Provided for non-commercial research and education use. Not for reproduction, distribution or commercial use.



**This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.**

**Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.**

**In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:**

**<http://www.elsevier.com/copyright>**

Physica A 387 (2008) 6605–6612

Contents lists available at ScienceDirect

# Physica A

journal homepage: www.elsevier.com/locate/physa

We study two weight-driven information spreading models for financial market. In these models, we find that the activity threshold below which the 'financial crash' occurs can be increased by uneven distribution of information weight, compared with Eguíluz and Zimmermann model [V.M. Eguíluz, M.G. Zimmermann, Phys. Rev. Lett. 85 (2000) 5659]. We also find that below the threshold the normalized return distribution, *P*(*Z*; ∆*t*) satisfies *P*(*Z* = 0;  $\Delta t$ ) ∼ exp( $-\Delta t/b$ ) whereas *P*(*Z* = 0;  $\Delta t$ ) ∼  $\Delta t^{-\tau}$  above the threshold. Here  $\Delta t$  is the time interval where the normalized return is defined,  $Z(t, \Delta t) = Z(t + \Delta t) - Z(t)$ . By approximating the relative increase of  $P(Z; \Delta t = 1)$  for large *Z* as Gaussian distribution with non-zero mean, we show that the non-zero mean of the Gaussian distribution can

# Herd behavior in weight-driven information spreading models for financial market

# Soon-Hyung Yook<sup>∗</sup> , Yup Kim

*Department of Physics and Research Institute for Basic Sciences, Kyung Hee University, Seoul 130-701, Republic of Korea*

A R S T R A C T

#### a r t i c l e i n f o

*Article history:* Received 11 July 2008 Available online 8 August 2008

*PACS:* 89.65.-s 89.65.Gh 87.23.Ge 02.50.Le

*Keywords:* Econophysics

# **1. Introduction**

# The analysis of financial systems using standard methods developed in physics has a long tradition [1] and has recently been one of the active research areas in physics [2]. Much of the research interest of physicists has been mainly focused on the analysis of stock markets [2,3] and foreign exchange markets [4] due to the large amount of accessible data. Among those empirical studies, the most remarkable finding is that many different markets share universal properties. For example, the fat-tailed distribution of returns [2,3,5], long-term volatility correlation [2,6,7] and herding behavior [8,9] have been observed in many different markets [3–5,10–14]. The existence of such universal nature in many different markets is striking and suggests that those markets should be governed by the similar underlying mechanisms. In order to investigate the universal phenomena observed in many real markets, many microscopic models such as percolation model [8] and Isinglike spin models [15] have been developed.

cause such exponentially decaying behavior of  $P(Z = 0; \Delta t)$ .

Among those studies, Eguíluz and Zimmermann (EZ) recently proposed an interesting model to investigate the relationship between the transmission of information and herding behavior [9]. In EZ model, groups of agents are dynamically formed by random dispersion of information. The agents in the same group make the same decision for trading activity which cause the herding behavior. EZ showed that when the information dispersion is slower than the trading activity the return distribution follows a power-law. On the other hand, if the information dispersion is much faster than the trading activity then the relative increase in the distribution of extremely high return is observed. This relative increase of return distribution is known to be related to the *financial crashes* [9,13,14]. As shall be seen in Section 2 we indeed find that the relative increase in the return distribution is observed during the 9.11 crash. This implies that the information dispersion rate plays a very important role in the market dynamics.

Although the EZ model succeeded to explain many interesting features of the financial market, there are still many important factors which are not reflected in the model. In particular, the ''*value*'' or ''*weight*'' of information that each agent

∗ Corresponding author. Tel.: +82 2 961 0253; fax: +82 2 957 8408.

*E-mail addresses:* syook@khu.ac.kr (S.-H. Yook), ykim@khu.ac.kr (Y. Kim).

0378-4371/\$ – see front matter © 2008 Elsevier B.V. All rights reserved. doi:10.1016/j.physa.2008.08.009





© 2008 Elsevier B.V. All rights reserved.

