be established. The aggregate state of the cluster is given by

$$S_t = \sum_{i=1}^{N} \xi_t^i \tag{2.2}$$

where $S_t > 0$ means excess demand in markets raises price and $S_t < 0$ means excess supply in markets decreases price. The size of the cluster has an effect on volatility. After trading, the cluster is broken up into isolated agents. Once the cluster collapses, all of agents belonging to this cluster are inactive with probability $P(\xi^i = 0) = 1 - a(a = a^+ + a^-)$ and this repeated process produces herd behavior (refer to Eguíluz and Zimmermann, 2000). In this percolation model the price evolves over time so that the following equation is considered as the simple update rule for the price:

$$p_{t+1} = p_t + \frac{1}{\lambda}\sum_{i=1}^{N} \xi_t^i \tag{2.3}$$

where λ is the excess demand needed to move the price by one unit and is a parameter measuring the sensitivity of price to fluctuations in excess demand. When considering the cluster size, the equation can be rewritten by price change:

$$\triangle p = p_{t+1} - p_t = \frac{1}{\lambda}\sum_{\chi=1}^{k} \omega^\chi \xi_t^\chi = \frac{1}{\lambda}\sum_{\chi=1}^{n_c} \zeta_t^\chi \tag{2.4}$$

where ω^χ is the size of cluster χ, ξ_t^χ is the individual demand of agents belonging to the cluster χ, n_c is the number of clusters, and ζ_t^χ is equal to $\omega^\chi \xi_t^\chi$. Then, following Cont and Bouchaud (2000), the probability density function of price change is

$$f(\Delta p = \nu) \underset{|\nu| \to \infty}{\sim} \frac{1}{|\nu|^{5/2}} e^{\nu/\nu_0} \tag{2.5}$$

where ν_0 denotes ν at the origin and the kurtosis of the probability density function is given by (Cont and Bouchaud, 2000)

$$k(\Delta p) = \frac{2c + 1}{\aleph(1 - c/2)A(c)(1 - c)^3} \qquad (2.6)$$

where c is a parameter that indicates the level of clustering between agents, $A(c)$ is a normalization constant with a value close to 1, and $\aleph$ is the average number of orders received during a given period. Thus, as $\aleph$ is smaller, excess kurtosis becomes larger, resulting in higher fluctuations of returns. It is well known that herd behavior in financial markets gives rise to high volatility of returns (Lux, 1995; Lux and Marchesi, 1999; Abreu and Brunnermeier, 2003; Chari and Kehoe, 2004, among others). According to Cont and Bouchaud (2000), it is assumed that λ is constant and c is a fixed external parameter so that the probability $(p_\nu = c/N)$ that any pair of agents i and j are linked together is also constant. Hence, the volatility depends on parameter a. For example, for $a \to 1$ each agent makes his investment decision individually. Since isolated agents trade independent of the other investment decision, large clusters are not built and herd behavior is not produced. In contrast, for small $a << 1$, large proportion of agents do not trade in financial markets and information is spread out over the agents. Thus, information asymmetry in the markets prevents agents from making investment decisions independently. This causes an increase in internal connectivity, which leads to large clusters and herd behavior. Therefore, according to this view, the herding parameter can be defined as

$$H = 1/a - 1 \qquad (2.7)$$

While no herding occurs for $a = 1$ $(H = 0)$, herding is produced for $a < 1$ $(H > 0)$.

In the Cont-Bouchaud model, probability (a) for buying and selling the asset is assumed to be fixed at a given value. However, Stauffer and Sornette (1999) shows that price changes do not change the number of traders, but rather, price changes alter traders' investment behavior. Further, Lux (1995), Lux and Marchesi (1999) argue that although price volatility in markets does not influence the number of traders, it transforms traders from fundamentalists to noise traders who are likely to regard the behavior of other traders as a source of information. This gives rise to a tendency towards herd formation. That is, high volatility tends to make it difficult for traders to invest in assets independently and instead react according to herd behavior.

The new idea proposed in this context is that the probability (a) is not fixed. That is, it is assumed that the probability evolves with reflection of previous price changes over time:

$$a_t = a_{t-1} + \gamma_1|\ln p_t - \ln p_{t-1}| - \gamma_2(\ln p_t - \ln p_{t-1})^2 \qquad (2.8)$$

where $0 < \gamma_1 < 1$, $0 < \gamma_2 < 1$ are parameters, presenting that while increase in small price changes leads to increase of a_t within the information-flow paradigm, increase in large price changes leads to decrease of a_t. The latter intuition, reflecting on the studies of Lux (1995), Lux and Marchesi (1999) and others, is also quite consistent with the research lines of Michaely and Vila (1996) and Park (2010), which show that trading volume is negatively associated with high volatility that is considered as risk or uncertainty. Further, it is supported by the empirical finding from Figure 4.2 that shows the nonlinear relationship

between volatility and trading volume (i.e. these variables have a positive relationship at low volatile periods but their relationship changes from positive to negative as trading volume become larger). The probability with which an agent buys or sells the asset in a given period is a nonlinear function of past volatility. Hence, herding parameter also evolves with time.

$$H_t = 1/a_t - 1 \tag{2.9}$$

Since at time t the probability that an agent trade independently is a_t during a given period, the average number of orders received is given by

$$\aleph_t \cong a_t N \tag{2.10}$$

That is, the decrease of a_t gives rise to the increase in H_t but the decrease of $\aleph_t$. As a result, during the period of high volatility, the probability of independent trading is likely to be low and the possibility of herding is likely to be high, leading to the small number of orders. Consequently, the market becomes more volatile. Of course, high volatility resulting from herd behavior is disappeared by flow of public information or news, which causes the clusters' collapse. Then, agents are also isolated again.

