On The Network Effect in The Stock Market, The $N^{3/2}$ Power Law, and Optimal Volatility Equilibrium

Stephen H.-T. Lihn
Piscataway, NJ 08854
stevelihn@gmail.com

Second Draft revised on Jan 24, 2010

Abstract

This working paper is an excerpt on the recent finding of the network effect in the stock market. Specifically I summarize the $N^{3/2}$ power law which states that the total value of the stock market is proportional to the three-half power of the number of stocks in the market. This power law is based on several intuitive assumptions on a many-body stochastic system, which will be described in this paper. The hypothesis of the optimal volatility equilibrium is introduced as a pillar leading to the power law, which states that such system must adjust itself around a level of optimal volatility in order to survive as a viable market. The connections to Fernholz's diversity index and Clark's lognormal subordinated process are discussed. This hypothesis also indicates that investing in a large market doesn't guarantee the "benefit" of diversification as prescribed in modern portfolio theory.

1 Introduction

The stock market consists a population of stocks interacting fiercely within an economy. Their individual market values add up to the total wealth of the capital system. Such a tightly coupled, many-body system is very similar to the condense matter systems in physics. As the number of stocks (N) grows, we expect certain statistical properties to emerge. In particular, there should be a scaling law between the number of stocks in the market and the total value of the market, $Z = f(N)$. This is the primary topic of this paper.

There have been several scaling laws floating around in microeconomics in the last few decades. First of all, when there is no additive effect to be expected, the total value of a system is linearly proportional to the number of constituents in it, $Z \sim N$. This is called Sarnoff's law, attributed to David Sarnoff. On the other hand, when each constituent is interconnected to other constituents and these connections contribute to the value of the system in addition to its own worth, one would expect a value-add effect, more than that of Sarnoff's. Such effect is called the "network effect." In particular, Metcalfe's law states that the total value of a network is proportional to the square of the number of constituents, $Z \sim N^2$. The square power law comes from $N(N-1)/2$, the number of connections when every constituent is connected to all other constituents and each connection is valued in equal weight.

There is yet the third popular scaling law, Reed's law, which is the assertion of David Reed that the utility of large networks, particularly social networks, can scale exponentially with the size of the network, $Z \sim 2^N$. This is due to the fact that, given a set of N constituents, it has 2^N possible subsets.

These scaling laws have profound influence on the estimated value of a network, particularly when senior management or a venture capitalist makes investment or acquisition decisions. While under-estimate could blind enterprenuers from persuing golden opportunities, over-estimate inevitably leads to significant waste of capital. During the golden years of Internet's explosive growth (1997-2009), we have witnessed two cycles of boom and bust. On one hand, several companies successfully harnessed the network effect to grow to Fortune 100 statue in less than 10 years (which normally would take 50-100 years). On the other hand, even more companies attempted, or even half way there in gathering the critical mass, but only to bring unspoken grief to their investors at the end. The capital loss is often 100 percent, and in billion dollars. These golden years unfortunately turned out to be one of the worst investment experiences in the stock market's history. In Jan, 1997, DOW was at 6500. In Mar, 2009, DOW retreated back to 6500 briefly. The 12-year investment return was a big zero.

On the other hand, since the stock market is a vast, complicated network with tens of thousands of stocks, one would expect certain network effect exists. This network effect could be also useful in determining the fair value of the stock market. Investors who can determine such fair value, even approximately, will have an upper hand compared to those who can't.

In this paper, I propose a new scaling law, not based on simple permutation, but on stochastic calculus. I call this the $N^{3/2}$ power law, which states that the total value of the stock market is proportional to the 3/2 power of the number of stocks in this market, $Z \sim N^{3/2}$, assuming that the mean of log-market capitalization in the population does change much in the long term (which is true for the U.S. stock market).

Obviously, $N^{3/2}$ is more optimistic than Sarnoff's law. The network effect provides a financial incentive to participate in the market (as we generally believe). However, the benefit is nothing like the nuclear chain reaction of 2^N. It is actually slightly less optimistic than Metcalfe's square power law, N^2.

