The Theory of Higher Order Lognormal Cascade Distribution and the Origin of Fat Tails in the Fluctuations of Stock Market Index

Stephen H.-T. Lihn

Piscataway, NJ 08854

stevelihn@gmail.com

First Draft revised on June 5, 2009

Abstract

This working paper presents the general theory of the higher order "skew lognormal cascade distribution" as a mathematical extension of the previously proposed skew lognormal cascade distribution (Stephen Lihn 2008, SSRN: 1273087, which is the first order cascade over the normal distribution). In particular, the second order distribution is studied in details, which incorporates the fat tails into the volatility (aka the volatility of volatility). We show that the second order distribution can handle very heavy tails and high kurtosis in the high frequency financial time series. It accurately fits the first four moments of the daily log-returns of Dow, whose kurtosis is 26. The framework of the higher order lognormal cascade distributions also provides a new way to study the capital distribution (aka firm size distribution), the market index, and the market entropy of the stock market. Such study in the context of stochastic portfolio theory reveals that the origin of the fat tails in the fluctuations of the market index is from the lognormal cascade structure of the capital distribution in the market. We show from a simple stochastic model that the contraction and expansion of the underlying capital distribution is the fundamental driving force of the bull-bear market cycles and the market volatility in the past 20 years. A stochastic equation is derived to establish the relation between the market index and the capital distribution, which is exactly the lognormal cascade equation in our theory. This shows that the fluctuations of the market index are a natural mathematical consequence of the stochastic calculus on the market portfolio in which weights are exponentially distributed. Therefore, we conclude that the phenomena of fat tails should exist everywhere in our financial system.

Acknowledgement: Complicated algebraic results in this paper are carried out by the open source computer algebra system, Maxima. Numerical computation is prototyped in GNU Octave.

1 Defining The Higher Order Lognormal Cascade Distribution

In this section, we define the higher order theory for the skew lognormal cascade distribution. We shall use a logarithmic model for continuous-time processes. The process X, which can be price or volatility or an abstract process, shall always be presented in its logarithm, $\log X$, which is abbreviated as χ (See 1.1 of Fernholz 2002).

Under the limit of $\tau_c \to 0$, the SIBM model (Lihn 2008, SSRN: 1149142) can be reduced to the following emperical stochastic equation [1] :

$$d_t\chi(t) = \Phi \cdot e^{\mathcal{H}(t)} \left[d_tW(t) + (\beta \cdot \mathcal{H}(t) + g) \, dt \right]. \quad (1)$$

The parameters are defined as following: Φ is a global scale constant, β is "the skewness parameter", and g is the constant growth term. $\mathcal{H}(t)$ is a time dependent, slow varying and mean reverting, cyclic process, representing the stochastic volatility term in this context. The "slow varying" assumption is called the higher order randomness hypothesis (HORN) (Lihn 2008, SSRN: 1149142). Since we are defining a static distribution with a delta time interval Δt (e.g. $\Delta t = 1$ day or 1 month) from Equation (1), the exact time dependency of $\mathcal{H}(t)$ is of less interest in this paper (A simple way to model a normally distributed $\mathcal{H}(t)$ is by the Ornstein–Uhlenbeck process). All we are interested in is the kind of static probability distribution $\mathcal{H}(t)$ produces over time ($\{\mathcal{H}(t_n) \mid t_n = n\,\Delta t, n \subset$ positive integers $\}$).

In the previous paper (SSRN: 1273087), we have discussed the case of a normally distributed $\mathcal{H}(t)$, i.e. $\{\mathcal{H}(t_n)\} \sim N\left(0, \eta^2\right)$. Our new framework here will absorb the case of a normally distributed $\mathcal{H}(t)$, and its resulting skew lognormal cascade distribution, $d\chi/dt$. In that case, the distribution of $\mathcal{H}(t)$ will be called "the zero-th order distribution" and the distribution of $d\chi/dt$ will be called "the first order distribution" in the new framework we will present in this paper.

In order to present this new framework, we want to first generalize what $\mathcal{H}(t)$ can be. Let's assume abstractly $\mathcal{H}(t)$ can be any skew lognormal cascade distribution with zero mean (This includes a normal distribution). And such generalization is followed with the definition of **the higher order theory**:

---

[1] There is a possible extension by adding one more parameter α:

$$d_t\chi(t) = \Phi \left[e^{\mathcal{H}(t)} \, d_tW(t) + e^{\alpha\,\mathcal{H}(t)} \, (\beta \cdot \mathcal{H}(t) + g) \, dt \right]. \quad (2)$$

where α could be a positive integer 0, 1, 2, etc. In this paper, we will discuss the emperical case of $\alpha = 1$. There is an indication from stochastic portfolio theory that $\alpha = 0$ or 2 could be potential choices too. The framework layed out in this paper can be easily extended to investigate these cases. Such investigation is left for future research.

Definition 1.1. *If $\mathcal{H}(t)$ is a $(n-1)$-th order skew lognormal cascade distribution, then we* **define** $d\chi/dt$ [2] **is "lifted"** *to the n-th order skew lognormal cascade distribution from $\mathcal{H}(t)$according to Equation (1). We use the term "lift" since we now have a properly defined higher order hierarchy. The requirement of zero mean is simply the fact that any long running average in $\mathcal{H}(t)$ can and should be detrended to Φ, and one redundant parameter is reduced by doing so.*

Let's now define the abstract probability density function (pdf) of the n-th order skew lognormal cascade distribution:

$$p_{(n)}(x; \langle \bullet \rangle_{(n)}) \tag{3}$$

in which $\langle \bullet \rangle_{(n)}$ represents all the parameters required to characterize the n-th order distribution: The β, g, Φ, and so on, and all of them with proper suffix of (i) where $i = 0..n$. More details on $\langle \bullet \rangle_{(n)}$ will follow shortly.

Next we define the kernel function as

$$K(x, \mathcal{H}; \beta, g, \Phi) = \frac{1}{\sqrt{2\pi}\,\sigma(\mathcal{H})}\, e^{-\frac{(x-\sigma(\mathcal{H})\ (\beta\cdot\mathcal{H}+g))^2}{2\sigma(\mathcal{H})^2}}, \tag{4}$$

where $\sigma(\mathcal{H}) = \Phi \cdot e^{\mathcal{H}}$. [3] Two special cases of the kernel function are worth mentioning. First, when $\beta = g = 0$, we have

$$K(x, \mathcal{H}; 0, 0, \Phi) = \frac{1}{\sqrt{2\pi}\,\sigma(\mathcal{H})}\, e^{-\frac{x^2}{2\sigma(\mathcal{H})^2}}, \tag{5}$$

which resembles the form of the typical (symmetric) stochastic volatility models. Second, when $\mathcal{H}$ is zero (recall that any constant in $\mathcal{H}$ is factored to Φ), we have $\sigma(\mathcal{H}) = \Phi$ and

$$K(x, 0; \beta, g, \Phi) = \frac{1}{\sqrt{2\pi}\,\Phi}\, e^{-\frac{(x-\Phi g)^2}{2\Phi^2}}, \tag{6}$$

which is a normal distribution of $N(\Phi g, \Phi^2)$. Notice that in this case, β is irrelevant and therefore should be set to zero for simplicity and convenience.

Definition 1.2. *Let's now define the n-th order distribution in terms of the kernel function and the lower order distributions:*

---

[2] We call $d\chi/dt$ either the returns of χ or the log-returns of X. It depends on the context of the discussion. In the financial market, the market indices and stock prices are represented by X, thus log-returns are mentioned most frequently.

[3] If Equation (2) is considered, the kernel function becomes

$$K(x, \mathcal{H}; \beta, g, \Phi, \alpha) = \frac{1}{\sqrt{2\pi}\,\sigma(\mathcal{H})}\, e^{-\frac{(x-\Phi\cdot e^{\alpha\ \mathcal{H}}\ (\beta\cdot\mathcal{H}+g))^2}{2\sigma(\mathcal{H})^2}}. \tag{7}$$

$$p_{(n)}(x;\beta_{(n)},g_{(n)},\Phi_{(n)};\langle\bullet\rangle_{(n-1)}) = \int_{-\infty}^{\infty} d\mathcal{H}\, p_{(n-1)}\Big(\mathcal{H};\langle\bullet\rangle_{(n-1)}\Big)\ K\big(x,\mathcal{H};\beta_{(n)},g_{(n)},\Phi_{(n)}\big). \tag{8}$$

Again, using the language of lifting, we say $p_{(n-1)}(\mathcal{H};\langle\bullet\rangle_{(n-1)})$ is **lifted** to $p_{(n)}(x;\langle\bullet\rangle_{(n)})$ by the kernel $K\big(x,\mathcal{H};\beta_{(n)},g_{(n)},\Phi_{(n)}\big)$. This recursive definition starts with the zero-th order distribution, which is a normal distribution:

$$p_{(0)}(x;0,g_{(0)},\Phi_{(0)}) = K(x,0;0,g_{(0)},\Phi_{(0)}). \tag{9}$$

If we substitute $\eta=\Phi_{(0)}$ and $\mu=g_{(0)}\Phi_{(0)}$, we have $p_{(0)}(x;0,\mu/\eta,\eta)\sim N(\mu,\eta^2)$.

The first order distribution is the "skew lognormal cascade distribution" defined in the previous paper (SSRN: 1273087):

$$p_{(1)}(x;\beta_{(1)},g_{(1)},\Phi_{(1)};\langle\bullet\rangle_{(0)}) = \int_{-\infty}^{\infty} d\mathcal{H}\, p_{(0)}(\mathcal{H};0,0,\eta)\, K(x,\mathcal{H};\beta_{(1)},g_{(1)},\Phi_{(1)}), \tag{10}$$

where $\langle\bullet\rangle_{(0)}$ is simply $\Phi_{(0)}$. For simplicity reason, $p_{(1)}(x;\beta_{(1)},g_{(1)},\Phi_{(1)};\Phi_{(0)})$ has been written in the form of $p_{(1)}(x;\eta,\beta,g,\Phi)$ in which $\eta=\Phi_{(0)}$ and the suffix is removed in $\beta_{(1)},g_{(1)},\Phi_{(1)}$. Notice that $p_{(0)}$ is made to have zero mean, $\mu=0$. The numerical implementation of $p_{(1)}$ is reasonably fast and it provides reasonable accuracy when fitting time series data with kurtosis less than 6. The successful application of $p_{(1)}$ to fit many financial time series has been demonstrated on the author's website (See http://www.skew-lognormal-cascade-distribution.org/apps/ for more details).

In later sections of this paper, we will explore the structure and numerical implementation of the second order distribution $p_{(2)}$. The need to explore $p_{(2)}$ arises from the fact that some important time series data exhibit very high kurtosis, which can not be explained well by $p_{(1)}$. One such example is the one-day log-returns of Dow, whose kurtosis is 26. Our financial exploration can not be called successful if we can't fit Dow data with desired accuracy. We believe the second order distribution gives us enough dimensions to handle the high frequency data of Dow properly.

The second order distribution is written as

$$p_{(2)}(x;\beta_{(2)},g_{(2)},\Phi_{(2)};\langle\bullet\rangle_{(1)}) = \int_{-\infty}^{\infty} d\mathcal{H}\, p_{(1)}(\mathcal{H};\eta,\beta,g,\Phi)\, K(x,\mathcal{H};\beta_{(2)},g_{(2)},\Phi_{(2)}), \tag{11}$$

where $\langle\bullet\rangle_{(1)}$ contains the four parameters that characterizes $p_{(1)}$: η,β,g,Φ. However, since $p_{(1)}$ should have zero mean, we have $g=-\beta\cdot\eta^2$. Therefore, $p_{(2)}$ has six independent parameters: $\beta_{(2)},g_{(2)},\Phi_{(2)}$ and $\beta_{(1)},\Phi_{(1)}$, and $\Phi_{(0)}$. We see that $p_{(2)}$ is more complicated than $p_{(1)}$. The complexity brings with it more flexibility and power; but at the same time also demands more computational resources. The reader should be able to derive more higher order distributions via the recursive relation if that interests you.