3. Accounting for herd behavior in stochastic volatility models

3.1. Realized volatility and realized bipower variation

Consider a simple diffusion process for the log of a price ($\wp_t$) with instantaneous volatility, σ_t^2. Conditional on the sample path $\{\sigma_{t+\tau}^2\}_0^1$, a natural measure of σ_t^2 is the integral of the instantaneous variances over the day t, $\bar{\sigma}_t^2 = \int_0^1 \sigma_{t+\tau}^2 d\tau$. Under general conditions, an unbiased estimator of the integrated variance, known as realized volatility or variation (RV), is obtained by summing intraday squared returns over many small intervals $(1/\triangle)$ within the day (Barndorff-Nielsen and Shephard, 2002; Andersen *et al.* 2003):

$$RV_{t+1}(\triangle) \equiv \sum_{j=1}^{1/\triangle} r_{t+j\cdot\triangle,\triangle}^2 \tag{3.1}$$

where $r_{t,\triangle}$ is the discretely sampled $\triangle$-period returns, $r_{t,\triangle} \equiv \wp_t - \wp_{t-\triangle}$ and without loss of generality it is assumed that $1/\triangle$ is an integer. Under weak regularity conditions, realized volatility converges in probability to the increment of the quadratic variation of the diffusion process over the day as $\triangle$ goes to zero (Andersen *et al.* 2003; Maheu and McCurdy, 2007). That is,

$$RV_{t+1}(\triangle) \to \int_t^{t+1} \sigma^2(s)ds + \sum_{t<s\leq t+1} k^2(s) \tag{3.2}$$

where k refers to the size of discrete jumps. Therefore, without jumps realized volatility is consistent for the integrated volatility.

On the other hand, lately Barndorff-Nielsen and Shephard (2004) define the standardized realized bipower variation, which provides a consistent estimator of the integrated volatility

unaffected by jumps, as

$$BV_{t+1}(\triangle) \equiv \mu_1^{-2} \sum_{j=2}^{1/\triangle} |r_{t+j\cdot\triangle,\triangle}||r_{t+(j-1)\cdot\triangle,\triangle}| \tag{3.3}$$

where $\mu_1 \equiv \sqrt{2/\pi} = E(|Z|)$ is the mean of the absolute value of standard normally distributed random variable Z. Thus, for $\triangle \to 0$,

$$BV_{t+1}(\triangle) \to \int_t^{t+1} \sigma^2(s)ds \tag{3.4}$$

Evidently for $\triangle \to 0$ $RV_{t+1}(\triangle) - BV_{t+1}(\triangle)$ is a consistent estimator of the pure jump contribution to realized volatility. To guarantee positive estimates, Barndorff-Nielsen and Shephard suggest truncating the actual empirical measurements at zero:

$$J_{t+1}(\triangle) \equiv \max[RV_{t+1}(\triangle) - BV_{t+1}(\triangle), 0] \tag{3.5}$$

Based on this idea, a test statistic for jumps can be derived (Andersen *et al.* 2007). Especially, with regards to applying the delta rule to the joint bivariate distribution, Huang and Tauchen (2005) show that the following statistic is closely approximated by a standard normal distribution.

$$\Im_{t+1}(\triangle) \equiv \triangle^{-1/2} \frac{[RV_{t+1}(\triangle) - BV_{t+1}(\triangle)]/RV_{t+1}(\triangle)}{[(\mu_1^{-4} + 2\mu_1^{-2} - 5)\max\{1, TQ_{t+1}(\triangle)BV_{t+1}(\triangle)^{-2}\}]^{1/2}} \tag{3.6}$$

where $TQ_{t+1}(\triangle)$ is the standardized realized tripower quarticity:

$$TQ_{t+1}(\triangle) \equiv \triangle^{-1}\mu_{4/3}^{-3} \sum_{j=3}^{1/\triangle} |r_{t+j\cdot\triangle,\triangle}|^{4/3}|r_{t+(j-1)\cdot\triangle,\triangle}|^{4/3}|r_{t+(j-2)\cdot\triangle,\triangle}|^{4/3} \tag{3.7}$$

and $\mu_{4/3} \equiv 2^{2/3}\Gamma(7/6)\Gamma(1/2)^{-1} = E(|Z|^{4/3})$. Thus, we can detect whether the significant jumps (or news) happen, and identify how big the size of jumps is by the statistic:

$$J_{t+1,\wp}(\triangle) \equiv I[\Im_{t+1}(\triangle) > \Phi_\wp] \cdot [RV_{t+1}(\triangle) - BV_{t+1}(\triangle)] \tag{3.8}$$

where $\Phi_\wp$ is the critical value of standard normal distribution at a significance level $\wp$ and $I[\cdot]$ is the indicator function.

3.2. A method for detecting herd behavior

Generally, since the arrival of new information induces agents to trade assets and change prices, it raises both volatility and trading volume, thereby leading to a positive relationship between them (refer to the Mixture of Distribution Hypothesis; Tauchen and Pitts, 1983). As explained earlier, however, herd behavior caused by noise traders causes a divergence from this relationship. Even in the absence of news or significant information, it increases

volatility but decreases trading volume. Due to this anomaly, we can regard time periods without news at which volatility is higher than conditionally predicted variation given trading volume as an exhibition of herd behavior.