6606 *S.-H. Yook, Y. Kim / Physica A 387 (2008) 6605–6612*



**Fig. 1.** (a) Change of S&P 500 index during the 9.11 crash period. The circle denotes September 11. (b) *R*(*t*, ∆*t* = 5 min) of S&P 500 index for the same period.

has should be different from agent to agent in real markets. A certain piece of information is more important than the others. Moreover, the value of information can be changed with time by combining with other piece of information. The cooperation between agents or individuals should be affected by the value of the information to maximize their profit. In order to investigate the effects of different weights of information, we assume that the profit is proportional to the weight of information for simplicity. Based on this assumption, we introduce two weight-driven information spreading models. One has time independent weight of information. The other one has dynamically changing weight as a result of synergetic cooperation among the agents. From the numerical simulations, we find that the financial crash can occur with higher activity rate compared to EZ model. We also suggest a novel criterion to determine activity threshold below which the financial crash occurs by analyzing the return distribution.

This paper is organized as follows. In Section 2 we provide empirical measurement of return distribution during financial crash. In Section 3 two information spreading models are introduced. And the simulation results are given Section 4. Summary and discussions are presented in the last section.

#### **2. S&P500 index**

The financial crashes in stock markets are usually defined by the striking drops of index or price of all stocks [16]. One of recent market crashes has been reported in September 11, 2001 which is known as 9.11 crash. Fig. 1(a) shows the 5-min change of Standards and Poor's (S&P) 500 index from August 1 to October 30 in 2001. The circle indicates the change of index on September 11. The data shows the index starts to decrease from the end of August and reaches the minimum around September 20. Then the rally period begins and the index becomes stable after October 10. For a time series *p*(*t*) of price or index value, the logarithmic return (or simply return) over integer time interval ∆*t*, *R*(*t*, ∆*t*), and the normalized return *Z*( $t$ ,  $\Delta t$ ) are defined as

$$
R(t, \Delta t) = \ln[p(t)] - \ln[p(t - \Delta t)], 0
$$
\n(1)

and

$$
Z(t, \Delta t) = \frac{R(t, \Delta t) - \langle R(t, \Delta t) \rangle}{\sigma},
$$
\n(2)

respectively. Here  $\sigma$  is the volatility,  $\sigma = (\langle R(t,\Delta t)^2 \rangle - \langle R(t,\Delta t)\rangle^2)^{1/2}$  and  $\langle \cdots \rangle$  denotes the average over time. The logarithmic return in the same period, *R*(*t*, ∆*t* = 5 min), is also displayed in Fig. 1(b). As shown in the data, the amplitude of *R*(*t*,  $\Delta t$  = 5 min) suddenly increases from the end of August (Fig. 1(b)). Although the financial crash is sometimes observed when the return distribution follows the Gaussian distribution [14], several studies have pointed out the possibility that the return distribution becomes asymmetric when the crash occurs [13] and deviates from the Lévy distribution [13,14]. From the data in Fig. 1 we find a very interesting result. Fig. 2 shows *P*(|*Z*|) measured from the data in Fig. 1. The obtained *P*(|*Z*|) has two distinctive regimes. When |*Z*| is small (|*Z*| < 0.05), *P*(|*Z*|) monotonically decreases. If we approximate the data to a power-law distribution in this regime, then we obtain *P*(|*Z*|) ∼ |*Z*| −1.7 (dashed line in Fig. 2). In the large |*Z*| regime *P*(|*Z*|) shows an explicit bump which can be approximated by the Gaussian distribution with a non-zero mean (solid line in Fig. 2). Thus, we expect that *P*(|*Z*|) during the 9.11 crash can be regarded as a combination of monotonically decreasing power-lawlike distribution and Gaussian distribution which has non-zero mean. (The detailed analysis for this kind of financial crash *S.-H. Yook, Y. Kim / Physica A 387 (2008) 6605–6612* 6607



**Fig. 2.** (color on line) Plot of *P*(|*Z*|) of S&P 500 index from August to September 2001. The dashed line represents *P*(|*Z*|) ∼ |*Z*| −1.7 and the sold line denotes Gaussian distribution with a non-zero mean.

shall be given in Section 4.2.) Our measurement of  $P(|Z|)$  confirms the expectation suggested by EZ [9] for the first time. This provides an evidence that the information dispersion rate becomes very important to understand many properties of financial markets during the financial crash.