2 The Moments of The Higher Order Distribution

In this section, we will study the statistic properties for the higher order distributions. We shall begin with the characteristic function $\Psi_{(n)}(t) = \int_{-\infty}^{\infty} dx\, e^{itx} p_{(n)}(x)$ which is

$$\frac{1}{\sqrt{2\pi}\,\Phi_{(n)}} \int \int_{-\infty}^{\infty} dx\, d\mathcal{H}\, e^{itx - \frac{e^{-2\mathcal{H}}}{2\Phi_{(n)}^2}\left(x - \Phi_{(n)}\, e^{-\mathcal{H}}(\beta_{(n)}\mathcal{H} + g_{(n)})\right)^2 - \mathcal{H}}\, p_{(n-1)}\Big(\mathcal{H}; \langle\bullet\rangle_{(n-1)}\Big). \tag{12}$$

The integral on x can be carried out exactly,

$$\Psi_{(n)}(t) = \int_{-\infty}^{\infty} d\mathcal{H}\, e^{-\frac{t^2}{2}\left(\Phi_{(n)} e^{\mathcal{H}}\right)^2 + it\, \Phi_{(n)} e^{\mathcal{H}}\left(\beta_{(n)}\mathcal{H} + g_{(n)}\right)}\, p_{(n-1)}\Big(\mathcal{H}; \langle\bullet\rangle_{(n-1)}\Big). \tag{13}$$

Let $q = it$, we can obtain the k-th moment $\mu_k^{(n)}$ via the Taylor series

$$\Psi_{(n)}(-iq) = \sum_{k=0}^{\infty} \mu_k^{(n)} q^k / k!, \tag{14}$$

or in a derivative form

$$\mu_k^{(n)} = \frac{d^k}{dq^k}\left(\Psi_{(n)}(-iq)\right)_{q=0}. \tag{15}$$

We'd like to carry out the first four moments here. The first moment is

$$\mu_1^{(n)} = \Phi_{(n)} \int_{-\infty}^{\infty} d\mathcal{H}\, e^{\mathcal{H}}\left(\beta_{(n)}\mathcal{H} + g_{(n)}\right)\, p_{(n-1)}\Big(\mathcal{H}; \langle\bullet\rangle_{(n-1)}\Big). \tag{16}$$

It is interesting to note that $\mu_1^{(n)}$ is simply the expected value of the skew terms in the kernel function in (4), $\langle \sigma(\mathcal{H})\,(\beta \cdot \mathcal{H} + g)\rangle$. The integral of $\mathcal{H}$ can be carried out by expanding $p_{(n-1)}(\mathcal{H})$ into $\int_{-\infty}^{\infty} dy\, p_{(n-2)}(y)\, K\Big(\mathcal{H}, y; \langle\bullet\rangle_{(n-1)}\Big)$ (via Equation (8)) and results in

$$\begin{aligned} \mu_1^{(n)} = \Phi_{(n)} \textstyle\int_{-\infty}^{\infty} dy\, &\Big(\beta_{(n)}\Big(\sigma_{(n-1)}(y)\mathbb{Y}_{(n-1)} + \sigma_{(n-1)}^2(y)\Big) + g_{(n)}\Big) \\ &e^{\sigma_{(n-1)}(y)\mathbb{Y}_{(n-1)} + \frac{\sigma_{(n-1)}^2(y)}{2}} \cdot p_{(n-2)}\Big(y; \langle\bullet\rangle_{(n-2)}\Big). \end{aligned} \tag{17}$$

where $\mathbb{Y}_{(n-1)} = \beta_{(n-1)}\, y + g_{(n-1)}$ and $\sigma_{(n-1)}(y) = \Phi_{(n-1)} \cdot e^y$. The integral of $\mathcal{H}$ in the higher moments can be carried out in a similar fashion and one level of integration is eliminated. Such simplification is very useful for the numerical implementation of descriptive statistics.

Let $\mathbb{H}_{(n)} = \beta_{(n)}\mathcal{H} + g_{(n)}$, we can effectively simplify the presentation of higher moments. The second moment is

$$\mu_2^{(n)} = \Phi_{(n)}^2 \int_{-\infty}^{\infty} d\mathcal{H}\, e^{2\mathcal{H}}\left(\mathbb{H}_{(n)}^2 + 1\right)\, p_{(n-1)}\Big(\mathcal{H}; \langle\bullet\rangle_{(n-1)}\Big). \tag{18}$$

When $\mu_1^{(n)}$ is zero, or more precisely $\mu_1^{(n)} \ll \sqrt{\mu_2^{(n)}}$, this indicates $\chi(t)$ is also an oscillating process with very large amplitude compared to its almost zero long-term rate of increase. When $n = 1$, this is the well known condition of $g \approx -\beta\eta^2$. In real world, this is the case of Dow, representing the US stock market. Dow can have large swings of ~50% in each market cycle while delivers 5-7% long-term rate of return.

The third moment is

$$\mu_3^{(n)} = \Phi_{(n)}^3 \int_{-\infty}^{\infty} d\mathcal{H}\, e^{3\mathcal{H}} \left(\mathbb{H}_{(n)}^3 + 3\mathbb{H}_{(n)}\right)\; p_{(n-1)}\Big(\mathcal{H}; \langle\bullet\rangle_{(n-1)}\Big). \tag{19}$$

The fourth moment is

$$\mu_4^{(n)} = \Phi_{(n)}^4 \int_{-\infty}^{\infty} d\mathcal{H}\, e^{4\mathcal{H}} \left(\mathbb{H}_{(n)}^4 + 6\mathbb{H}_{(n)}^2 + 3\right)\; p_{(n-1)}\Big(\mathcal{H}; \langle\bullet\rangle_{(n-1)}\Big). \tag{20}$$

It should be obvious to the reader that the general form of the moments can be calculated via the generating function of the Hermite Polynomials (See Appendix of Lihn 2008, SSRN: 1273087). The analytical form of the moments for $(n) = (1)$ have been calculated previously. The reader can compute them from these formula and compare them to the results in SSRN 1273087. We list these moments here since some of them will be used in later sections:

$$\begin{aligned} \mu_1^{(1)} &= \Phi e^{\frac{\eta^2}{2}} \beta_1 = \Phi e^{\frac{\eta^2}{2}} (\beta\eta^2 + g), \\ \mu_2^{(1)} &= \Phi^2 e^{2\eta^2} (\gamma + \beta_2^2), \\ \mu_3^{(1)} &= \Phi^3 e^{\frac{9\eta^2}{2}} \beta_3 (3\gamma\; + \beta_3^2), \\ \mu_4^{(1)} &= \Phi^4 e^{8\eta^2} (3\gamma^2 + 6\beta_4^2\gamma + \beta_4^4), \end{aligned} \tag{21}$$

in which $\gamma = \beta^2\eta^2 + 1$ and $\beta_n = n\beta\eta^2 + g$.

From these moments, the cumulants can be computed numerically. The second cumulant κ_2 is $\mu_2 - \mu_1^2$, which is the variance. The third cumulant κ_3 is $\mu_3 - 3\kappa_2\kappa_1 - \kappa_1^3$. The skew is defined as $\kappa_3/\kappa_2^{3/2}$. The fourth cumulant κ_4 is $\mu_4 - 4\kappa_3\kappa_1 - 3\kappa_2^2 - 6\kappa_2\kappa_1^2 - 1\kappa_1^4$. And the kurtosis is κ_4/κ_2^2.

3 The Second Order Distribution

In this section, we will analyze the second order distribution $p_{(2)}$, based on the abstract higher order definitions in Sections 1 and 2. We will also discuss how to implement the numerical method of computing the pdf.

First of all, we repeat that $p_{(2)}$ has six independent parameters: $\beta_{(2)}, g_{(2)}, \Phi_{(2)}$ and $\beta_{(1)}, \Phi_{(1)}$ and $\Phi_{(0)}$. We'd like to rename these parameters slightly to simplify the notation within the second order context. We shall denote $p_{(2)}$ as $p_{(2)}(x; \beta, g, \Phi, \eta_2, \beta_2, \Phi_2)$. For $\Phi_{(0)}$ it becomes η_2. For parameters from the first order, $\beta_{(1)}, \Phi_{(1)}$ become β_2, Φ_2, and to satisfy the zero-mean requirement, we have the implicit relation for $g_{(1)}$ which is $g_2 = -\beta_2 \cdot \eta_2^2$.

For parameters from the second order, the suffix is removed; that is, $\beta_{(2)}, g_{(2)}, \Phi_{(2)}$ become β, g, Φ.

It follows from Equation (11) that

$$p_{(2)}(x;\beta,g,\Phi,\eta_2,\beta_2,\Phi_2) = \int_{-\infty}^{\infty} d\mathcal{H}\,\frac{p_{(1)}(\mathcal{H};\eta_2,\beta_2,g_2,\Phi_2)}{\sqrt{2\pi}\,\sigma(\mathcal{H})}\,e^{-\frac{(x-\sigma(\mathcal{H})\ (\beta\cdot\mathcal{H}+g))^2}{2\sigma(\mathcal{H})^2}}, \tag{22}$$

where $\sigma(\mathcal{H}) = \Phi\cdot e^{\mathcal{H}}$. And we repeat that

$$p_{(1)}(\mathcal{H};\eta_2,\beta_2,g_2,\Phi_2) = \int_{-\infty}^{\infty} dy\,\frac{1}{2\pi\,\eta_2\,\Phi_2} e^{-\frac{y^2}{2\eta_2^2}-y}\,e^{-\frac{(\mathcal{H}-\sigma(y)\ (\beta_2\cdot y+g_2))^2}{2\sigma(y)^2}}, \tag{23}$$

where $\sigma(y) = \Phi_2\cdot e^y$ and $g_2 = -\beta_2\cdot\eta_2^2$. Therefore, $p_{(2)}(x)$ contains two gaussian-like integrals. It is easy to verify that $\int_{-\infty}^{\infty} p_{(2)}(x)\,dx = 1$. But we find no general way to simplify the double integrals in $p_{(2)}(x)$ due to the e^{e^y} term. However, in the case studies we know of (Such as the daily log-returns of Dow), η_2 is typically much smaller than 1.0 (e.g. ˜0.2). A Taylor expansion around η_2 seems very promising.

Let's examine the first four moments with the new symbols we define for $p_{(2)}(x)$. The first moment is

$$\mu_1^{(2)} = \Phi\int_{-\infty}^{\infty} d\mathcal{H}\,e^{\mathcal{H}}\,(\beta\,\mathcal{H}+g)\ p_{(1)}(\mathcal{H};\eta_2,\beta_2,g_2,\Phi_2). \tag{24}$$

The integral on $\mathcal{H}$ can be carried out analytically

$$\mu_1^{(2)} = \Phi\int_{-\infty}^{\infty} dy\,\frac{\beta\,(\sigma^2(y)+\mathbb{Y}\,\sigma(y))+g}{\sqrt{2\pi}\,\eta_2} e^{-\frac{y^2}{2\eta_2^2}+\frac{\sigma^2(y)}{2}+\mathbb{Y}\,\sigma(y)}. \tag{25}$$

where $\mathbb{Y} = \beta_2\,y + g_2$ and $\sigma(y) = \Phi_2\cdot e^y$ and $g_2 = -\beta_2\cdot\eta_2^2$. Such simplication is very helpful to speed up the numerical computation.

Let $\mathbb{H} = \beta\,\mathcal{H}+g$, we can effectively simplify the presentation of higher moments. The second moment is

$$\mu_2^{(2)} = \Phi^2\int_{-\infty}^{\infty} d\mathcal{H}\,e^{2\mathcal{H}}\left(\mathbb{H}^2+1\right)\ p_{(1)}(\mathcal{H};\eta_2,\beta_2,g_2,\Phi_2), \tag{26}$$

which is simplified to

$$\mu_2^{(2)} = \Phi^2\int_{-\infty}^{\infty} dy\,\frac{\beta^2\sigma^2(y)\ +(\beta\,\sigma(y)\ (\mathbb{Y}+2\sigma(y))+g)^2+1}{\sqrt{2\pi}\,\eta_2} e^{-\frac{y^2}{2\eta_2^2}+2\sigma^2(y)+2\,\mathbb{Y}\,\sigma(y)}. \tag{27}$$

The third moment is

$$\mu_3^{(2)} = \Phi^3\int_{-\infty}^{\infty} d\mathcal{H}\,e^{3\mathcal{H}}\left(\mathbb{H}^3+3\mathbb{H}\right)\ p_{(1)}(\mathcal{H};\eta_2,\beta_2,g_2,\Phi_2). \tag{28}$$

The fourth moment is

$$\mu_4^{(2)} = \Phi^4 \int_{-\infty}^{\infty} d\mathcal{H}\, e^{4\mathcal{H}} \left(\mathbb{H}^4 + 6\mathbb{H}^2 + 3\right)\ p_{(1)}(\mathcal{H}; \eta_2, \beta_2, g_2, \Phi_2). \tag{29}$$

The simplified formula for $\mu_3^{(2)}$ and $\mu_4^{(2)}$ can be carried out easily using Maxima, but they are too complicated to be listed here.

Next, we discuss how to implement $p_{(2)}$ for numerical computation. In most cases, it is safe to assume $\eta_2 \ll 1$ or $\beta_2 \approx 0$. Then intuitively speaking, $p_{(1)}(\mathcal{H})$ has a peak near zero with the standard deviation of $\Phi_2 e^{\eta_2^2}$. Therefore we can employ the same numerical integration technique we used in the previous paper (Lihn 2008, SSRN: 1273087). That is, first locate where the the maximum probability is in the integrand (finding z_p); and second, determine approximately the "width" of integrand. Then simply integrate far beyond the "width" to obtain high precision. The caveat here is that, with the fat tails in $p_{(1)}$, the implementation has to cover a larger region than when it was Gaussian in order to provide good convergence.