This approach intuitively allows for the derivation of the following process for detecting days with herd behavior. First, calculate realized volatility at each day. Second, estimate the relationship between volatility and trading volume through nonparametric estimation methods (e.g., spline function; Rice and Rosenblatt, 1983), reflecting their nonlinear relationship. Based on the estimated relationship, realized volatility is predicted at a given trading volume. Third, detect days with news by the jump test, described in previous section. Fourth, identify days during which news is not detected but realized volatility is higher than predicted one in the second step, as the exhibition of herd behavior.

3.3. Herd-behavior-stochastic-volatility model and an estimation method

Herd-behavior-stochastic-volatility model specification is inspired by the idea that the state of volatility in the presence of herding should be different from that of volatility without herding. Thus, to incorporate herd behavior with the stochastic volatility model, which can consider volatility dynamically, we classify volatility as belonging to one of two states of markets - 'with herding' and 'without herding', using appropriately defined indicator function θ_t : $\theta_t = 1$ with herding and $\theta_t = 0$ otherwise. θ_t is determined by the method for detecting herd behavior, which is described in Section 3.2.Thus, the herd-behavior-stochastic-volatility model has the following representation:

$$r_t = \exp(h_t/2)\varepsilon_t$$
$$h_t = (\mu_1 + \phi_1(h_{t-1} - \mu_1) + \beta_1 V_t)\theta_t + (\mu_2 + \phi_2(h_{t-1} - \mu_2) + \beta_2 V_t)(1 - \theta_t) + \eta_t \quad (3.9)$$

where r_t is returns, $h_t = \ln \sigma_t^2$, and V_t is trading volumes. $\varepsilon_t \sim iid\ N(0,1)$ and $\eta_t \sim iid\ N(0, \sigma_\eta^2)$ are independent and the persistent parameter is assumed to satisfy $|\phi| < 1$, implying that h_t is stable.

It is difficult to estimate the herd-behavior-stochastic-volatility model by standard maximum likelihood methods because the unobserved volatility vector as well as the parameter vector have to be estimated simultaneously and further the observation equation is nonlinear in the state variable. Indeed, from the herd-behavior-stochastic-volatility model, the log volatility vector is defined as $h_t = (h_1, \cdots, h_\Gamma)$ and the parameter vector is defined as $\Psi = (\mu_\gamma, \mu_1, \mu_2, \phi_1, \phi_2, \beta_1, \beta_2, \sigma_\eta^2)$. Therefore, the likelihood function of the model is the conditional density of data y :

$$L(\Psi) = f(y|\Psi) = \int f(y, h|\Psi)dh$$
$$= \int f(y|h, \Psi)f(h|\Psi)dh \quad (3.10)$$

Denoting $Y_{t-1} = (y_1, \cdots, y_{t-1})$, the density of the data can be represented as a mixture

over the log volatility vector using law of total probability:

$$
\begin{aligned}
f(y|\Psi) &= \prod_{t=1}^{T} f(y_t|Y_{t-1}, \Psi) \\
&= \prod_{t=1}^{T} \int f(y_t|h_t, Y_{t-1}, \Psi) f(h_t|Y_{t-1}, \Psi) dh
\end{aligned}
\tag{3.11}
$$

where this likelihood function is intractable because the density function $f(h_t|Y_{t-1}, \Psi)$ has no closed form so that $y_t|Y_{t-1}$ cannot be analytically expressed. Due to the reasons, Jacquier *et al.* (1994) suggested, the Markov Chain Monte Carlo (MCMC) method that is more efficient than a quasi-maximum likelihood method, and Shephard and Pitt (1997), Kim *et al.* (1998), among others, have developed the MCMC method. According to the results of simulation for estimating stochastic models, the MCMC method is superior to the quasi-maximum likelihood method or the GMM method in terms of sampling properties (Jacquier *et al.* 1994). Much more detail for the MCMC method is available in the articles by Broto and Ruiz (2004) and Asai *et al.* (2006) among others.

4. Empirical evidence

4.1. Data and preliminary statistics

In this section, we present an empirical application to illustrate the plausibility of the herd-behavior-stochastic-volatility model. The data employed in this application are daily and high-frequency KOSPI and trading volumes (the number of trading stocks) in the stock market of South Korea, covering the period between January 2, 2004 and February 29, 2008. The sample contains a total of 1029 observation days. The span of the sampling period has high volatility due to 'subprime mortgage crisis,' which has caused panic in financial markets, and other several pieces of big news. For representative news, we can take China shock (April 29, 2004), triggered by Chinese premier Wen Jiabao's comments on cooling down the overheating Chinese economy and dealing a sharp blow to the Korean financial markets, Announcement that North Korean produced a nuclear weapon (February 11, 2005; the announcement was actually made February 10, a holiday.), and North Korea nuclear test (October 9, 2006).