# **3. Models**

In our models, it is assumed that the number of agents who want to cooperate with a certain agent *i* (''carrying capacity'' of agent *i*) is proportional to the weight of the information of the agent *i*, w*<sup>i</sup>* . Quantifying the value of information or finding its distribution is not trivial. In EZ model, each agent has the same weight of information [9]. Therefore, the information spreading of EZ model is described by a Gaussian process in which the fluctuation becomes irrelevant. However, there is no reason that the weight of all information should be the same. To study the effect of different weights on the market dynamics, we assume that the resulting dynamical process is different from the Gaussian process. One of the simplest nontrivial distributions which cause a non-Gaussian process is a power-law distribution. The power-law distribution can be observed in various systems. For example, the popularity of a file in peer-to-peer network, which can be regarded as a value of the file, is known to follow a power-law [17–20]. Like the popularity distribution of file, we assign a weight to each node which is drawn from the power-law distribution

$$
P(w) \sim w^{-\delta}.\tag{3}
$$

Here  $\delta$  determines the heterogeneity in the weight of information.

#### *3.1. Weight-driven information spreading (WDIS) model*

We consider a system of *N* agents (or nodes). Each agent *i* can have one of the three states  $\phi_i \in \{-1, 0, 1\}$  and weight of information  $w_i$  drawn from the distribution (3). Each state corresponds to an inactive ( $\phi = 0$ ), selling ( $\phi = -1$ ), or buying ( $\phi = 1$ ) state. If  $\phi = \pm 1$  then the agent is in active state. Initially, all agents are inactive and isolated (no links or cooperation among them). The network of the cooperation evolves as follows: At each time step *t* (I) an agent *i* is randomly selected. (II) *Activity*: With probability *a* the agent *i* randomly chooses one state between  $\phi_i(t) = 1$  and  $\phi_i(t) = -1$ , and instantly all agents in the cluster to which *i* belongs choose the same state with agent *i* simultaneously. The aggregated state of the system  $s_i(t) = \sum_{j=1}^N \phi_j(t)$  is calculated. Thus,  $|s_i(t)|$  represents the active cluster size or the size of order at time *t*. Then disconnect all the links between active agents. This process defines the economic activity or simply activity. (III) *Cluster formation*: With the probability  $1 - a$ , the state of agent *i* remains inactive and new  $[w_i - k_i]$  links between agent *i* and a set of randomly selected agents  $\{j\}$  are established if  $k_i\leq w_i$  and  $k_j\leq w_j$ . Here [x] denotes the maximum integer which is not greater than *x*, and *k<sup>i</sup>* represents the degree or number of nodes already linked to the agent *i* with undirected connections. This procedure implies that another link is added if the carrying capacity of both agents involved is not exhausted.

#### *3.2. Weight-driven information spreading with synergetic cooperation (WDISSC) model*

The initial condition and the procedure (I)–(III) of the WDISSC model are the same with those of WDIS model. But when the two nodes are connected in the process (III) the value of information is increased synergetically, i.e.  $w_i \to w_i + 1$  and  $w_i \rightarrow w_i + 1$  [21]. After the activity (II), the weights are restored to the initial weight which is originally drawn from the distribution (3).