An implementation for GNU Octave is released on the author's website (See http://www.skew-lognormal-cascade-distribution.org/impl/ for more details). As expected, $p_{(2)}$ calculation is very resource intensive. To run a SQP fit on a small computer, one must simplify the diff algorithm in some ways. For instance, in the SQP diff function, you can perform diff on the first four moments and a few important points in the pdf. It will slow you down significantly if you insist on diffing the entire pdf with many sampling points. The author has found that diffing the pdf by choosing several points between the peak and the standard deviation is sufficient to get a reasonably good fit.

4 Better Fit to the Heavy Tails in Stock Market Index

The first order fit to Dow's high frequence log-returns has a flaw of overestimating the standard deviation by 20% when the focus is to fit the skewness and kurtosis. This makes the fitted curve wider than the actual data and predicts too high of the peak probability density near the center. The first order distribution is not able to get all three moments right at the same time. The author attributes this discrepancy to the presence of the fat tails in the volatility term $\mathcal{H}$. Therefore it can not be solved without the second order distribution. The second order distribution introduces the concept of the fat tails into the volatility term $\mathcal{H}$ in the context of the higher order lognormal cascade. It turns out the second order distribution works much better for the time series data with very high kurtosis.

Figure 1 shows the fit of Dow's daily log-returns from 1928 to 2008 by the first order distribution $p_{(1)}$. The standard deviation of the fitted curve is 20% higher than that of the data when our focus is to fit the skewness and the kurtosis precisely. The fitted curve is not only wider than the data, but the peak probability density is also 20% higher than that is shown in the data (it is hard to visualize in the log(pdf) plot).

When we use the second order distribution $p_{(2)}$ to fit the same data, the result is much better as shown in Figure 2. The peak probability density and the first four moments are fitted precisely. We want to point out that $\eta_2 = 0.19$ indicates that there are observable fat tails in the volatility. A precise measure of the probability density opens up many applications in risk management. For instance, the daily VaR (value at risk) can be calculated from $p_{(2)}$. Risk averse investors can even perform stress test on potential large losses from the small probabilities in the fat tails.

At the end of Section 6, we will further point out that η_2 (0.19) in Figure 2 is close to the standard deviation η_ψ (0.25 in raw data, 0.21 in model) of the half normalized variance $\psi(t)/2$. This is an indication that the second order feature in Dow's log-returns could have an origin from the small tails in the underlying capital distribution. The two have profound connection.

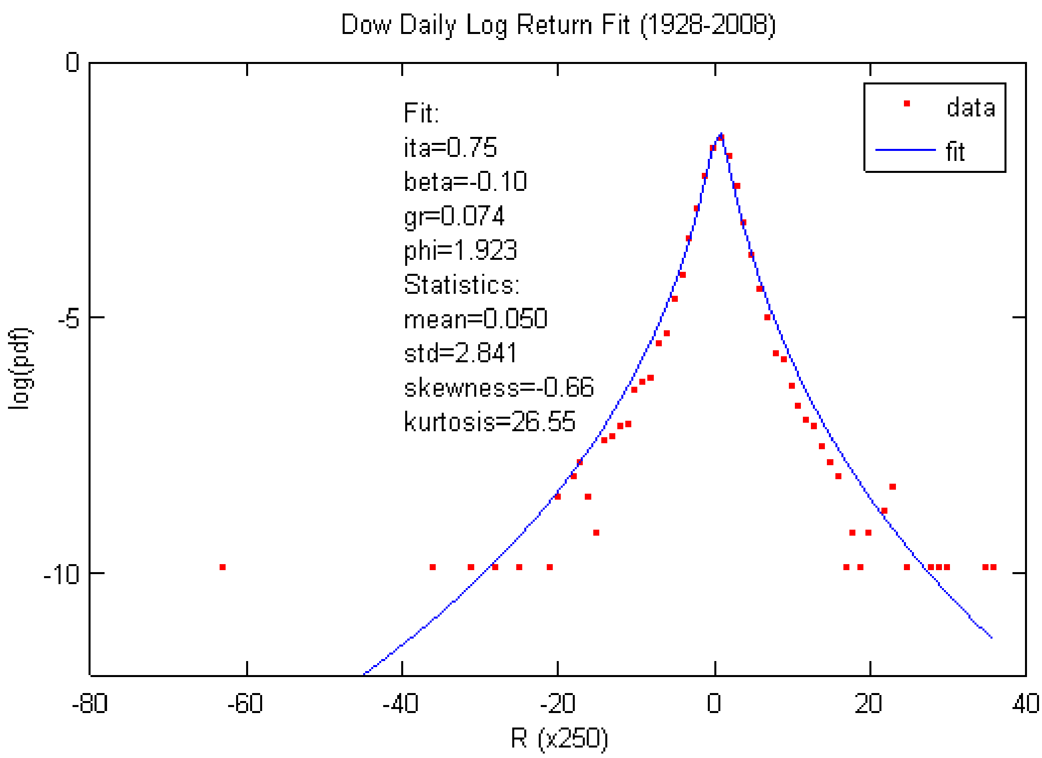

Figure 1: The First Order Fit to the Dow's Daily Log-Returns (1928-2008). The standard deviation is 20% higher than that of the data.

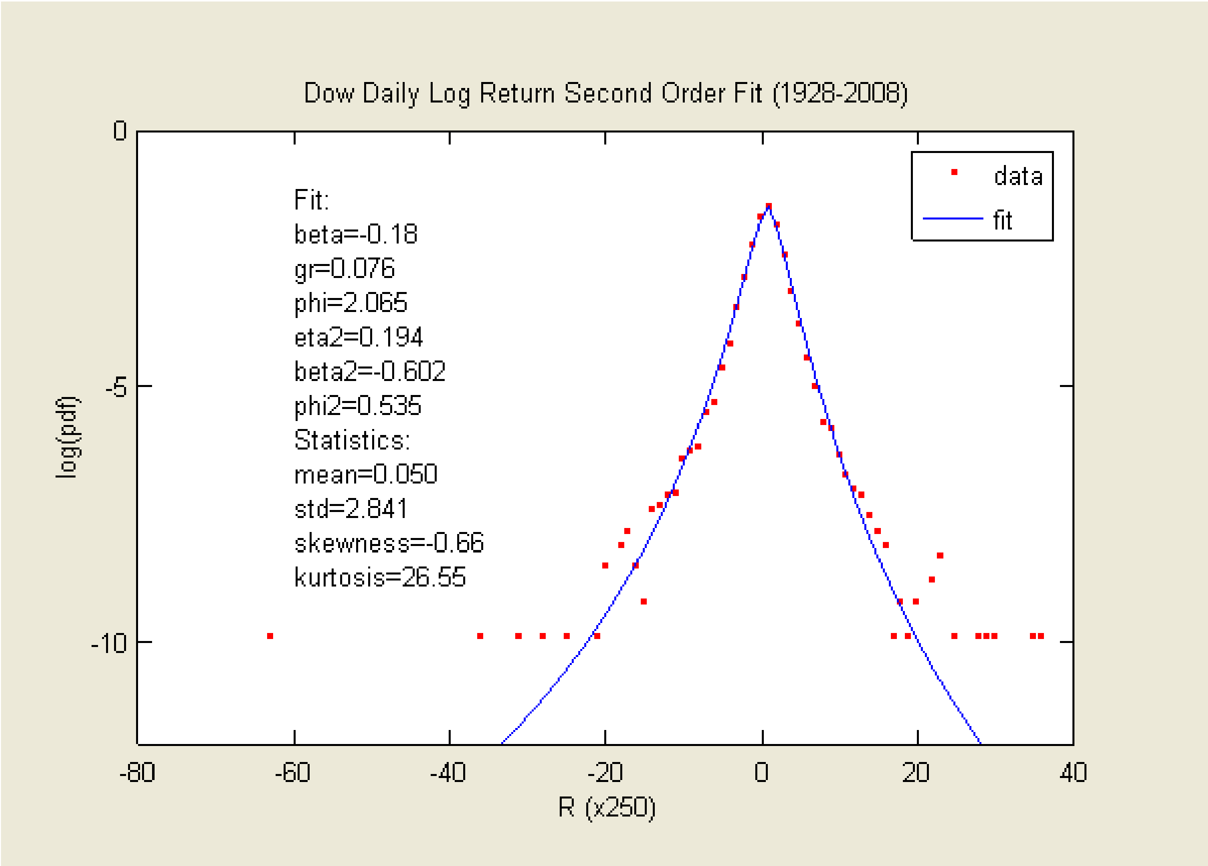

Figure 2: The Second Order Fit to the Dow's Daily Log-Returns (1928-2008). The first 4 moments and the peak probability density are precisely fitted. This is the successful demonstration of the second order distribution's capability to handle very heavy tails (high kurtosis) in high frequency financial time series.

5 A New Look at the Market Portfolio, Market Entropy, and Capital Distribution

In this section, we introduce a new way of studying the capital distribution (also known as the firm size distribution), the market portfolio (aka market index), and the market entropy. The best known broad stock market indices, such as S&P1500, or Russell 3000, typically cover a significant portion of publicly traded stocks and are weighted by market capitalization. Among them, the Dow Jones Wilshire 5000 Total Market Index represents the broadest index for the U.S. equity market. Only some smallest stocks (still a few thousands of them) are left out due to liquidity and research coverage reasons. However, ignoring the smallest stocks does not pose any problem to representativeness of the broad market indices because the differences of market capitalization between the largest stocks and the smallest stocks are so big (exponentially) that the weight of these smallest stocks is negligibly small in the broad market indices. These broad market indices are capable of "representing the entire market" by capturing the majority of market fluctuations.

In quantitative finance, the broad market index is studied via the market portfolio. As in Definition 1.2.3 of Fernholz 2002, the market portfolio is defined as the portfolio $\{X_i\}$ composed of all the stocks in the stock market, whose value Z is the total market capitalization of this market,

$$Z = \sum_{i=1}^{N} X_i, \tag{30}$$

where X_i is the market capitalization of the i-th stock and Σ_i sums over the entire stock market and N is the number of stocks in this market. Assuming all stocks have only one share outstanding, X_i is also the per-share price.

The weight of each stock is $\rho_i = X_i/Z$. It is trivial that $\sum_{i=1}^{N} \rho_i = 1$. Following Section 2.3 of Fernholz 2002, the market entropy is defined as

$$S = -\sum_{i=1}^{N} \rho_i \log(\rho_i). \tag{31}$$

We can define the auxiliary market entropy function $\widetilde{S}$ to bridge the relation between S and Z:

$$\widetilde{S} = -\sum_{i=1}^{N} X_i \log(X_i). \tag{32}$$

It is easy to obtain

$$S = \frac{\widetilde{S}}{Z} + \log(Z). \tag{33}$$

Traditionally, the distribution of ρ_i is presented in the log-log plot of (i, ρ_i), in which $\{\rho_i\}$ is sorted in a descending order; therefore, the suffix i represents the rank in weight, and ρ_1 is the weight of the largest stock in the market. This plot is called "rank distribution". One such plot from the US stock market as of 12/31/2008 is shown in Figure 3. As also shown in Figure 5.1 of Fernholz 2002 and discussion thereof, the curves of the rank distribution since 1929 have a similar shape. This indicates some type of stability is present in such distribution and it is likely such distribution follows a predictable dynamics. Some have argued that the log-log plot should follow the Pareto distribution for the largest stocks, which is a straight line with a negative slope. But in reality the linear Pareto distribution is only a convenient concept. The rank distribution is generally concave and it deviates further from the ideal linear distribution as the rank increases. Banner et al (Banner, Fernholz, Karatzas, 2005) have proposed the Atlas model and a "first order model" to explain the rank distribution. (See also Section 5.5 of Fernholz 2002.)