The high-frequency data used to measure realized volatility are five-minute observations of spot markets, which are provided by the Korea Exchange (KRX, http://sm.krx.co.kr), because five-minute observations are close to optimal sampling intervals derived by previous studies (Andersen *et al.* 2001; Bandi and Russell, 2005). The use of a five-minute frequency, corresponding to 60 intraday observations ($1/\triangle = 60$), means that the total data used in this study were obtained from 61,740 observations. It is known that realized volatility suffers from a bias problem resulting from market microstructure noise, causing autocorrelation in the intraday returns (Hansen and Lunde, 2006). To remove this autocorrelation and resolve the problem, filtering techniques such as a moving average (MA) filter have been used by Andersen *et al.* (2001), Maheu and McCurdy (2007), and others. Hence, using the method in section 3.1, realized volatility estimates are calculated by summing squares of MA(1)-filtered

intraday returns.

$$RV_{t+1} \equiv \sum_{j=1}^{60} r_{t+j/60}^2 \tag{4.1}$$

To avoid problems arising from the non-stationary behavior usually observed in stock prices, we take the natural logarithmic differences between two successive trading days. The first panel of Figure 4.1 shows the movements of returns, r_t, which supports the view that the return series is highly dynamic. The return series also tends to be clustered together over time. The second panel of Figure 4.1 shows the movement of absolute return residuals $|\hat{\varepsilon}_t|$, which are widely used to estimate daily volatility, and the third panel of Figure 4.1 shows the turbulent movement of realized volatility. As explained above, this tendency may indicate herd behavior in the stock market. Although their movements seem to be similar, absolute return residuals might have more variation than realized volatility except when measured at several extreme values. Absolute return residuals are obtained from estimating the following regression model:

$$r_t = c + \sum_{i=1}^{2} a_i r_{t-i} + \sum_{j=1}^{4} b_j D_{jt} + \varepsilon_t \tag{4.2}$$

where c is a constant, D_{jt} values' are day-of-the-week dummies used to capture differences in mean returns, and r_{t-i} values' are lagged returns. Based on Schwarz's information criterion (SIC), the two lag length is chosen to control for serial dependence in returns.
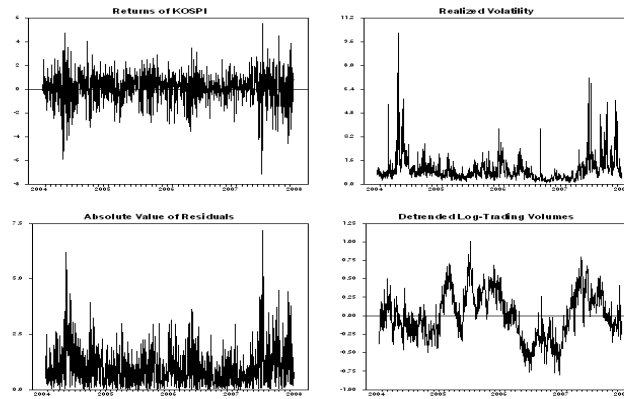

Figure 4.1 KOSPI returns, absolute return residuals, RV, and DTLV

Table 4.1 reports the preliminary statistics for returns, trading volume, absolute return residuals, and realized volatility. The coefficients of skewness and kurtosis for realized volatility and trading volumes support the view that each distribution is not normal, coinciding with the other empirical findings. In particular, the high kurtosis is attributed to large outliers that should be strongly associated with not only news but also herding in the market. Jarque-Bera (JB) statistics for the series also reject normality at the conventional 5-percent

level. Therefore, the series is commonly transformed with a logarithm in the existing literature.

We turn to the joint test of serial correlation. The Ljung-Box Q statistics are computed up to the fifth lags. Under the null hypothesis with no serial correlation, such statistics have an asymptotic chi-square distribution with five degrees of freedom. The Ljung-Box Q statistics indicate a high serial correlation in the series except returns. In particular, the high correlation in the realized volatility series indicates that herd behavior may be serially correlated over time because the volatility is closely associated with herd behavior, and partially supports the reliability of Equation (2.8).

As in previous literature (e.g., Fleming, 2006; Park, 2007), the raw trading volume is transformed by a natural logarithm to stabilize the variability of the trading volume series and more reduce the non-normality of its distribution that is found by JB test statistic in Table 4.1. We test for the stationarity of the log-trading volume and the realized volatility series using the augmented Dickey-Fuller (1979) and Phillips-Perron (1988) tests of the null of unit root against the stationary series. The results, provided in Table 4.2, indicate that both the log-trading volume and the realized volatility series are clearly stationary. The stationarity nature of these series allows us to apply traditional economic models to empirical works.

Table 4.1 Summary statistics for r_t, $|\hat{\varepsilon}_t|$, realized volatility, and trading volume

Statistics	r_t	Trading volume	Realized Volatility	$|\hat{\varepsilon}_t|$
Mean	0.0728	368600635	1.0215	0.9689
Median	0.1499	359542950	0.7468	0.7045
Maximum	5.5223	967839370	10.1767	7.195
Minimum	-7.1825	136329089	0.082	0.0032
Variance	1.7582	1.45E+16	0.9099	0.8094
Skewness(Sk=0)	-0.5615(0.0000)	0.6649(0.0000)	3.6523(0.0000)	1.7998(0.0000)
Kurtosis(Ku=0)	2.2368(0.0000)	0.5411(0.0004)	19.9037(0.0000)	5.0748(0.0000)
JB	268.35(0.0000)	88.373(0.0000)	19273.04(0.0000)	27.4030(0.0000)
Q(5)	8.3916(0.0782)	3310.81(0.0000)	1475.99(0.0000)	147.207(0.0000)
Q2(5)	204.569(0.0000)	2851.03(0.0000)	687.582(0.0000)	196.212(0.0000)

JB is the Jarque-Bera normality test, and Q (M) and Q2 (M) are the Ljung-Box Q statistics at lag M for series and squared series. P-values are in parentheses.