#### **4. Simulation results**

#### *4.1. Normalized return distribution*

Before discussing our simulation results, let us define the price and return. As one of the simplest assumptions, the evolution of the price  $p(t)$  at *t* is proportional to both the size  $s(t)$  of the order and  $p(t)$  [22]. The price might increase (decrease) when the net activity is buying (selling). Thus, we assume that the price change is given by

$$
\frac{\mathrm{d}p(t)}{\mathrm{d}t} = \frac{s(t)}{\lambda}p(t),\tag{4}
$$

where  $\lambda$  is a parameter which controls the size of price. In the numerical simulations we use  $\lambda=5\times10^4$  as in Ref. [9]. The value of λ does not affect the analysis of the return distribution. In our model we consider only a discrete time series of *p*(*t*),  $(t = 0, 1, 2, \ldots)$ . In the discrete time series  $p(t)$  evolves as  $p(t + 1) = p(t) \exp(s(t)/\lambda)$  from Eq. (4). Due to the symmetry of the model,  $\langle R(t, \Delta t) \rangle = 0$  for all  $\Delta t$  ( $\Delta t = 1, 2, 3, ...$ ). In Ref. [9], the authors mainly focused on the properties of *P*(*Z*( $\Delta$ *t* = 1)) to understand the relationship between the information dispersion and the occurrence of financial crash. In the following analysis, we will show that the scaling properties of  $P(Z = 0; \Delta t)$  can provide a novel criterion to determine the threshold *a*<sup>∗</sup> below which the financial crash occurs.

From the measurements of stock indices it has been shown that the distribution of *R* or equivalently the distribution of *Z* satisfies a leptokurtic distribution with power-law or fat-tailed distribution for small ∆*t* [2,5]. The obtained distribution has been approximated by Lévy stable distribution for small  $\Delta t$  and  $0 < \alpha < 2$ :

$$
P(Z) = \frac{1}{\pi} \int_0^\infty \exp\left(-\gamma |q|^\alpha\right) \cos(qZ) dq. \tag{5}
$$

Here  $\alpha$  and  $\gamma$  are the index and scaling factor, respectively. The distribution (5) is known to be stable when  $0 < \alpha \leq 2$  [12]. The distribution (5) approaches the power-law [2]

$$
P(|Z|) \sim |Z|^{-(1+\alpha)},\tag{6}
$$

when  $|Z| \gg 1$  and  $\alpha < 2$ . However, most empirical studies on the stock market indices have shown that the *P*(*Z*) with a finite second moment deviates from the Lévy stable distribution for large *Z*. As an alternative, the truncated Lévy distribution has been suggested to have finite variance [2].

Now let us discuss the numerical simulation results for WDIS and WDISSC models. In the simulations we use  $P(w) \sim w^{-\delta}$ with  $\delta = 2.5$  because the variance of  $P(w)$  diverges for  $\delta < 3$ . We rescale the time to place an activity (II) at every time on the average:  $t \to at$  [9]. In Fig. 3 we display the obtained  $P(|Z(\Delta t = 1)|)$  and  $P(Z(\Delta t = 1))$  of both WDIS and WDISSC models when  $a = 0.9$  which corresponds to the large activity regime. From the best fit of Eq. (6) to the data of  $P(|Z(\Delta t = 1)|)$  we obtain  $\alpha \simeq 1.5$  for both models (see Fig. 3(a) and (b)). In Fig. 3(c) and (d), we compare the obtained *P*( $Z(\Delta t = 1)$ )'s of our models with Eq. (5). With the obtained  $\alpha \simeq 1.5$  we find that the simulation results agrees well with Eq. (5) when  $\gamma = 0.05$ for WDIS model and  $\gamma = 0.005$  for WDISSC model. Note that, in this high activity limit ( $a \to 1$ ), there is no sufficient time to spread information or to form a large cluster. Thus, if we choose an agent for an activity at random, the rest of agents in the active cluster are mostly restricted to the directly connected neighbors of the selected agent. Since the degree of the chosen agent is proportional to the assigned weight, the distribution of the number of active agents or active cluster size is simply proportional to the given weight distribution. From Eqs. (4) and (1), the active cluster size  $|s(t)|$  satisfies  $|s(t)| = \lambda |R(t, \Delta)|$ when  $\Delta t = 1$ . Therefore, the (normalized) return distribution is proportional to Eq. (3) and  $1 + \alpha = \delta$ . We obtained the similar results for other moderated values of  $\delta$ . In the limit  $\delta \to \infty$  we expect that  $P(Z(\Delta t = 1))$  approaches the exponential distribution. The power-law tails of data in Fig. 3 also indicate that there can be large clusters in which the majority of the agents sharing the same information and produce the large return. This behavior is usually known as ''herding behavior'' [8,9,15].