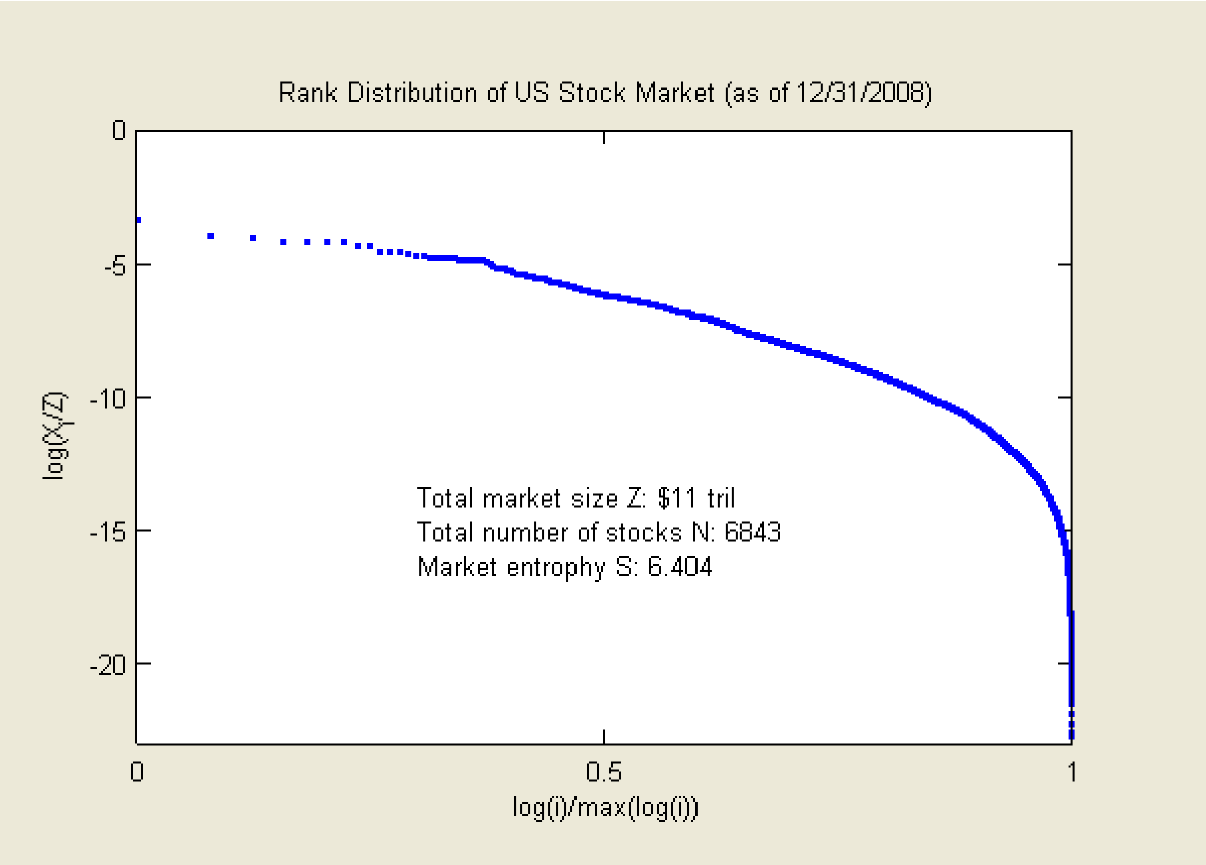

Figure 3: Rank Distribution of the U.S. Stock Market as of 12/31/2008.

In this paper, we will take a different approach to understand the rank distribution. Now since the market contains a large collection of stocks, such as in the U.S. stock market, it is reasonable to assume $\{\rho_i\}$ can be approximated by a continuous distribution density, for instance, $\mathbb{P}\,(\mathcal{H};\mathcal{H}_c)\;d\mathcal{H}$ where $\mathcal{H}$ is **the logarithm of the market capitalization** (abbreviated as **log-market cap**) substracted by the quantity $\mathcal{H}_c$. The unity equation $\sum_{i=1}^{N}\rho_i = 1$ is transformed to $\int_{-\infty}^{\infty}\mathbb{P}\,(\mathcal{H};\mathcal{H}_c)\;d\mathcal{H} = 1$. The quantity $\mathcal{H}_c$ is defined as the mean of the log-market cap distribution, such that $\int_{-\infty}^{\infty}\mathcal{H}\,\mathbb{P}(\mathcal{H};\mathcal{H}_c)\,d\mathcal{H} = 0$. Therefore, the market cap X_i is transformed to $e^{\mathcal{H}+\mathcal{H}_c}$. (The purpose of $\mathcal{H}_c$ will be shown shortly.) The new concept we want to introduce is following.

Equation (30) can be written in a continuous notation as

$$Z \approx N\int_{-\infty}^{\infty}\mathbb{P}(\mathcal{H};\mathcal{H}_c)\,d\mathcal{H}\,e^{\mathcal{H}+\mathcal{H}_c}. \tag{34}$$

This is trivial enough. What is nontrivial is to identify that $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ follows one of the skew lognormal cascade distributions in our higher order framework. For convenience sake, we assume $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ is the $(n-1)$-th order lognormal cascade distribution with zero mean (Definition 1.2). The requirement of zero mean is satisfied naturally by the introduction of $\mathcal{H}_c$. Then, according to Equation (16), Z in Equation (34) is the first moment $\mu_1^{(n)}$ of the returns $d\chi/dt$ of an abstract process χ, whose returns $d\chi/dt$ is lifted to an n-th order lognormal cascade distribution with $\beta_{(n)} = 0$, $g_{(n)} = 1$, and $\Phi_{(n)} = Ne^{\mathcal{H}_c}$. We can call χ the "total market" process. Although this process is abstract, this is a quick way to see that the market portfolio Z should exhibit fat tails.

Let's also transform the auxiliary market entropy. Equation (31) can be written in a continuous notation as

$$\widetilde{S} \approx -N \int_{-\infty}^{\infty} \mathbb{P}(\mathcal{H};\mathcal{H}_c)\, d\mathcal{H}\, e^{\mathcal{H}+\mathcal{H}_c}(\mathcal{H}+\mathcal{H}_c). \tag{35}$$

The reader should also notice that, according to Equation (16), $\widetilde{S}$ in Equation (35) is the first moment $\widetilde{\mu}_1^{(n)}$ of the returns $d\widetilde{\chi}/dt$ of another abstract process $\widetilde{\chi}$, whose returns $d\widetilde{\chi}/dt$ is lifted to an n-th order lognormal cascade distribution with $\beta_{(n)} = 1$, $g_{(n)} = \mathcal{H}_c$, and $\Phi_{(n)} = -Ne^{\mathcal{H}_c}$. We can call $\widetilde{\chi}$ the "auxillary market entropy" process.

Let's now verify what $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ really looks like in reality and how good it fits our proposition of a skew lognormal cascade distribution if any. Table 1 lists the mean, standard deviation, skewness, and kurtosis of the log-market cap distribution of the U.S. stock market for each year-end from 1991 to 2008. Market cap is in the unit of million \$USD. In our notation, the mean of the distribution is $\mathcal{H}_c$ whose unit is log(million \$USD).The standard deviation η_c is a measurement of the "width" of the distribution (in a naive way). N is the number of stocks in our data set. Z is the total market capitalization in the data set. Due to data quality issue, some stocks are filtered out when their calculated market capitalizations are too out of range (overestimated) compared to their company fundamentals.

Table 1: The mean, standard deviation, skewness, and kurtosis of the log-market cap distribution of the US stock market for each year-end from 1991 to 2008. Market cap is in the unit of million $USD. N is the number of stocks in our data set. Z is the total market capitalization in the data set.

Year	mean ($\mathcal{H}_c$)	std (η_c)	skewness	kurtosis	N	Z (tril $)
1990	18.67	2.311	0.0059	-0.3070	1871	2.514
1991	18.96	2.312	-0.0888	-0.2458	1994	3.355
1992	19.09	2.228	-0.1068	-0.1013	2174	3.773
1993	19.18	2.128	-0.0556	-0.0469	2557	4.366
1994	19.04	2.126	0.0059	-0.1651	2801	4.261
1995	19.18	2.163	-0.0318	-0.0378	3055	5.808
1996	19.21	2.185	-0.0465	0.0389	3407	7.096
1997	19.29	2.225	-0.0173	0.1006	3666	9.568
1998	19.02	2.372	-0.0967	0.5757	4134	11.59
1999	19.11	2.334	0.0809	0.2399	4697	15.69
2000	18.79	2.483	0.0572	-0.0251	4979	14.14
2001	18.73	2.545	-0.1890	0.1782	5148	12.40
2002	18.47	2.548	-0.2890	0.6192	5343	9.800
2003	18.96	2.432	-0.4001	1.094	5479	12.77
2004	19.16	2.334	-0.3325	1.058	5757	14.29
2005	19.18	2.309	-0.2216	0.6703	5980	15.03
2006	19.29	2.332	-0.2529	0.5257	6241	17.44
2007	19.12	2.407	-0.3070	0.8672	6725	18.30
2008	18.35	2.621	-0.3858	0.7038	6843	11.40

We can see from Table 1 that the skewness and kurtosis in the data is relatively small (but not zero, and especially so since 2002). It is reasonable to expect the log-market cap distribution is roughly a normal distribution. Furthermore, the small skewness and kurtosis can be easily captured by a first order skew lognormal distribution, which we will discuss later. Now let's first examine some important long term trends in the data.

Figure 4 shows the long term trend of $\log(N)$, $\mathcal{H}_c$, and η_c^2 from the same data set behind Table 1. The data set is retrieved at every month-end from 1990 to 2008. They are mean detrended in order to emphasize the fluctuations over time. $\mathcal{H}_c$ is pretty much rising and falling with the bull-bear market cycles, while its average level did not change much in the past 20 years. The most obvious trend is that, apart from small variations,

N is almost linearly increasing with time,

$$N(t) = 1453 + 289 \times (t - 1990), \tag{36}$$

in which t is in the unit of year. However, there is an important discovery here: The baseline of $\eta_c(t)^2$ is increasing slowly, following the trend line of $\log(N)$, which is slightly concave. We can define the **"normalized variance"** $\psi(t)$ for the log-market cap distribution

$$\psi(t) = \eta_c(t)^2 - \log(N) + \psi_c, \tag{37}$$

where ψ_c is a constant called the **"characteristic average normalized variance"** such that $\psi(t)$ becomes a mean-reverting quantity, that is $\langle\psi(t)\rangle = 0$. Numerically ψ_c is about 3.0 in our data set. The normalized variance $\psi(t)$ plays a critical role in our later discussion of the origin of the fat tails in the market fluctuations. (Curious readers can peek at Figure 8.)

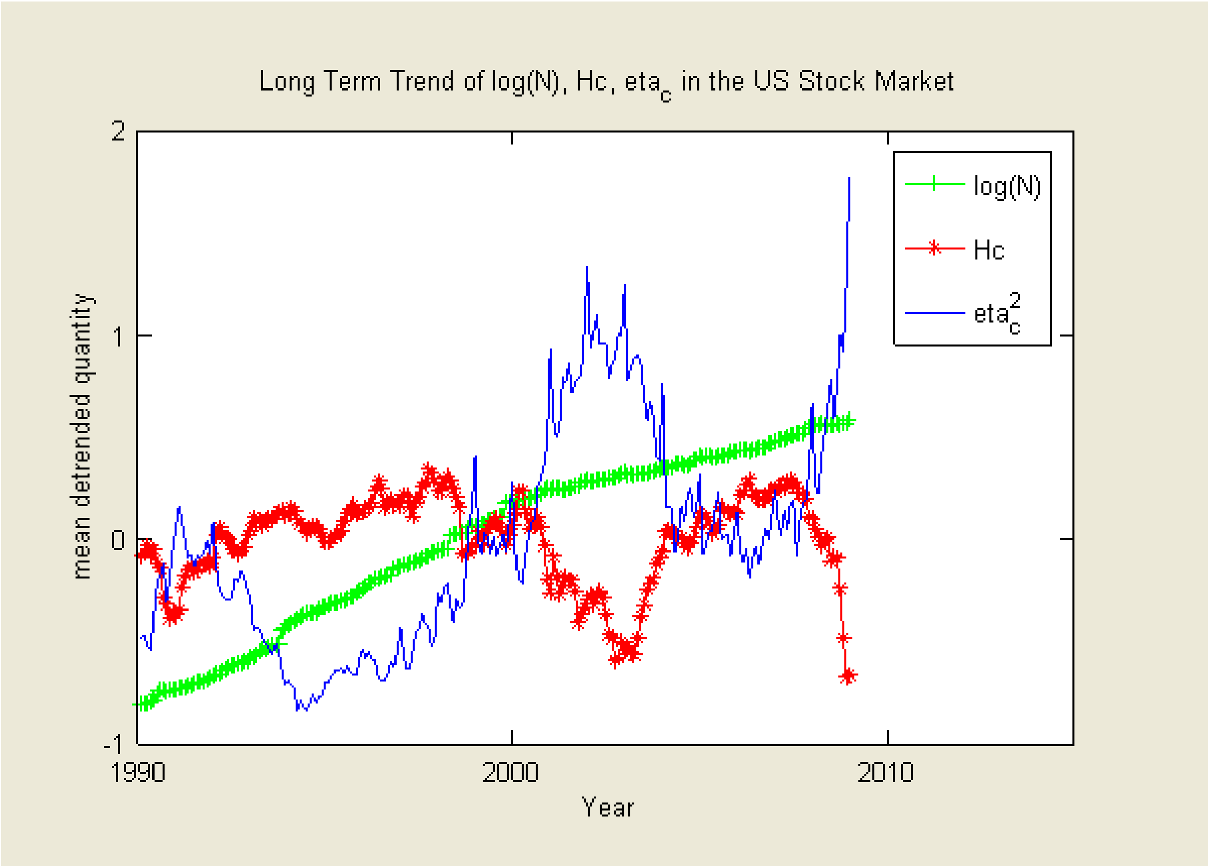

Figure 4: The Long Term Trend of $\log(N)$, $\mathcal{H}_c$, and η_c^2 in the U.S. Stock Market. The data set is retrieved at every month-end from 1990 to 2008. They are mean detrended in order to emphasize the fluctuations over time. The important discovery here is: The baseline of $\eta_c(t)^2$ is increasing slowly, following the trend line of $\log(N)$, from which a mean-reverting quantity called the **"normalized variance"** $\psi(t)$ is extracted. $\mathcal{H}_c$ is pretty much rising and falling with the bull-bear market cycles, while its average level did not change much in the past 20 years.