Table 4.2 Unit root tests on the stationarity of log-trading volumes and realized volatility

	M	Dickey-Fuller (M)	Phillips-Perron (M)
Log-Trading Volumes	5	-29.150(0.00)	-83.881(0.00)
	10	-18.920(0.00)	-100.940(0.00)
Realized Volatility	5	-84.520(0.00)	-335.221(0.00)
	10	-80.724(0.00)	-437.080(0.00)

P-values are in parentheses and M is the number of lags.

To filter out the trend and, among others, a day-of-the-week effect, we estimate a regression of the log-trading volumes (LV_t) on a constant, a linear trend (T), and a quadratic trend (T^2), using four day-of-the-week dummies (D_{jt}) for Monday through Thursday by OLS estimation method. The fourth panel of Figure 4.1 shows the detrended log-trading volumes, which will be used in our empirical study.

4.2. Detecting trading days with herd behavior

Trading days with herd behavior in the South Korean stock market are detected by the method described in the section 3.2. The calculation of realized volatility belonging to the first step previously completed. For the next step, a cubic smoothing spline function as a nonparametric method is used to estimate the relationship between volatility and trading volume. Figure 4.2 graphs the estimated spline function. The noteworthy estimation result is that while the variables have a positive relationship at low volatile periods, their relationship changes from positive to negative as trading volumes become larger. Intuitively, this phenomenon might be largely due to herd behavior which induces their negative relationship.
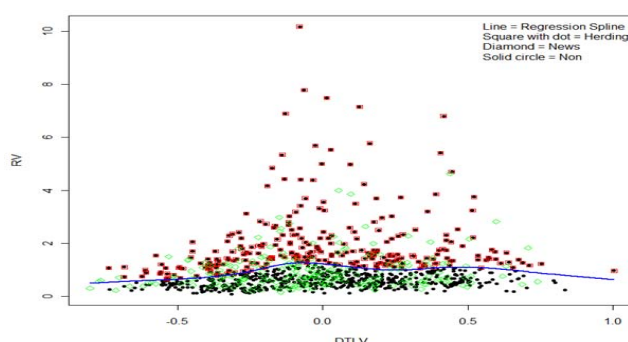

Figure 4.2 Regression spline and trading days with herding
(RV = realized volatility, DTLV = detrended log-trading volumes)

From the results of the third step, Figure 4.3 was drawn. The second panel shows the jump components defined in Equation (3.5), and the third panel shows the significant jumps estimated by Equation (3.8) corresponding to $a = 0.99$. From the Figure 4.3, we find that the most of the days with significant jumps have high volatility, which is consistent with the result of Andersen *et al.* (2007). A more interesting finding from the figure is that there are days that have no the significant jumps but much higher volatility. For example, on Jan. 22, 2008, the jump component is only 0.3780 while realized volatility is quite high, i.e., 5.6829. Obviously, this empirical finding cannot be explained by the information-flow paradigm in which market volatility is attributed to news (or public information), and therefore, this finding supports the possibility of herd behavior in the stock market. In the final step, we select days that have no news (i.e. no significant jumps) but have higher volatility than estimated volatility using the cubic smoothing spline function in the second step. As a consequence of this process, the selected days are considered as the days with herd behavior. These days are depicted by red dots in Figure 4.2. According to the results, 22 percent of total trading days include significant jumps and 26 percent of trading days without significant jumps exhibit herd behavior. The high probability of herding stands in existing literatures on the herd behavior such as Chang *et al.* (2000) and Kim *et al.* (2004) among others.
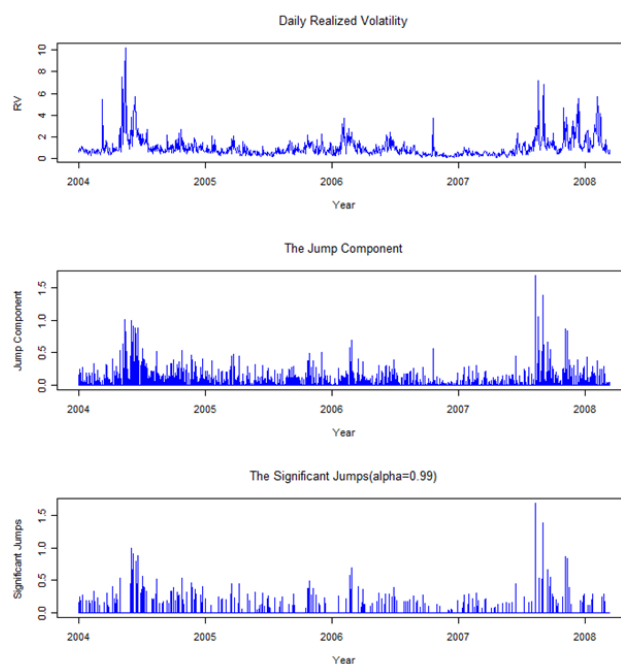
Figure 4.3 The jump component and the significant jumps

4.3. Estimation of the herd-behavior-stochastic-volatility model

Given our emphasis on herd behavior, we now turn to estimation of the herd-behavior-stochastic-volatility model using KOSPI returns and trading volumes, and compare the estimation result with other models:

Model 1 (SV):

$$r_t = a_0 + a_1 r_{t-1} + a_2 r_{t-2} + \exp(h_t/2)\varepsilon_t$$
$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t \quad (4.3)$$