The scaling law that describes the stock market has a profound impact to our understanding of network effect. From the resource conversion's perspective, the stock market is the ultimate capital allocation engine in the modern economy. If the stock market follows $N^{3/2}$ power law, than nothing else of value in the economy can grow faster than $N^{3/2}$. Otherwise, we can simply IPO these N^2 or 2^Nnetworks and inject the fountain of youth into our beloved stock market. The zero return in the past 12 years seems to disapprove the existence of such fountain.

2 Formulation

In this section, the four pillars that lead to the $N^{3/2}$ power law will be described. In order to keep the framework generic, we can use the market to represent an abstract system that can be a stock market, or a social network (Facebook, Youtube), an electronic market place (Amazon.com, eBay), a telephone user base (AT&T), or an Internet network. And the constituent of the market is a stock for a stock market, a user in a social network, a buyer (or seller) on eBay, a telephone subscriber, and an endpoint on the Internet network. However, we do prefer to use the language of the stock market since it is our main interest in this paper.

2.1 Population Distribution

The first pillar of the $N^{3/2}$ power law states that, at any given time t, the population distribution of the market is approximately a normal distribution when measured by a canonical coordinate. For a stock market, such canonical coordinate is the log-market capitalization. The distribution is defined by the probability density function

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

in which $\mathcal{H}_c(t)$ is the mean log-market capitalization of the population, $\eta_c(t)^2$ is the variance of the normal distribution, and $N(t)$ is the number of constituents in the market. Obviously,

$$\int_{-\infty}^{\infty} \mathbb{P}(\mathcal{H}, t)\, d\mathcal{H} \;= N(t). \tag{2}$$

2.2 Exponential Influence

The second pillar of the $N^{3/2}$ power law is to assume the "influence" of the constituents is proportional to the exponential of its coordinate, $e^{\mathcal{H}}$. This is a bit abstract. For the stock market, the "influence" of a stock is its market capitalization, (that is, its contribution to the value of the market). Since the canonical coordinate is the log-market capitalization, it is obvious that the a stock's "inflence" to the market is the exponential of its log-market capitalization. Thus, the total value of the market is

$$Z(t) = \int_{-\infty}^{\infty} \mathbb{P}(\mathcal{H}, t)\, d\mathcal{H}\, e^{\mathcal{H}} = N(t)\, e^{\mathcal{H}_c(t)+\eta_c(t)^2/2}. \tag{3}$$

The output of the integral is simply the expected value of a lognormal distribution.

2.3 The Dynamic Equilibrium of A Growing Population and The $N^{3/2}$ Power Law

The third pillar of the $N^{3/2}$ power law is to define the dynamic equilibrium when the population is growing. It assumes that the growing population tends to reach (or oscillate around) an equilibrium level when $\eta_c(t)^2$ is expanding in $\log N(t)$ when $N(t)$ changes. That is,

$$\eta_c(t)^2 \to \log N(t) + \psi_c, \tag{4}$$

in which ψ_c is a constant independent of $N(t)$. (The stochastic details of Equation (4) will be explained in Section 2.4.) The point here is that, for a stock market, when more and more stocks are added into the financial supermarket [1] , the width of the capital distribution $\eta_c(t)$ is not going to stay the same, but will expand in a systematic way. Such expansion has a profound effect on the growth of wealth as well as the volatility characteristics of the market, as will be shown in Section 2.4. With Equations (3) and (4), the total market value becomes

$$Z(t) \to N(t)^{\frac{3}{2}} e^{\mathcal{H}_c(t)+\psi_c/2}. \tag{5}$$

If we further assume $\langle \mathcal{H}_c(t) \rangle$ remains constant over time, then we obtain the $N^{3/2}$ power law:

$$Z(t) \sim N(t)^{\frac{3}{2}}. \tag{6}$$

At its face value, this scaling law is useful in gauging the approximate level of market index, such as S&P500 or DOW. For example, one can examine how many stocks were included in the Vanguard total market index fund in the past 12 years (as of early 2010). One will see that the number at the end of the period is not higher than at the start of the period. Therefore, one should not expect DOW to be higher than 12 years ago.