Let's now examine how well $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ fits to a normal distribution (which is also the zero-th order lognormal cascade distribution $p_{(0)}(\mathcal{H};\eta_c)$) and a first order lognormal cascade distribution $p_{(1)}(\mathcal{H};\eta,\beta,g,\Phi)$. Figure 5 shows the fit to a normal distribution with 12/31/2008 data. Green line is the best fit of a normal distribution, but the indicated total market value is 30% higher and the market entropy is 6% lower. Blue line indicates the theoretical curve derived from Equation (39) [5] , which produces correct total market value and the market entropy, but the "implied distribution" deviates from the raw data. [4] Overall, we consider a normal distribution describes the data roughly okay.

[4]Since both methods capture different aspects of the data, this is the reason why we present the normalized variance from both methods (See Figure 8) as a compromise when we discuss the origin of

On the other hand, $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ can be very precisely fitted by the first order lognormal cascade distribution $p_{(1)}(\mathcal{H};\eta,\beta,g,\Phi)$ where $g = -\beta\,\eta^2$, as shown in Figure 6 , with the same 12/31/2008 data. We know that the first order skew lognormal distribution works very well for the distributions with kurtosis < 6.0. Thus, undoubtly the first order fit is able to accurately characterize the standard deviation, skewness and kurtosis; and produce the total market value and the market entropy with smaller than 10^{-4} error. It is a much superior fit. Careful readers can verify the fit via Equation (21). The small η (0.16) also indicates that $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ does not deviate much from a normal distribution.

Since $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ is either roughly a $p_{(0)}(\mathcal{H})$ or more precisely a $p_{(1)}(\mathcal{H})$, it indicates the returns of the total market process χ have significant fat tails since its distribution must be lifted one order higher than that of $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$.

Our observation of a normal distribution in the log-market cap distribution is equivalent to a lognormal distribution in market cap. Such observation is in agreement with the recent work of Cabral and Mata (Cabral and Mata 2003) on the firm size distribution. Furthermore, Growiec et al (Growiec, Pammolli, Riccaboni, and Stanley, 2008) have studied the tail structure of the firm size distribution and concluded that it is a lognormal distribution multiplied by a stretching factor which can lead to a Pareto upper tail. Whether such a stretching factor is the same as our approach of the first order lognormal cascade distribution remains to be verified in the future. But by and large, our use of a normal distribution in log-market cap is a reasonable approximation.

---

the fat tails when assuming a normally distributed market cap distribution.

[5] Due to the exponential weight in Z and $\widetilde{S}$, they are very sensitive to the right tail, but far less sensitive to the left tail. This is expected due to the $e^{\mathcal{H}}$ terms in the higher order distributions. Thus one should pay more attention on the behavior of the right tail region, but be somewhat ignorant on how exactly the fitted curve or theoretical curve behave in the left tail region.

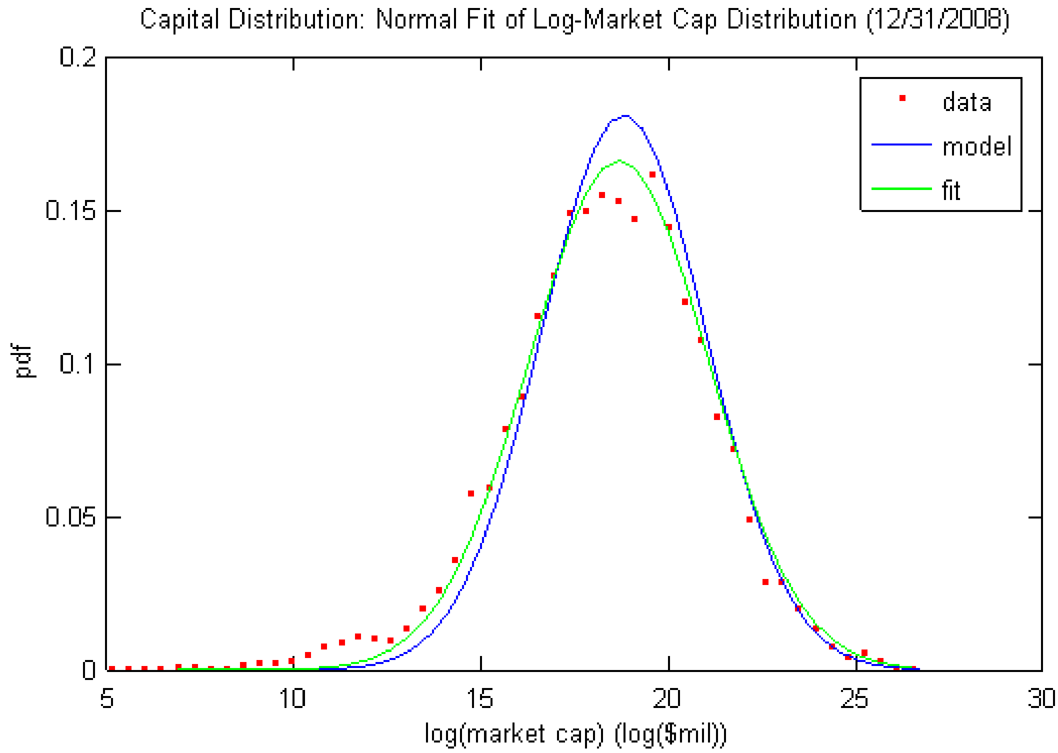

Figure 5: Normal fit of Log-Market Cap Distribution as of 12/31/2008, $p_{(0)}(\mathcal{H};\eta_c)$. $\mathcal{H}_c$ is added back to the $\mathcal{H}$ axis of $p_{(0)}(\mathcal{H};\eta_c)$ to match the raw data's distribution. Green line is the best fit of a normal distribution, but the indicated total market value is 30% higher and entropy is 6% lower. Blue line indicates the theoretical curve derived from the simple model, i.e. Equation (39). Note that an accurate description of the right tail is much more important than the left tail due to the exponential weight in Z and $\widetilde{S}$.

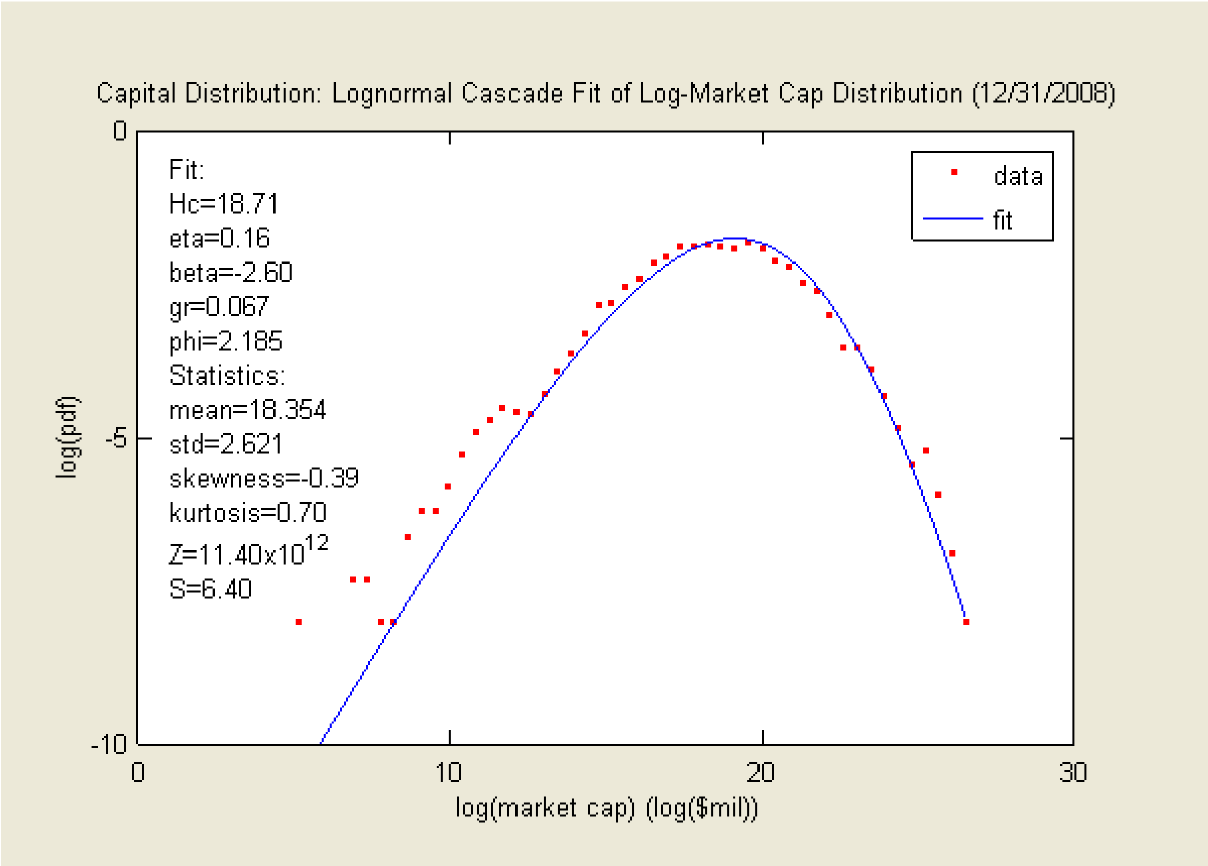

Figure 6: First order fit of Log-Market Cap Distribution as of 12/31/2008. The first order lognormal cascade distribution $p_{(1)}(\mathcal{H};\eta,\beta,g,\Phi)$ is able to accurately characterize the standard deviation, skewness, and kurtosis; and produce the total market value and the market entropy with smaller than 10^{-4} error. $\mathcal{H}_c$ is added back to the $\mathcal{H}$ axis of $p_{(1)}(\mathcal{H};\eta,\beta,g,\Phi)$ to match the raw data's distribution. Note that an accurate description of the right tail is much more important than the left tail due to the exponential weight in Z and $\widetilde{S}$.

Next, we are going to introduce a simple model based on the assumption that the capital distribution is a normal distribution in the log-market cap space. We shall call this model "**normally distributed log-capitalization model**" in this paper. We will continue to enrich this model in this paper, which allows us to gain some insight on the total market portfolio Z, the auxiliary entropy $\widetilde{S}$, and eventually the origin of fat tails. The model here is static, which will be expanded into a stochastic one in next section.

Now we assume $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ is a normal distribution:

$$\mathbb{P}(\mathcal{H};\mathcal{H}_c) = \frac{1}{\sqrt{2\pi}\,\eta_c}\, e^{-\frac{\mathcal{H}^2}{2\eta_c^2}}, \tag{38}$$

in which $\mathcal{H}_c$ is the mean log-market cap and η_c^2 is the variance of the log-market cap distribution. The integral of Z and $\widetilde{S}$ (and therefore S) can be carried out exactly:

$$\begin{gathered} Z = N\, e^{\mathcal{H}_c + \frac{\eta_c^2}{2}}, \\ \widetilde{S} = -\left(\mathcal{H}_c + \eta_c^2\right)\, Z \text{ and } S = \log(N) - \tfrac{\eta_c^2}{2}. \end{gathered} \tag{39}$$

If N, $\mathcal{H}_c$ and η_c are given, Z and S can be calculated. Inversely, since in reality Z, N, and S are known, we can "derive" the "implied value" of $\mathcal{H}_c$ and η_c from the model as

$$\begin{gathered} \mathcal{H}_c = \log(Z) + S - 2\log(N), \\ \tfrac{\eta_c^2}{2} = \log(N) - S\,. \end{gathered} \tag{40}$$

Such method of obtaining $\mathcal{H}_c$ and η_c are illustrated in Figure 5, shown as the blue solid line, "Model".

It is interesting to note that, if our proposition on the mean reverting process $\psi(t)$ is correct (See Equation (37)), then strictly speaking, S is not mean reverting, but increases with time in the long term (or, more accurately, increases with N). We can define the **"normalized market entropy"** S_ψ as $S - \log(N)/2$, which is the truly mean reverting quantity with regard to the market entropy that won't increase with time in the long term (apart from a constant).

The concept of the normalized variance $\psi(t)$ implies **an important scaling law between** $\log(N)$ **and** $\log(Z)$. By combining Equations (37) and (39), we have

$$\log(Z) = \frac{3}{2}\log(N) + (\mathcal{H}_c - \frac{\psi_c - \psi(t)}{2}). \tag{41}$$

From the discussion around Figure 4, we recognize that $\langle\psi(t)\rangle = 0$, ψ_c and $\langle\mathcal{H}_c\rangle$ are constants over the long term. We arrive that $\log(Z)$ and $\log(N)$ have a linear relation with the slope of ˜1.5. Figure 7 shows such scaling law in the US stock market. The linear fit gives the scaling exponent of 1.438, very close to our theoretical value of 1.5 in Equation (41). This is an amazing scaling law. The total market value created (Z) is more than proportional to the number of companies put into the market (N). There is an additional synergy effect of the free market that Adam Smith was talking about in "Wealth of Nations". As Warren Buffett always advocates, the free market system in the United States really works well in creating wealth.