Model 2 (SVJ):

$$r_t = a_0 + a_1 r_{t-1} + a_2 r_{t-2} + \exp(h_t/2)\varepsilon_t$$
$$h_t = \mu + \phi(h_{t-1} - \mu) + \lambda_t q_t + \eta_t \quad (4.4)$$

Model 3 (HBSV):

$$r_t = a_0 + a_1 r_{t-1} + a_2 r_{t-2} + a_3\theta_t + a_4 JD_t + \exp(h_t/2)\varepsilon_t$$
$$h_t = (\mu_1 + \phi_1(h_{t-1} - \mu_1) + \beta_1 V_t)\theta_t + (\mu_2 + \phi_2(h_{t-1} - \mu_2) + \beta_2 V_t + \delta J_t)(1 - \theta_t) + \eta_t \quad (4.5)$$

The discrete-time representation of SV is the approximation of the continuous-time one. Thus, to maintain consistency with the continuous-time diffusion, including the terms for previous returns should not be valid. From a statistical point of view, however, the general discrete-time representation of SV do have serious disadvantage due to unrealistic assumption of no autocorrelation in returns. Even in KOSPI returns used in this paper, there is

significant autocorrelation. With the statistical problems, the advantage of consistency has diminished considerably. Hence, sharing the same spirit as in Harvey *et al.* (1994) or Shephard (1996), it is assumed that returns can be decomposed into two components: shift and scale components. So, unlike Equation (3.9), the terms for constant and previous returns are included for shift component in the empirical models.

Model 1 is a simple stochastic volatility model and Model 2 is a stochastic volatility model including a jump variable whose coefficient is changed over time. In this model q_t is a Bernoulli random variable that takes 1 with probability ζ and 0 with probability 1 - ζ and the size of jumps is denoted by $\lambda_t \sim N(m, \nu)$, where m and ν are random variables with normal and gamma distributions respectively. Model 3 is the herd-behavior-stochastic-volatility model whose stochastic volatility equation is switched corresponding to states of herd behavior, and it incorporates a jump dummy (JD_t) in mean equation, which is 1 with a jump component and 0 without a jump component. This dummy is obtained from days with significant jumps in Section 3.1. In Model 3, J_t denotes jumps estimated by test statistics for jumps in Equation (3.8).

The estimation results of Model 1 and Model 2 are reported in Table 4.3 and those of Model 3 are reported in Table 4.4. Each table contains the mean, standard deviation, 95% confidence interval of the posterior distribution for parameters, and deviance information criterion (DIC). For the posterior computation in the stochastic volatility models, WinBUGS is used because it is well known that WinBUGS provides an efficient implementation of the MCMC algorithm. WinBUGS uses Gibbs sampling (Geman and Geman, 1984) and the Metropolis method (Metropolis *et al.* 1953) to generate a Markov chain by sampling from conditional distributions. The posterior quantities are computed from 10,000 draws of the MCMC algorithm, collected after an initial burn-in period of 1,000 iterations. Note that although it is known that 5,000 draws are large enough to get accurate estimation using an efficient Metropolis algorithm (e.g., Chib *et al.* 2002), this paper takes 10,000 draws for optimal implementation in terms of speed and accuracy.

To estimate the herd-behavior-stochastic-volatility model via the MCMC algorithm, we should specify suitable prior distributions on unknown parameter vector Ψ and assume each parameter to be independent. For instance, a prior density function of $\Psi = (\mu_\gamma, \mu, \phi, \sigma_\eta^2)$ for Model 1 is

$$f(\Psi) = f(\mu_\gamma)f(\mu)f(\phi)f(\sigma_\eta^2) \tag{4.6}$$

In this empirical study, prior distributions on Ψ basically follow from the specification of Chib *et al.* (2002) or Yu and Meyer (2006). Thus, the following prior distributions are adopted: $a_i \sim N(0, 100), i = 0, \cdots, 4$, $\mu_i \sim N(0, 100), i = 1, 2$, $\beta_i \sim N(0, 100), i = 1, 2$, $\delta \sim N(0, 100)$, $\phi_i^* \sim B(20, 1.5)$, $i = 1, 2$, where $\phi_i^* = (\phi_i + 1)/2$, $\lambda_t \sim N(m, \nu)$, where $m \sim N(0, 100), \nu \sim \Gamma(2.5, 0.025)$.

As expected, the estimation results of Model 1 show that the parameter of volatility persistence ϕ is highly credible and its value, 0.9690, close to one confirms a strong daily persistence in volatility. The estimation results of Model 2 with jump components also support the strong persistence, in accordance with typical estimates reported in the literature. Note that the probability ζ is quite low and jumps are negatively related to volatility. In both models, since the estimates for α_1, α_2, and μ are close to 0 in terms of 95% CI and not credible at all, the parameters can be negligible in the models.

Table 4.3 Estimation results of Model 1 (SV) and Model 2 (SVJ)

Parameter	Model 1			Model 2		
	Mean	SD	95% CI	Mean	SD	95% CI
α_0	0.1564	0.0356	(0.0851 0.2268)	0.1635	0.0352	(0.0941 0.2329)
α_1	0.0339	0.032	(-0.0282 0.0968)	0.0272	0.0316	(-0.0345 0.0902)
α_2	-0.0338	0.0318	(-0.0969 0.0285)	-0.0267	0.0322	(-0.0914 0.0355)
μ	0.2816	0.1918	(-0.1285 0.6474)	0.2891	0.1806	(-0.0686 0.6524)
ϕ	0.9690	0.0122	(0.9431 0.9894)	0.9650	0.0137	(0.9304 0.9856)
ζ				0.0017	0.0012	(0.0002 0.0049)
m				-14.920	0.0536	(-15.050 -14.830)
σ_η	0.1705	0.0311	(0.1156 0.2268)	0.1938	0.0327	(0.1543 0.2908)
DIC	3277			3270		

'95% CI' denotes 95% confidence interval and DIC stands for deviance information criterion.