As *a* decreases, we find a threshold  $a^*$  below which a bump in  $P(|Z(\Delta t = 1)|)$  for large  $|Z(\Delta t = 1)|$  is observed in both WDIS and WDISSC models (see Fig. 4(a) and (b)). This relative increase in  $P(|Z(\Delta t = 1)|)$  for large  $|Z(\Delta t = 1)|$  is expected to be related to the financial crash of a market [9] (see Section 2). The similar behavior was reported in EZ model. a<sup>\*</sup> of EZ model was reported as a<sup>\*</sup> < 0.10 in Ref. [9]. To find more precise value of a<sup>\*</sup> we measure the successive slope  $1+\alpha_s \equiv -\delta[\log P(|Z(\Delta t = 1)|)/\delta[\log |Z(\Delta t = 1)|]$ . When  $P(|Z(\Delta t = 1)|)$  follows Eq. (6),  $1+\alpha_s$  should be the same with 1 + α. On the other hand, if there is a bump in *P*(|*Z*(∆*t* = 1)|) then 1 + α*<sup>s</sup>* becomes less than 1 + α, and if *P*(|*Z*(∆*t* = 1)|) decays exponentially then there is an abrupt increase of  $1+\alpha_s$ . This can be clearly seen from the data in the insets of Fig. 4. For  $a\approx a^*$ ,  $1+\alpha_s$  coincides with  $1+\alpha$  which is obtained from the best fit of  $P(|Z(\Delta t=1)|)$  to Eq. (6) and abruptly increases (open circles in the insets). However, if *a* decreases further then  $1 + \alpha_s$  becomes smaller than  $1 + \alpha$  for moderate value of |*Z*(∆*t* = 1)| which indicates the occurrence of the bump in *P*(|*Z*(∆*t* = 1)|)(see the open squares in the insets). The detailed measurements of  $1+\alpha_s$  reveal that  $a^*\approx 0.06$  for EZ model (which is not shown). From the data in Fig. 4(a) and (b) we find that  $a^*\approx 0.11$  and  $a^*\approx 0.4$  for WDIS and WDISSC model, respectively. The values of  $a^*$  of our models are larger than that of EZ model. This implies that the financial crash can occur not only in the low activity limit but also in the moderate activity regime compared to EZ model, for example  $a\lesssim$  0.4 in WDISSC model. The obtained scaling exponents around  $a\approx a^*$  are  $\alpha \simeq 0.6 \pm 0.1$  for WDIS model and  $\alpha \simeq 0.7 \pm 0.1$  for WDISSC model (see the dashed lines in Fig. 4(a) and (b)). In EZ model  $\alpha \simeq 0.5$  for  $a \simeq a^*$  was obtained.

*S.-H. Yook, Y. Kim / Physica A 387 (2008) 6605–6612* 6609



**Fig. 3.** (color online) Plots of  $P(|Z|)$  against |*Z*| for (a) WDIS model and (b) WDISSC model with  $a = 0.9$  and  $\Delta t = 1$ . The solid lines represent the power-law  $P(|Z|) \sim |Z|^{-(1+\alpha)}$  with  $\alpha \simeq 1.5$  for both models. The value of  $\alpha$  is obtained from the best fit of Eq. (6) to the data. Plots of *P*(*Z*) for (c) WDIS model and (d) WDISSC model with  $a = 0.9$  and  $\Delta t = 1$ . (The solid lines correspond to the Lévy distribution (Eq. (5)) with  $\alpha = 1.5$ , (b)  $\gamma = 0.05$  and (c)  $\gamma = 0.001$ .).