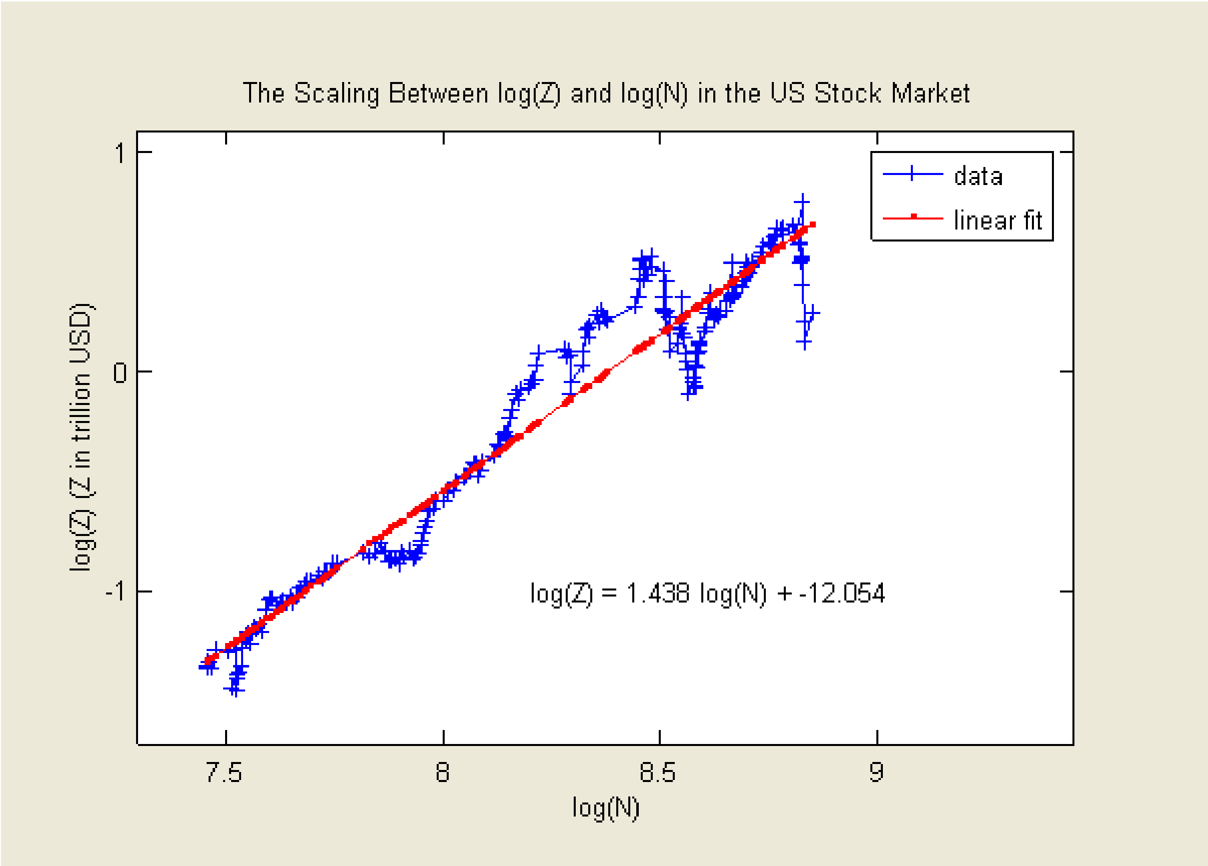

Figure 7: The Scaling Law Between $\log(Z)$ and $\log(N)$ in the US Stock Market. The linear fit gives the scaling exponent of 1.438, very close to the theoretical value of 1.5 in Equation (41).

We also notice that $\log(Z)$ is affected by both $\mathcal{H}_c$ and $\eta_c(t)^2/2$ (linearly), yet S is affected by $\eta_c(t)^2/2$ only and in an opposite direction. That is, the market index can fluctuate due to the movement of the mean and/or the changing shape in the underlying capital distribution. However, the market entropy won't change unless the shape changes. Furthermore, if $\eta_c(t)$ increases, so does Z, but the market diversity, as indicated by S, decreases. This is the consequence of large cap stocks dominating the market. Such opposite movements caused by $\eta_c(t)$ have a profound influence on market volatility as we will show in next section.

It would also be interesting to see how these relations hold when $\mathbb{P}(\mathcal{H};\mathcal{H}_c)$ is characterized by the more accurate first order distribution $p_{(1)}(\mathcal{H};\eta,\beta,g,\Phi)$ as shown in Figure 6. However, due to the significant mathematical complexity involved, this is left for future research.

6 The Origin of Fat Tails in The Stock Market

So far we have studied the static properties of the capital distribution. Let's turn our attention to the stochastic properties of the capital distribution. Our goal in this section is to demonstrate that the fluctuations of the market portfolio's value, $d\log Z(t)/dt$, will have fat tails due to the presence of the mean reverting process $\psi(t)$. A stochastic equation will be derived to establish the relation between $Z(t)$ and $\psi(t)$, which is exactly the lognormal cascade equation in our theory.

Our starting point is Corollary 1.1.6 of Fernholz 2002:

$$\begin{aligned} d\log Z(t) &= \sum_{i=1}^{N} \rho_i(t)\ d\log X_i(t) + \gamma^*(t)\,dt, \\ d\log X_i(t) &= \gamma_i(t)\,dt + \Sigma_\nu\, \xi_{i\nu}(t)\,d\,W_\nu(t), \end{aligned} \tag{42}$$

in which $\gamma^*(t)$ is excess growth rate of the market portfolio, $\gamma_i(t)$ is the growth rate of the stock X_i, and $\xi_{i\nu}(t)$ is the volatility matrix. We need to make a few assumptions: First, our interest is in the fat tails. In whatever time series we look at, the tail distribution stretches far beyond the small "long-term average return". Therefore, the γ^* and γ_i terms are significantly smaller than the stochastic processes. For instance, this is very obvious in Figures 1 and 2 where Δt is one day and the range of log-returns R (-20 to +20) is more than 400 times of the mean (0.05) (See http://www.skew-lognormal-cascade-distribution.org/apps/ for more examples). Thus we can safely drop the γ^* and γ_i terms for our purpose.

The second assumption is that, in our continuous notation, when we arrange ρ_i into the rank distribution, we also rewrite $\Sigma_\nu\, \xi_{i\nu}(t)\,d\,W_\nu(t) \to \sigma_i(t)\,d\,W_i(t)$. That is, the long-range correlation of volatility between stocks is simplified in such rank presentation. Such assumption is similar to the rank-based variance $\sigma^2_{k:k+1}$ in Section 5.5 of Fernholz 2002. In the "first order model" discussion thereof (Definition 5.5.1), $\sigma^2_{k:k+1}$ is further simplifed to σ^2_k by assuming the stocks are independent. Apparently, we also made a similar claim here.

Next we apply the additive rule of the normal distributions: When X and Y are normal processes, the combined process $Q = X + Y$ has the variance of $\sigma(Q)^2 = \sigma(X)^2 + \sigma(Y)^2$. Therefore, we can simplify Equation (42) to $d\log Z(t) = \sigma_{\log\,Z}(t)\,d\,W(t)$ in which $\sigma_{\log\,Z}(t)$ is the volatility of the market index and

$$\sigma_{\log\,Z}(t)^2 = \sum_{i=1}^{N} \left(\rho_i(t)\,\sigma_i(t)\right)^2. \tag{43}$$

The third assumption is to further simplify $\sigma_i(t)$. We know that the volatility of the small cap stocks is larger than that of the large cap stocks so σ_i should have a dependency on $\mathcal{H}$. We can write $\sigma_i(t) \sim \sigma(\mathcal{H}, t)$, which can be viewed as the net volatility that stocks generally have in their ability to affect the capital distribution near their log-market cap neighborhood. This allows us to write Equation (43) in a continuous fashion:

$$\sigma_{\log\,Z}(t)^2 \approx N \int_{-\infty}^{\infty} \mathbb{P}(\mathcal{H}, t\,; \mathcal{H}_c)\,d\mathcal{H} \left(\frac{e^{\mathcal{H}+\mathcal{H}_c(t)}}{Z(t)}\right)^2 \sigma^2(\mathcal{H}, t). \tag{44}$$

Equation (44) is similar to Equation 5.5.6 of Fernholz 2002 written in the continuous notation (except we choose to ignore the growth rate term γ_μ^* described in Equation 5.5.7 there).

Next, we will study Equation (44) in the **normally distributed log-capitalization model**. In this simple model, we observe that since $\rho_i \sim e^{\mathcal{H}+\mathcal{H}_c}/Z$, the variation of $\sigma(\mathcal{H},t)$ on $\mathcal{H}$ is insignificant in the context of Equation (44). Therefore, we choose to let $\sigma(\mathcal{H})$ be a constant σ_c across $\mathcal{H}$ as a first order approximation. σ_c can be viewed as the characteristic average volatility that stocks have across the market. Such σ_c is similar to the constant volatility in the Atlas model (Example 5.3.3 of Fernholz 2002).

We also want to express an alternative hypothesis that can reach the same effect, that is,$\mathbb{P}(\mathcal{H},t\,;\mathcal{H}_c)\,\sigma(\mathcal{H},t)$ preserves the normality in its distribution. We know from Figure 6 and related discussion that $\mathbb{P}(\mathcal{H},t\,;\mathcal{H}_c)$ has small tails all the time. And the presence of the right tail will affect $\sigma_{\log Z}(t)$ in a great deal due to the $e^{\mathcal{H}+\mathcal{H}_c(t)}$ term. However, we also know that the volatility $\sigma(\mathcal{H},t)$ is smaller for large stocks (and larger for small stocks). Thus we can expect that the right tail in $\mathbb{P}(\mathcal{H},t\,;\mathcal{H}_c)$ is cancelled out by the decrease of $\sigma(\mathcal{H},t)$ in large $\mathcal{H}$, resulting in a more or less normally distributed $\mathbb{P}(\mathcal{H},t\,;\mathcal{H}_c)\,\sigma(\mathcal{H},t)$.

Furthermore, we assume the time dependency of $\mathbb{P}(\mathcal{H},t\,;\mathcal{H}_c)$ is characterized by $\eta_c(t)$ and $\mathcal{H}_c(t)$; that is,

$$\mathbb{P}(\mathcal{H},t\,;\mathcal{H}_c) \to \mathbb{P}(\mathcal{H};\mathcal{H}_c(t),\eta_c(t)) = \frac{1}{\sqrt{2\pi}\,\eta_c(t)}\, e^{-\frac{\mathcal{H}^2}{2\eta_c(t)^2}}. \tag{45}$$

We can then carry out $\sigma_{\log Z}(t)$ in a closed form:

$$\sigma_{\log Z}(t) = \frac{\sigma_c}{\sqrt{N}}\, e^{\frac{\eta_c(t)^2}{2}} = \sigma_c\, e^{-\frac{\psi_c}{2}} e^{\frac{\psi(t)}{2}}, \tag{46}$$

and therefore,

$$d\log Z(t) = \sigma_c\, e^{-\frac{\psi_c}{2}} e^{\frac{\psi(t)}{2}}\, d\,W(t). \tag{47}$$

This is an important result: Equation (47) resembles the form of Equation (1). We assume σ_c is a constant term (that is, remove its time dependency for the time being) [6] . From Equation (37) and discussion thereof, we know ψ_c is also a constant and $\psi(t)/2$ is a slow varying, mean reverting process. If $\psi(t)/2$ produces a skew lognormal cascade distribution (including a normal distribution), than $d\log Z(t)/dt$ **is lifted** to a lognormal cascade distribution one order higher than that of $\psi(t)/2$. Therefore, $d\log Z(t)/dt$ will exhibit fat tails and how heavy the tails are is directly driven by the changes of the normalized variance, $\psi(t)$. To examine $\psi(t)$ as a mean reverting process, it is interesting to relate $\psi(t)/2$ to the normalized market entropy, $(-\,S(t)+\log(N)/2)$. It is well known that $S(t)$ is cyclic over long period of time, and the normalized market entropy qualifies our requirement of being a "slow varying, mean reverting" process ($S(t)$ has been calculated since 1939 using CRSP database. See Figure 6.2 of Fernholz 2002). Therefore, $d\log Z(t)/dt$ exhibits fat tails according to Equation (46) and the cyclic nature of $\psi(t)$.