For the stock market, the assumption of constant $\langle \mathcal{H}_c(t) \rangle$ means that the average (logarithmic) market value of stocks does not change much over the long term. There are two forces that affect $\mathcal{H}_c$. On one hand, there is the inflationary force that executives want to grow their companies bigger (by either organic growth or acquisitions) and avoid being marginalized (if stay small). Years of mergers and acquisitions are making Fortune 500 companies bigger and bigger. On the other hand, there is the deflationary force, particularly in a mature economy such as in the U.S.. First, emerging markets around the world are constantly challenging the status quote of large corporations, reducing their market shares and large moats of profits. Secondly, the progress of technology is also deflationary – destroying many older products and services while new ones invented in completely different directions. The two forces tend to balance each other and keep $\mathcal{H}_c$ more or less constant over the decades.

[1] In reality, $N(t)$ are increasing in the long term, but did suffer temporary contraction during recessions. In recent memory, $N(t)$ increased particularly fast during the Internet bubble years (1997-2000), but decreased in the bear markets of 2001-2 and 2008-9.

As the readers will find later that, for the stock market, such dynamic equilibrium is far from "static". The market never "converged" to such equilibrium, but rather "oscillated" around such equilibrium. The market often overshot for a period of time, and then undershot for another period of time. The "overshoot" periods approximately coincide the periods of high volatility, while the "undershoot" periods that of low volatility. Such cycles can be very extensive, in the order of 10 years for each.

The origin of such expansions and moody oscillations is still unknown. But the effects have profound (negative) impact to the anticipated volatility reduction as a prescribed "benefit" of diversification in modern portfolio theory. What is known to happen is that, after each cycle, the large cap stocks become even larger and a portfolio is again dominated by a few large stocks. The smaller stocks no longer provide enough weights to "balance" the portfolio. Therefore, we have been having, and will continue to have, market turmoils approximately every 7-10 years. And it is not going to get better by putting even more stocks into the stock market.

2.4 Optimal Volatility Equilibrium

The assumption that $\eta_c(t)^2$ is expanding in $\log N(t)$ is based on the stochastic calculus in a lognormal cascade, which is explained in this section. The fundamental hypothesis is that a growing population will reach (or oscillate around) a dynamic equilibrium such that the volatility of the market value process $Z(t)$ always stays in a reasonable range and satisfies the following requirements:

1. The market volatility doesn't depend on the number of constituents in the market. This is the "scalability" requirement.
2. The market volatility can't be very high since high volatility makes the market unstable, thus cannot survive.
3. The market volatility can't be very low. Market complacency entices people to make stupid decisions, thus is not sustainable either.

What remains is that, a healthy long-lasting market will oscillate around an optimal level of volatility over the long term. Such condition is called the **Optimal Volatility Equilibrium**.

Let's formulate these requirements in the context of stochastic calculus. We start with a group of stocks $\{X_i\}$ in the market with the market value process $Z(t)$: (The readers can refer to Corollary 1.1.6 of Fernholz 2002.)

$$\begin{aligned} d\log Z(t) &= \sum_{i=1}^{N(t)} \mu_i(t)\ d\log X_i(t) + \gamma^*(t)\,dt, \\ d\log X_i(t) &= \gamma_i(t)\,dt + \Sigma_\nu\,\xi_{i\nu}(t)\,dW_\nu(t), \end{aligned} \tag{7}$$

in which $\gamma^*(t)$ is excess growth rate of the market portfolio, $\gamma_i(t)$ is the growth rate of the i-th stock, and $\xi_{i\nu}(t)$ is the volatility matrix. $\mu_i(t)$ is the weight of the i-th stock, $X_i/Z(t)$. We will only focus on the (geometric) Brownian motion terms and the growth rate terms are ignored. We also want to simplify the volatility matrix $\xi_{i\nu}(t)$ such that each stock is one Wiener process and all the stocks share the same characteristic volatility σ_c. Therefore, we have

$$\begin{gathered} d\log Z(t) = \sum_{i=1}^{N(t)} \mu_i(t)\ d\log X_i(t), \\ d\log X_i(t) = \sigma_c\, d\,W_i(t). \end{gathered} \tag{8}$$