The empirical evidence on herd behavior from Figure 4.2 motivates us to specify Model 3 that appears to be an improvement over the other models. From the estimates for important parameters (ϕ, β, δ) in the stochastic volatility equation, we can derive several interesting findings: First, when we compute $\exp(\mu\ /2)$, interpreted as mode volatility, its values are 1.9251 with herding and 0.8080 without herding, respectively. This means that the daily volatility remarkably increases under herding phenomenon. The empirical result is consistent with many previous studies which argue that herding causes market unstability (e.g., Lux and Marchesi, 1999; Sornette and Malevergne, 2001). Furthermore, whereas the daily volatility is still persistent, it is significantly reduced by the indicator function for herd behavior, i.e., $\phi_1 = 0.8322$, $\phi_2 = 0.6230$, and this reduction is relatively prominent on days without herding. In these results, there is quite important point that needs to be addressed. The volatility persistence that is closely associated with herding and high volatility does not persist in the absence of herding. Second, the estimates of parameter for trading volume are $\beta_1 = -0.1088$, $\beta_2 = 0.0336$, implying that while the relationship between volatility and trading volume is positive in the absence of herding, the relationship is negative in the presence of herding. This is consistent with the theoretical intuition from the modified Cont-Bouchaud model, that herd behavior gives rise to high volatility but low trading volume in contrast with the flows of information because herding reduces heterogeneity in agents, leading to reduction of the probability of independent trading. However, since the estimates are not highly credible, the finding has little empirical implication. Third, the parameter δ for jumps, which presents the effect of news on volatility separately from herd behavior, is 1.4020 and is highly credible. This result shares the same spirit as the information paradigm that news is likely to increase the magnitude of change in price, i.e., volatility.

To highlight the adequacy of the herd-behavior-stochastic-volatility model, we can compute the deviance information criterion (DIC) (Berg *et al.* 2004) as a model comparison method due to its advantage in increasingly complex statistical models because it combines a Bayesian measure of fit with a measure of model complexity. Although the Akaike information criterion (AIC) is a general method for comparing alternative models, it is hard to apply AIC for stochastic volatility models. The reason is that unlike a nonhierarchical Bayesian model with only parameter vector Ψ, the number of parameters exceeds the number of observations in stochastic volatility models where the parameter vector should be augmented by an unobserved volatility vector $h = (h_1, \cdots, h_\mathrm{T})$.

Therefore, this study adopts DIC that is an extension of AIC to complex hierarchical

Table 4.4 Estimation results of Model 3 (HBSV)

Parameter		Model 3		
		Mean	SD	95% CI
	α_0	0.2525	0.0432	(0.1674 0.3387)
	α_1	4.3660	1.0180	(2.3210 6.0210)
	α_2	-4.3570	1.0200	(-6.0120 -2.3090)
	α_3	-0.6178	0.0995	(-0.8123 -0.4212)
	α_4	-0.0228	0.0882	(-0.1988 0.1488)
μ	μ_1	1.3100	0.6640	(-1.1450 2.1170)
	μ_2	-0.4263	0.1773	(-0.6936 0.1053)
ϕ	ϕ_1	0.8322	0.0769	(0.6919 0.9888)
	ϕ_2	0.6230	0.0917	(0.3383 0.7175)
β	β_1	-0.1088	0.1385	(-0.3617 0.1621)
	β_2	0.0336	0.0865	(-0.1153 0.2484)
	δ	1.4020	0.1494	(1.0310 1.6410)
	σ_η	0.2752	0.0521	(0.1959 0.4013)
DIC		3235		

'95% CI' denotes 95% confidence interval and DIC stands for deviance information criterion.

models. DIC provides an efficient approach to identify the most appropriate model with the smallest value like AIC. In Tables 4.3 and 4.4, estimated values of DIC for the models are reported. Model 3 with herd behavior has a remarkably smaller value of DIC than the other models, and thereby, it appears to be the most appropriate model according to DIC. The implication taken from the model comparison is that the results empirically support the validity of both the method of detecting herd behavior and the specification of the herd-behavior-stochastic-volatility model.

5. Concluding remarks

Extending the percolation model, developed by Cont and Bouchaud (2000), this paper has derived an interesting result that the relationship between volatility and trading volume can be negative in the presence of herding, which cannot be explained by the models within information-flow paradigm. Based on the result and the concepts of realized volatility and realized bipower variation, this paper has presented a method for detecting days with herd behavior. Furthermore, from a practical perspective, the paper has specified the herd-behavior-stochastic- volatility model and suggested the Markov chain Monte Carlo (MCMC) method to estimate the model efficiently.