### *4.2. Scaling property of return distribution*

In order to investigate the crossover observed in Figs. 3 and 4 in more details, we consider the scaling arguments on Lévy stable distribution suggested by Mantegna and Stanley [2,5]. They showed that the self-similarity of Eq. (5) is related to the 'probability of return to origin' which corresponds to  $P(Z = 0; \Delta t)$  and scales as

$$
P(Z=0; \Delta t) \sim \frac{1}{(\Delta t)^{1/\alpha}} \sim \Delta t^{-\tau}.
$$
\n(7)

From the empirical studies on S&P 500 index, Mantegna and Stanley showed that the return distribution of the stock indices satisfy Eqs. (6) and (7) with  $\alpha \simeq 1.40$  (or  $\tau \simeq 0.7$ ) up to a certain value of  $\Delta t$ . As increasing  $\Delta t$  further, the returns are not correlated any more and the return distribution becomes Gaussian. Thus,  $\tau$  becomes 0.5 since  $\alpha = 2$  for the Gaussian distribution. This crossover from the Lévy distribution to the Gaussian distribution is well described by the truncated Lévy flight (TLF) distribution [2]. The TLF is not a stable distribution. However, TLF is quite similar to the Lévy stable distribution for small ∆*t* and converges to the Gaussian distribution when ∆*t* → ∞. We now apply this scaling argument to our models. As shown in Fig. 5 (a) and (b) when  $a > a^*$  we find that  $\tau \simeq 1.7$  for WDIS model and  $\tau \simeq 1.4$  for WDISSC model when  $\Delta t$  < 200. The estimated  $\alpha$ 's from the relation  $\alpha = 1/\tau$ , Eq. (7), agree well with those obtained from the data in Fig. 4. As we increase Δ*t* further ( $Δt > 200$ ), the correlation disappears and  $P(Z; Δt)$  becomes Gaussian. Thus, we obtain  $τ = 0.5$ (see the dashed lines in Fig. 5(a) and (b)) as expected. These results consistent with the empirical results. However, if  $a < a^*$ then we find that  $P(Z = 0; \Delta t)$  decays exponentially (Fig. 5(c) and (d)).

The exponentially decaying  $P(Z=0; \Delta t)$  for  $a < a^*$  can be understood from the following analysis. If we approximate the bump for large |Z| to a Gaussian distribution with non-zero mean  $\pm\mu$  and variance  $D^2$ , then  $P(Z;\,\varDelta t)$  can be expressed as (see Section 2)

$$
P(Z; \Delta t) \sim P_L(Z; \Delta t) \left[ P_G^+(Z(\Delta t)) \Theta(Z(\Delta t)) + P_G^-(Z(\Delta t))(1 - \Theta(Z(\Delta t))) \right],\tag{8}
$$

6610 *S.-H. Yook, Y. Kim / Physica A 387 (2008) 6605–6612*



**Fig. 4.** (color online) Plot of *P*(|*Z*|) against |*Z*| in log-log scale for small values of *a* when ∆*t* = 1. (a) WDIS model and (b) WDISSC model. As *a* decreases the relative increase of *P*(|*Z*|) for large |*Z*| is observed in both models. Dashed lines denotes the relation (a) *P*(|*Z*|) ∼ |*Z*|<sup>-1.6</sup> for WDIS model and (b)  $P(|Z|) \sim |Z|^{-1.7}$  for WDISSC model. Insets of (a) and (b) display the successive slope of  $P(|Z|)$ . Horizontal lines indicate the relations (a)  $(1 + \alpha) = 1.6$ and (b)  $(1 + \alpha) = 1.7$ .