[6] By assuming σ_c a constant, it can actually be calculated from other observeables. $\log(\sigma_c) = \psi_c/2 + \langle\log(\sigma_{\log Z}(t))\rangle$. Since $\langle\log(\sigma_{\log Z}(t))\rangle = -2.95$, we obtain the monthly volatility of $\sigma_c = 0.24$,

Does Equation (46) really describe what is going on in our financial market? Let's compare our model to the market data. In the case of the U.S. stock market, $\sigma_{\log Z}(t)$ can roughly equate to the S&P500 volatility index VIX, which is the annualized volatility that investors expect for the next one month period. The monthly volatility of the market can be calculated by $\frac{\text{VIX}}{100}/\sqrt{12}$. Figure 8 shows the mean-detrended logarithm [7] of the monthly volatility from VIX (blue cross solid line) compared to half of the normalized variance, $\psi(t)/2$, for each month from 01/1990 to 12/2008. The logarithm of monthly volatility is a 12-month moving average. The red dotted solid line is based on the normalized variance calculated directly from η_c and N in the data set. The green solid line is the "implied normalized variance" derived from the market entropy, that is, mean detrended $(-\,S + \log(N)/2)$. We found that the evolution of the normalized variance matches strikingly well with the evolution of market volatility and the bull-bear market cycles over the past 20 years. And due to the presence of a negative skewness in log-returns of the stock market index, the high volatility periods typically relate to the bear market cycles (See Lihn 2008, SSRN 1149142 for more details). In fact, by mimicing Equation (1), we can write down **the full emperical dynamic equation for the market index in terms of the normalized variance of the log-market cap distribution**:

$$d_t \log Z(t) = \Phi \cdot\, e^{\frac{\psi(t)}{2}} \left[d_t W(t) + (\,\beta \cdot \frac{\psi(t)}{2} + g)\, dt \right]. \tag{48}$$

where β is the negative default coefficient and g is the positive growth rate. Assume $\psi(t)$ is distributed normally, $\psi(t)/2 \sim N\left(0, \eta_\psi^2\right)$ where η_ψ is the standard deviation of $\psi(t)/2$'s distribution over time. In order to maintain a long-term positive growth in the market, it is required that $g > -\beta\, \eta_\psi^2$. From our monthly data set, we can calculate $\eta_\psi = 0.214$ and the VIX data $\langle \log(\sigma_{\log Z}(t)) \rangle$ implies the standard deviation of Dow's monthly log-returns $\langle \sigma_{\log Z} \rangle = 0.63$.

We can validate our model by comparing these emperical values (η_ψ and $\langle \sigma_{\log Z} \rangle$) to the fitted values on Dow's monthly log-returns from 1990 to 2008, using the first order lognormal cascade distribution. The first order fit, as shown in Figure 9 is very accurate on the first four moments. We can see that the fitted η_ψ (0.25) is within 20% range to our model prediction (0.214). The fitted standard deviation $\sqrt{\mu_2^{(1)}}$ (0.503) is also with 20% range compared to the standard deviation $\langle \sigma_{\log Z} \rangle$ implied by the VIX data (0.63). Consider that 20 years are a short history in terms of the economic cycles [8] , both numbers coming out within 20% range is fairly good. This is a strong validation to our model.

---

which is annualized to 83%. This is approximately the volatility expected of stocks in general.

[7] The meaning of mean-detrended logarithm is explained as following. Suppose we have a function $f(x) = C\, e^{a \sin(b\,x)}$, then the mean-detrended logarithm of $f(x)$ is simply $a \sin(b\,x)$, which is the mean reverting process in $f(x)$ that we are most interested in.

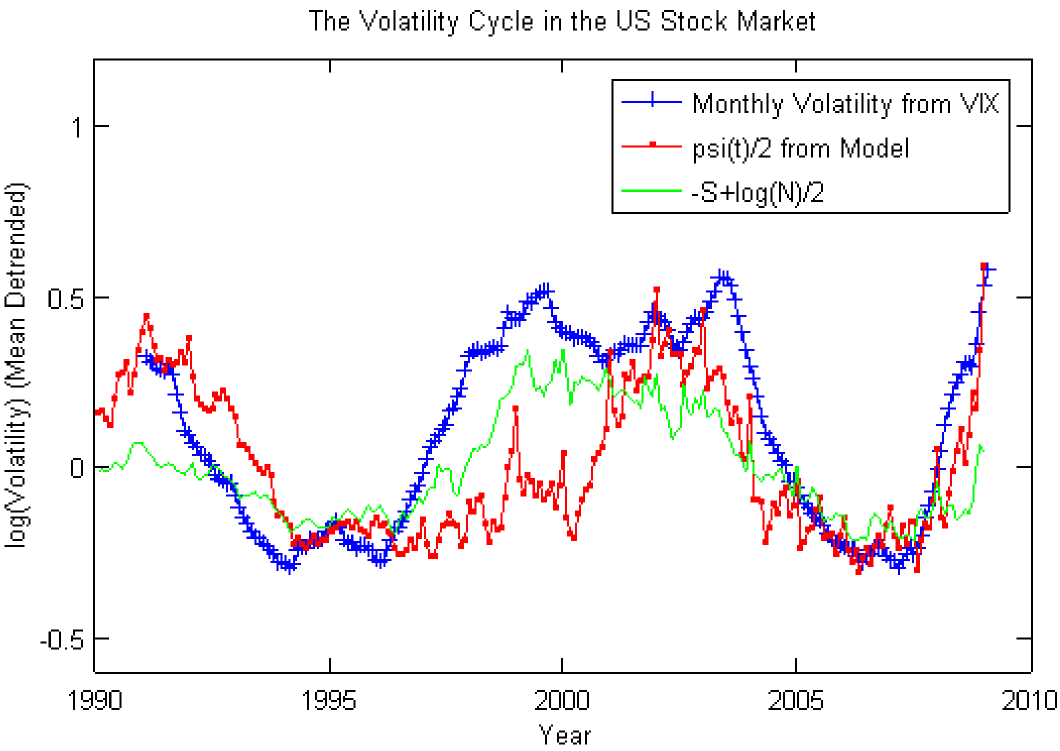

Figure 8: The mean-detrended logarithm of monthly volatility of S&P500 (VIX) (cross-blue line) compared to half of the normalized variance $\psi(t)/2$ for each month from 01/1990 to 12/2008. The logarithm of monthly volatility is a 12-month moving average. The red dotted solid line is the normalized variance calculated directly from η_c and N in the data set. The green solid line is the "implied normalized variance" derived from the market entropy, that is, mean detrened $(-S + \log(N)/2)$. The evolution of the normalized variance matches strikingly well with the evolution of market volatility and the bull-bear market cycles over the past 20 years.

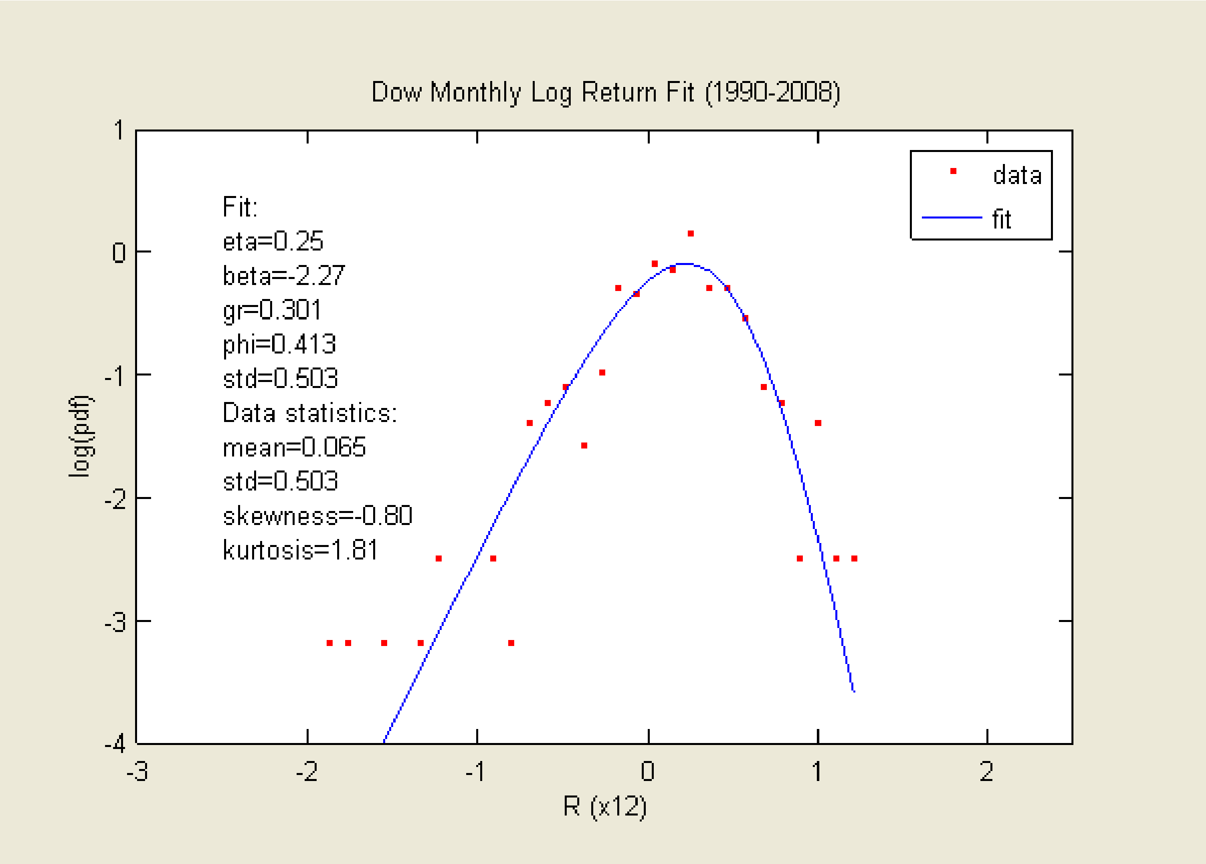

Figure 9: The first order lognormal cascade distribution fit on Dow's monthly log-returns from 1990 to 2008. The first order fit is very accurate on the first four moments. We can see that the fitted η_ψ (eta = 0.25) and $\sqrt{\mu_2^{(1)}}$ (std = 0.503) are close (within 20% range) to our model predictions: Our capital distribution data implies $\eta_\psi = 0.214$; and the VIX data $\langle \log(\sigma_{\log Z}(t)) \rangle$ implies the standard deviation of Dow's monthly log-returns $\langle \sigma_{\log Z} \rangle = 0.63$. This is a strong validation to our model.

Therefore, **the contraction and expansion of the log-market cap distribution is the fundamnetal driving force of the bull-bear market cycles**. We see that the normalized variance peaked during the 2002 bear market, and again it was moving up sharply in 2008-2009 as the stock market is heading towards the worst bear market in 50 years. On the other hand, when the normalized variance was decreasing, the market volatility also decreased, and the stock market enjoyed its bull runs, such as the years

[8] A note on the 20-year history: The kurtosis of the 20-year data is 5 times smaller than the kurtosis of the 80-year data (which is 11). Since kurtosis relates to η_ψ directly and η_ψ is the standard deviation of $\psi(t)/2$, what it means in our new understanding is that historically not only the market volatility should be much higher in the future as was in the past, but the capital distribution's fluctuation should also be much more volatile. The 20-year history we live is indeed a very peaceful and prosperous period.

from 1993 to 1998 and again from 2003 to 2007, The market cycles can be explained strikingly well by the expansion and the contraction of the log-market cap distribution, as illustrated by $\psi(t)/2$. The magnitude of market volatility can be derived from $\psi(t)/2$ accordingly.

Conclusion: The fat tails in the fluctuations of market index in the stock market are a natural mathematical consequence of **the structure of the market capitalization distribution and simple stochastic calculus.** We've explained such structure here. This is the main conclusion of this section. In a more general sense, fat tailed distribution is the manifestation of an institution composed of many normally populated constituents **whose influences are exponentially distributed**. This is the essence of the lognormal cascade structure. This is also known as a pyramid structure. It is well known that most human societies, including corporations and governments, are all based on some sort of pyramid structure. Thus, their aggregate behaviors are all subject to the phenomena of fat tails if we can quantify their behavior changes over long enough time horizon. The statistics under such structure **deviate significantly** from that of a normal distribution as we would've expected from **Central Limit Theorem** (CLT). It is evident from the market data that not only the log-returns of a market index exhibit fat tails, the log-returns of individual stocks also exhibit fat tails. The author has also shown that the commodities (oil, gold, etc) and Treasury yields all exhibits fat tails and can be fitted well with the first order lognoral cascade distributions (See http://www.skew-lognormal-cascade-distribution.org/apps/). Thus, fat tails in our financial market are universal and unavoidable. The large swings in the financial market are with us forever.

This is a somewhat frustrating conclusion. The mathematics of the normal distribution and CLT has been beautifully developed for 300 years. CLT is regarded as the second fundamental theorem of probability. It is implicitly believed by many participants and policy makers in the financial market that, as the financial system gets bigger and more complicated, the stability of the market will get better – since the aggregate distribution will approach a normal distribution in some way and the tail probability will become negligibly small. But Equation (48) indicates we are far from that utopia as long as $\psi(t)/2$ continues to make large swings every 10 years or so.