To detect herd behavior and its effect on both volatility persistence and the relationship between volatility and trading volume in empirical work, we have used five-minute frequency data on KOSPI in the stock market of South Korea, covering the period between January 2, 2004 and February 29, 2008. The estimation results worth mentioning are as follows: First, a significant proportion of trading days (i.e., 26 percent) exhibit herd behavior. That is, they have higher volatility than estimated one given trading volume but no significant jumps. Second, the daily volatility remarkably increases under herding behavior. Third, the volatility persistence is closely associated with herding and high volatility does not persist in the absence of herding. Fourth, the relationship between volatility and trading volume is negative in the presence of herding even if it is not highly credible. Finally, the herd-

behavior-stochastic-volatility model provides a substantial improvement over the standard stochastic volatility models, as is evident from the estimated values of DIC.

One of the most important next steps in this research stream is to develop a more sophisticated test statistic for detecting herd behavior. Further, we might also consider a stochastic volatility representation for herding as a more reliable model that accounts for the asymmetric impact of herding on price fluctuation.

References

Abreu, D. and Brunnermeier, M. K. (2003). Bubbles and crashes. *Econometrica*, **71**, 173-204.

Admati, A. R. and Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies,* **1**, 3-40.

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics,* **89**, 701-720.

Andersen, T. G., Bollerslev, T. Diebold, F. X. and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics* 61, 43-76.

Andersen, T. G., Bollerslev, T. Diebold, F. X. and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica,* **71**, 579-625.

Asai, M., McAeer, M. and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews,* **25**, 145-175.

Bandi, F. M. and Russell, J. R. (2005). *Microstructure noise, realized volatility and optimal sampling*, Working Paper, University of Chicago.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics,* **17**, 457-478.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics,* **2**, 1-37.

Berg, A., Meyer, R. and Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics,* **22**, 107-120.

Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys,* **18**, 613-649.

Chang, E., Cheng, J. and Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance,* **24**, 1651-1679.

Chari, V. V. and Kehoe, P. (2004). Financial crises as herds: Overturning the critiques. *Journal of Economic Theory,* **119**, 128-150.

Chib, S., Nardari, F. and Shephard, N. (2002). Markov Chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics,* **108**, 281-316.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative process. *Econometrica,* **41**, 135-155.

Copeland, T. E. (1976). A model of asset trading under the assumption of sequential information arrival. *Journal of Finance,* **31**, 1149-1168.

Cont, R. and Bouchaud, J.-P. (2000). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics,* **4**, 170-195.

Dickey, D. and Fuller, W. A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association,* **74**, 427-431.

Eguíluz, V. M. and Zimmermann, M. G. (2000). Transmission of information and herd behavior: An application to financial markets. *Physical Review Letters,* **85**, 5659-5662.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **6**, 721-741.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics,* **24**, 127-161.

Harvey, A. C., Ruiz, E. and Shephard, N. G. (1994). Multivariate stochastic variance models. *Review of Economic Studies,* **61**, 247-264.

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics,* **3**, 456-499.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics,* **12**, 371-417.

Jennings, R. H., Starks, L. T. and Fellingham, J. C. (1981). An equilibrium model of asset trading with sequential information arrival. *Journal of Finance,* **36**, pp. 143-161.

Kim, K., Yoon, S.-M. and Kim, Y. (2004). Herd behaviors in the stock and foreign exchange markets. *Physica A*, **341**, 526-532.

Kim, S. and Kim, J. (2009). Analyzing financial time series data using the GARCH model. *Journal of the Korean Data & Information Science Society,* **20**, 475-483.

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies,* **65**, 361-393.

Lux, T. (1995). Herd behavior, bubbles, and crashes. *Economic Journal,* **105**, 881-896.

Lux, T. and Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature,* **397**, 498-500.

Maheu, J. and McCurdy, T. (2007). *Modeling foreign exchange rates with jumps*, University of Toronto, Department of Economics, Working Papers.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics,* **21**, 1087-1092.

Michaely, R. and Vila, J. (1996). Trading volume with private valuation: Evidence from the ex-dividend day. *The Review of Financial Studies*, **9**, 471-509.

Park, B.-J. (2009). Risk-return relationship in equity markets: Using a robust GMM estimator for GARCH-M models. *Quantitative Finance,* **9**, 93-104.

Park, B.-J. (2010). Surprising information, the MDH, and the relationship between volatility and trading volume. *Journal of Financial Markets,* **13**, 344-366.

Park, B.-J. (2011). Study on time-varying herd behavior in individual stocks. *Journal of the Korean Data & Information Science Society,* **22**, 423-436.

Park, I., Kwon, O. and Kim, T. (2009). KOSPI directivity forecasting by time series model. *Journal of the Korean Data & Information Science Society,* **20**, 991-998.

Phillips, P. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrica,* **75**, 335-346.

Rice, J. and Rosenblatt, M. (1983). Smoothing splines: Regression, derivatives and deconvolution. *Annals of Statistics,* **11**, 141-156.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In *Time Series Models in Econometrics, Finance and Other Fields*, edited by Cox, D. R., Hinkley, D. V. and Barndorff-Nielsen, O. E., Chapman & Hall, London, 1-67.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika,* **84**, 653-667.

Shim, J. and Lee, J. (2010). Estimation of nonlinear GARCH-M model. *Journal of the Korean Data & Information Science Society,* **21**, 831-839.

Stauffer, D. and Sornette, D. (1999). Self-organized percolation model for stock market fluctuations. *Physica A*, **271**, 496-506.

Tauchen, G. E. and Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica*, **5**, 485-550.

Yu, J. and Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometrics Reviews,* **25**, 361-384.