**Fig. 5.** Plot of  $P(Z = 0; \Delta t)$  against  $\Delta t$  for (a) WDIS model and (b) WDISS model. The slope of the solid lines satisfy the relation (a)  $\tau \approx 1.7$  and (b)  $\tau \approx 1.4$ . The dashed lines denote  $1/\alpha \approx 0.5$  for both (a) and (b) when  $a > a^*$ . (c) and (d) display  $P(Z = 0; \Delta t)$  when  $a < a^*$  for WDIS model (c) and WDISS model (d). The solid lines in (c) and (d) represent the exponentially decaying  $P(Z = 0; \Delta t)$ .

for  $a < a^*$ . Here  $P_L(Z(\Delta t))$  denotes Lévy distribution given by Eq. (5) and  $P_G^{\pm}(Z(\Delta t))$  represents the Gaussian distribution with mean  $\pm\mu$ .  $\Theta(Z)$  is the Heaviside theta function satisfying  $\Theta(Z< 0) = 0$  and  $\Theta(Z \ge 0) = 1$ . By the convolution *S.-H. Yook, Y. Kim / Physica A 387 (2008) 6605–6612* 6611



**Fig. 6.** Plot of the averaged distribution of cluster size *s* for (a) WDIS model and (b) WDISS model. The solid lines denote the relation  $P(s) \sim s^{-\beta}$ . From the best fit, we obtain  $\beta \simeq 2.7$  for  $a \lesssim a^*$  (solid lines) and  $\beta \simeq 3.5$  when  $a = 0.9$  (dashed lines).

theorem, we obtain,

$$
P(Z=0; \Delta t) \sim \int_0^\infty \exp\left[-\Delta t \left(\gamma q^\alpha + \frac{D^2}{2}q^2 - i\mu q\right)\right] dq.
$$
\n(9)

When  $\alpha \ll$  2 and  $\gamma \ll D^2/2$ , Eq. (9) can be approximated by

$$
P(Z=0;\,\Delta t)\sim e^{-\frac{\mu^2}{2D^2}\Delta t}.\tag{10}
$$

Thus, Eq. (10) provides an evidence that the product of Lévy distribution and Gaussian-like distribution with non-zero mean can cause exponentially decaying  $P(Z = 0; \Delta t)$ . This result indicates that the emergence of bump causes a crossover from Eq. (7) to Eq. (10). Thus, Eqs. (7) and (10) can be used as another criterion to find the threshold *a* ∗ .

#### *4.3. Cluster size distribution*

Since the probability to choose a cluster of size *s* is proportional to *s*, the  $P(|Z|)$  is known to be directly related to the cluster size distribution, *P*(*s*). More specifically, if *P*(*s*) follows the power-law, *P*(*s*) ∼ *s*<sup>-β</sup>, then *P*(|*Z*|) ∼ |*Z*|<sup>-(1+α)</sup> ∼ *s*<sup>-(β-1)</sup> [9]. Thus, the exponents  $\alpha$  and  $\beta$  satisfy the relation

$$
\beta = \alpha + 2. \tag{11}
$$

In Fig. 6 we show the measured *P*(*s*) for both models. From the data we find that *P*(*s*) satisfies power-law *P*(*s*)  $\sim s^{-\beta}$  as assumed in Eq. (11). From the best fit to the data we obtain  $\beta \simeq 2.7$  when  $a \lesssim a^*$  and  $\beta \simeq 3.5$  for  $a=0.9$  (>  $a^*$ ). The values of  $\alpha$ 's and  $\beta$ 's obtained from Figs. 4 and 6 satisfy the relation (11).

# **5. Summary and discussions**

We study the effects of non-trivially distributed weight on the herding behavior and the outbreak of financial crash through two agent based weight-driven information spreading models. From the empirical analysis on S&P 500 index, we observe an explicit bump in *P*(|*Z*|) for large |*Z*| during the crash period. Through the numerical simulations, we show that the uneven distribution of the weight of information in our models can increase the  $a^*$ , below which a financial crash is expected, compared to the EZ model. The increment a<sup>\*</sup> becomes more significant when the synergetic cooperation occurs. Moreover, in EZ model  $P(Z(\Delta t = 1))$  decays exponentially for  $a > a^*$ . Such exponentially decaying  $P(Z)$  is not observed in real markets, but in our models  $P(Z(\Delta t = 1))$  follows the power-law or Lévy distribution even in the large activity regime. Therefore, we expect that our models are more appropriate to describe realistic situations in financial markets than EZ model.