Comment on wealth accumulation: According to Equation (39), there are two ways to increase the total wealth of a market, that is, to increase Z by (1) increasing the average $\mathcal{H}_c$; or (2) by increasing the variance$\eta_c(t)^2$ of the log-market cap distribution. However, now we know that the increase in the variance $\eta_c(t)^2$ has an unpleasant consequence of pumping up the market volatility and incurring market instability (if not social instability). Intuitively speaking, this makes sense. Increasing the average without changing the shape of the distribution means everybody shares the prosperity. However, Increasing the variance means some people become ultra-rich while some are pushed to the poverty. Even worse, if the right tail is lengthened, such is the case since 2002, the social injustice is significant. Thus, it seems very important that the policy maker should watch the shape of the capital distribution carefully. The **most serious question** is – Can human beings "manage" the capital distribution of the stock market? In particular, can FED and SEC minimize the variations of the normalized variance $\psi(t)$ as one of their missions to stablize the financial market?

Many details remain to be studied and improved upon for future research. First, how will the model look like if $\mathbb{P}(\mathcal{H}, t\,; \mathcal{H}_c)$ is characterized by the first order distribution as in Figure 6, instead of a normal distribution? This will lift $d \log Z(t)/dt$ to a second order distribution. We know from Section 4 that, in order to expand our study of the market portfolio from 20 years to 80 years, we must incorporate the second order distribution. The fact that η_ψ (0.25) is close to η_2 (0.19) in Figure 2 is not an accident, but an indication that they have profound connection.

Second, the γ^* and γ_i terms need to be included to give a wholly picture of the market dynamics. Are they related to the negative default coefficient β and the positive growth rate g in Equation (48)? Study of $\xi_{i\nu}(t)$ (or the covariance process, $\sigma_{i\,j} = \sum_\nu \xi_{i\nu}(t)\, \xi_{j\nu}(t)$) is essential to answer such question. If not, then what is the origin of β and g in Equation (48)? Third, the higher order terms in $\sigma(\mathcal{H}, t)$ needs to be considered to account for the market-cap effect.

Comment on democracy: As a side note, the relation between fat tails and a pyramid structure could also have social applications. Democracy proclaims an equal weight system. Election reduces the chance of excessive concentration of power. It seems to be an efficient way to counterbalance the pyramid structure. And the balance of powers is an central theme of a democratic system. If our mathematical argument is correct, it is also a proof that a democratic society will be subject to less catastrophic swings, compared to a tyrannic society.

7 Appendix I: Stochastic Portfolio Theory in Continuous Notation

The Stochastic Portfolio Theory has been comprehesively constructed in Fernholz (Fernholz 2002). The basic building blocks are the stochastic calculus (Ito rule) on the logarithmic model of stock prices and portfolio processes; and several important properties about the market. The market is assumed to be (1) nondegenerate; (2) having bounded variance; (3) diverse; and (4) coherent. In addition, portfolios can be constructed via generating functions and the arbitrage opportunities can be evaluated. We can attempt to understand some of these topics in our new found continuous notation.

"Nondegenerate" and "bounded variance" are properties related to the covariance process $\sigma_{i\,j}(t) = \sum_\nu \xi_{i\nu}(t)\, \xi_{j\nu}(t)$. We are not able to address that in this context.

Coherent: (Definition 2.2.1 of Fernholz 2002) The market is coherent if, for all i,

$$\lim_{t\to\infty} \frac{\log \rho_i(t)}{t} = 0,\ a.s. \tag{49}$$

This condition is simply saying none of the stocks declines too rapidly (Chapter 2 Summary of Fernholz 2002). In the continuous notation, this condition describes that the heaviness (kurtosis) of the left tail should not develop too rapidly. The market's coherence has a strong relation to the growth rates of the stocks in the market. The market is coherent if and only if the growth rates are all equal.

The market is coherent most of the time. But we see that when the variance of the log-market cap distribution is expanding rapidly, the market's coherence assumption is greatly challenged. Such is the case of the market crash in late 2008. Many stocks declined rapidly in a few months, and then subsequently were removed from the market.

Diverse: (Definition 2.2.1 of Fernholz 2002) The market is diverse if there exists a number $\delta > 0$,

$$\rho_1(t) \leq 1 - \delta, t \in [0, \infty),\ a.s. \tag{50}$$

This condition is simply saying the heaviness (kurtosis) of the right tail should be limited. To be more precise, let's define ξ such that

$$\int_{\xi}^{\infty} \mathbb{P}(\mathcal{H}; \mathcal{H}_c)\, d\mathcal{H} \ = \frac{1}{N}. \tag{51}$$

Then Equation (50) is equivelant to

$$\int_{\xi}^{\infty} \mathbb{P}(\mathcal{H}; \mathcal{H}_c)\, d\mathcal{H}\, e^{\mathcal{H}+\mathcal{H}_c} \leq Z\,(1-\delta). \tag{52}$$

Intuitively speaking, this is saying the few largest stocks should not dominate the entire market in terms of their market cap weights. Otherwise, the market will become very unstable.

Functionally Generated Portfolio: We will focus on the rank based generating functions in our context. Let $\boldsymbol{S}(\{\rho_i\})$ be the generating function. Then the weights of the generated portfolio are (Theorem 3.1.5 of Fernholz 2002)

$$\pi_i = \left(\mathcal{D}_i \log \boldsymbol{S}(\cdot) + 1 - \sum_{j=1}^{N} \rho_j \mathcal{D}_j \log \boldsymbol{S}(\cdot) \right) \rho_i\ . \tag{53}$$

We will discuss how to transform the portfolio weight into the weight density in the continuous notation. We also use the normal distribution model to study the generated portfolio.

The Market Entropy Generating Function $\boldsymbol{S}(\rho)$: We now discuss the portfolio generated by the market entropy function

$$\boldsymbol{S}(\cdot) = -\sum_{i=1}^{N} \rho_i \log \rho_i, \tag{54}$$

which yields

$$\pi_i = \frac{-\rho_i \log \rho_i}{\boldsymbol{S}(\rho)}. \tag{55}$$

We can rewrite $-\rho_i \log \rho_i$ into the continuous notation

$$-\rho_i \log \rho_i \sim \mathbb{P}(\mathcal{H}; \mathcal{H}_c)\, d\mathcal{H}\, \frac{N e^{\mathcal{H}+\mathcal{H}_c}}{Z} \left(\log Z - (\mathcal{H} + \mathcal{H}_c)\right). \tag{56}$$

We notice that $\log Z - (\mathcal{H} + \mathcal{H}_c)$ becomes negative for very large $\mathcal{H}$. This indicates such generating function will encounter some problem when generalized into the continuous notation. The continuous distribution density for entropy generated portfolio is

$$\Pi(\mathcal{H}; \mathcal{H}_c) = \mathbb{P}(\mathcal{H}; \mathcal{H}_c) \frac{N}{Z \cdot \boldsymbol{S}_0(\rho)} \left(\log Z - (\mathcal{H} + \mathcal{H}_c)\right). \tag{57}$$

The normalization factor $\boldsymbol{S}_0(\rho)$ is fixed by requiring $\int_{-\infty}^{\infty} \Pi(\mathcal{H}; \mathcal{H}_c)\, d\mathcal{H} = 1$, which yields

$$\boldsymbol{S}_0(\rho) = e^{-\mathcal{H}_c - \frac{\eta_c^2}{2}} \left(\log(N) + \frac{\eta_c^2}{2}\right). \tag{58}$$

The portfolio value Z_Π will be proportional to

$$\int_{-\infty}^{\infty} N\, \Pi(\mathcal{H}; \mathcal{H}_c)\, d\mathcal{H}\, e^{\mathcal{H}+\mathcal{H}_c}. \tag{59}$$

Consider the **normally distributed log-capitalization model**, we can use Equation (39) to simplify $\Pi(\mathcal{H}; \mathcal{H}_c)$ and we get

$$\Pi(\mathcal{H}; \mathcal{H}_c) = e^{-\frac{\mathcal{H}^2}{2\eta_c^2}} \frac{\log(N) + \frac{\eta_c^2}{2} - \mathcal{H}}{\sqrt{2\pi}\, \eta_c \left(\log(N) + \frac{\eta_c^2}{2}\right)}. \tag{60}$$

Such portfolio is a normally distributed portfolio similar to the market (the left part), butcomplicated by a polynomial term (the right part). The $(\log(N) + \frac{\eta_c^2}{2} - \mathcal{H})$ term injects underweight in large stocks and overweight in small stocks. This term is linear in $\mathcal{H}$, so the effect is not drastic. We shall study a similar generating function, which produces a very elegant result.

Diversity Generating Function $\boldsymbol{D}_p$: (Example 3.4.4 in Fernholz 2002) Another interesting generating function is

$$\boldsymbol{D}_p(\cdot) = \left(\sum_{i=1}^{N} \rho_i^p\right)^{1/p}, 0 < p \leq 1, \tag{61}$$

which yields

$$\pi_i = \frac{\rho_i^p}{\boldsymbol{D}_p(\rho)^p}. \tag{62}$$

This diversity generating function has been used to generate institutional investment strategies that produces excellent relative returns over the long term.

It is obvious that, when $p \to 1$, such portfolio approaches the market portfolio. We can rewrite ρ_i^p into the continuous notation

$$\rho_i^p \sim \mathbb{P}(\mathcal{H};\mathcal{H}_c)\, d\mathcal{H}\, \frac{e^{p(\mathcal{H}+\mathcal{H}_c)}}{(Z/N)^p}. \tag{63}$$

Therefore, the continous distribution density is

$$\Pi(\mathcal{H};\mathcal{H}_c) = \mathbb{P}(\mathcal{H};\mathcal{H}_c) \frac{e^{(p-1)\,(\mathcal{H}+\mathcal{H}_c)}}{(Z/N)^p \cdot \boldsymbol{D}_0^p(\rho)}, \tag{64}$$

where $\boldsymbol{D}_0^p(\rho)$ is a normalization factor that enforces $\int_{-\infty}^{\infty} \Pi(\mathcal{H};\mathcal{H}_c)\, d\mathcal{H} = 1$. Consider the **normally distributed log-capitalization model**, we can use Equation (39) to simplify $\Pi(\mathcal{H};\mathcal{H}_c)$ and we get

$$\Pi(\mathcal{H};\mathcal{H}_c) = \frac{1}{\sqrt{2\pi}\,\eta_c} e^{-\frac{1}{2\eta_c^2}\left(\mathcal{H}+(1-p)\eta_c^2\right)^2}. \tag{65}$$

Notice that $\boldsymbol{D}_0^p(\rho)$ is fixed by requiring $\int_{-\infty}^{\infty} \Pi(\mathcal{H};\mathcal{H}_c)\, d\mathcal{H} = 1$, which yields

$$\boldsymbol{D}_0^p(\rho) = e^{\frac{\eta_c^2}{2}(p^2-3p+1)-\mathcal{H}_c}. \tag{66}$$

And the portfolio value at point in time Z_Π is proportional to $N\, e^{\mathcal{H}_c + p\frac{\eta_c^2}{2}}$, which converges to Equation (39) when $p \to 1$.

From Equation (65), we find that the portfolio generated by $\boldsymbol{D}_p$ is very elegant when viewed in the normally distributed log-capital model. It is still a normally distributed portfolio in the log-market cap space, just like the market portfolio, except that the peak of its distribution is shifted down by $(1-p)\eta_c^2$. That is, it overweights on the smaller stocks in a systematic way. The strength of such overweight is controlled by the paramter p.

Intuitively speaking, such a normally distributed portfolio is aided by a natural force that pushes it toward the market portfolio. Because its peak is smaller than $\mathcal{H}_c$, the force is always a positive one and this portfolio will outperform the market over time. By fine-tuning p in different strages of a market cycle, the fund manager can exploit different arbitrage opportunities. Since periods of decreasing diversity (or increasing normalized variance $\psi(t)$) are often associated with bear markets, the fund manager can even incorporate a large-cap ($p > 1$) short portfolio to hedge the market exposure in times of trouble. Such hedged portfolio strategy shall be addressed in a future research.

Lastly, it is interesting to note that the polynomial $(p^2 - 3p + 1)$ in Equation (66) has the roots of $\pm(\sqrt{5} \pm 1)/2$ when expressed in $(p-1)$. The higher root is the golden ratio.

References

[1] Arfken, George, 1985, Mathematical Methods For Physicists (Academic Press, 1985).

[2] Fernholz, Robert E., 2002, Stochastic Portfolio Theory (Springer-Verlag New York, Inc.).

[3] Banner, A., Fernholz, R. & Karatzas, I. [BFK], 2005, On Atlas models of equity markets. Annals of Applied Probability 15, 2296-2330.

[4] Lihn, Stephen H., 2008, The Analytical Study and Numerical Method of the Skew Lognormal Cascade Distribution. Available at SSRN: http://ssrn.com/abstract=1273087.

[5] Lihn, Stephen H., 2008, The Scale-Invariant Brownian Motion Equation and the Lognormal Cascade in the Stock Market. Available at SSRN: http://ssrn.com/abstract=1149142.

[6] Cabral, L. M. B., and Mata, J., 2003, On the Evolution of the Firm Size Distribution: Facts and Theory. American Economic Association, vol. 93(4), pages 1075-1090.

[7] Growiec, J., Pammolli, F., Riccaboni, M., and Stanley, E. H., 2008, On the size distribution of business firms. Economics Letters, vol. 98, issue 2, pages 207-212.