We also find that the scaling behavior of  $P(Z=0;\Delta t)$  satisfies the power-law when the market does not crash ( $a>a^*$ ). However, in the financial crash regime ( $a < a^*$ ),  $P(Z = 0; \Delta t)$  exponentially decays. By assuming the relative increase of

#### 6612 *S.-H. Yook, Y. Kim / Physica A 387 (2008) 6605–6612*

normalized return for large *Z* as the Gaussian distribution with non-zero mean, we analytically show that  $P(Z = 0; \Delta t)$ decays exponentially. This can provide a novel criterion to find the threshold a<sup>∗</sup> below which the financial crash is expected. The same scaling behavior is also observed in EZ model.

### **Acknowledgements**

We thank to Mr. Ko for the help of computational work and Prof. Jeong for providing S&P 500 data during the crash period. This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST) (No. R01-2007-000-10910-0 and No. R01-2006-000-10470-0).

#### **References**

- [1] L. Bachelier, Ann. Sci. École Norm. Sup. III-17 (1900) 21.
- [2] R.N. Mantegna, H.E. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, New York, 2000.
- [3] P. Gopikrishnan, M. Meyer, L.A.N. Amaral, H.E. Stanley, Eur. Phys. J. B 3 (1998) 139.
- [4] U.A. Müller, M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz, C. Morgenegg, J. Banking Finance 14 (1995) 1189.
- [5] R.N. Mantegna, H.E. Stanley, Nature 376 (1995) 46.
- [6] J.-P. Bouchaud, M. Potters, Theory of Financial Risk, Cambridge University Press, Cambridge, 2000.
- [7] J.D. Farmer, Comput. Sci. Eng. 1 (1999) 26.
- [8] R. Cont, J.P. Bouchaud, Macroeconomic Dyn. 4 (2000) 170.
- [9] V.M. Eguíluz, M.G. Zimmermann, Phys. Rev. Lett. 85 (2000) 5659.
- [10] T. Lux, Appl. Financial Econom. 6 (1996) 463.
- [11] A. Allison, D. Abbott, Unsolved Problems of Noise, in: D. Abbott, L. Kish (Eds.), AIP proceedings, Melville, New York, 2000.
- [12] R.N. Mantegna, H.E. Stanley, Phys. Rev. Lett. 73 (1994) 2946.
- [13] F. Lillo, R.N. Mantegna, Eur. Phys. J. B 15 (2000) 603.
- [14] T. Mart, Y. Surya, Physica A 344 (2004) 198.
- [15] D. Chowdhury, D. Stauffer, Eur. Phys. J. B 8 (1999) 477.
- [16] D. Sornette, New Thesis 1 (2004) 5.
- [17] V. Almeida, A. Bestavros, M. Crovella, A. de Oliveira, Characterizing reference locality in the WWW, in: Proceedings of 1996 International Conference on Parallel and Distributed Information Systems, 1996.
- [18] K. Sripanidkulchai, The popularity of gnutella queries and its implications on scalability, in: O'Reilly's www.openp2p.com, February 2001.
- [19] Q. Lv, et al. Search and replication in unstructured peer-to-peer networks, in: Proceedings of International Conference on Supercomputing, 2002. [20] S. Lee, S.-H. Yook, Y. Kim, Physica A 385 (2007) 743.
- [21] We also considered some variations in the increase of weight, for example, depending on the weight of cooperating partner and found the qualitatively the same results with WDISSC model.
- [22] M. Bartolozzi, A.W. Thomas, Phys. Rev. E 69 (2004) 046112.