This simple model allows us to obtain the market volatility in closed form by using the additive rule of normal distributions:

$$\sigma^2_{\log Z}(t) = (\sigma_c\, V(t))^2, \tag{9}$$

in which $V(t)$ is the **volatility reduction factor**:

$$V(t)^2 = \sum_{i=1}^{N(t)} \mu_i(t)^2. \tag{10}$$

Notice that $V(t)$ is in the same form of the diversity generating function $\boldsymbol{D}_p(\cdot)$

$$\boldsymbol{D}_p(\cdot) = \left(\sum_{i=1}^{N} \mu_i^p\right)^{1/p}, \tag{11}$$

in which $p = 2$ (See Example 3.4.4 in Fernholz 2002; or Appendix in Lihn 2009). Therefore, $V(t)$ also serves as a measure of diversity. Since it will be shown below that $V(t)$ is a measure of market volatility, there is a strong connection between market diversity and market volatility. $p = 2$ is a setting to measure how large cap stocks dominate the market. Thus, it indicates that large cap stocks can't outperform the small cap stocks without either increasing the market volatility or having a large pile of IPOs to supply enough fuel.

In modern portfolio theory, the volatility reduction factor describes how much volatility can be reduced by diversification. Its closed form can be carried out in our model as

$$V(t)^2 \approx \int_{-\infty}^{\infty} \mathbb{P}(\mathcal{H}, t)\, d\mathcal{H}\, \frac{e^{2\mathcal{H}}}{Z(t)^2} = \frac{1}{N(t)}\, e^{\eta_c(t)^2}. \tag{12}$$

If the width of the capital distribution $\eta_c(t)$ remains the same, then the market volatility (as characterized by $V(t)$) will indeed be lowered by adding more stocks to the market (an increasing $N(t)$). This is the "benefit" of diversification that we mentioned before. However, what will be shown below is unfortunately the opposite: Such "benefit" of diversification doesn't exist due to the requirements of Optimal Volatility Equilibrium.

Our scalability requirement (Requirement 1) demands certain behavior of $\eta_c(t)$ such that $V(t)$ is independent of $N(t)$. The most general form for $\eta_c(t)$ is (a revision of Equation

(4))

$$\eta_c(t)^2 = \log N(t) + \psi_c + \psi(t), \tag{13}$$

in which $\psi(t)$ is independent of $N(t)$ and is detrended such that $\langle \psi(t) \rangle = 0$. Therefore, Equation (13) leads to the $N^{3/2}$ power law, Equation (5).

Furthermore, the market volatility becomes

$$\sigma_{\log Z}(t) = \sigma_c e^{\frac{1}{2}\psi_c} e^{\frac{1}{2}\psi(t)}. \tag{14}$$

Notice that $\sigma_{\log Z}(t)$ doesn't depend on $N(t)$ any more. That is, the scalability requirement ensures that the size of the market is irrelevant to the market volatility, and the side effect is that diversification is ultimately meaningless. To fulfill Requirement 2 and 3, $\psi(t)$ must be a mean-reverting process that has bounded variance [2] . Thus the **Mean Reversion of Market Volatility** is a natural outcome of Optimal Volatility Equilibrium.

Lognormal Cascade Distribution: If $\frac{1}{2}\psi(t)$ is normally distributed with variance η^2, then the log-return of the market value process $Z(t)$ follows a first-order symmetric lognormal cascade distribution (Lihn 2009):

$$\mathrm{pdf}(x) = \int_{-\infty}^{\infty} dy \, \frac{1}{2\pi \, \eta \, \sigma(y)} e^{-\frac{y^2}{2\eta^2}} e^{-\frac{x^2}{2\sigma(y)^2}}, \tag{15}$$

where x is the log-return of $Z(t)$, and $y = \psi/2$, and $\sigma(y) = \sigma_c e^{\frac{1}{2}\psi_c} e^y$.

This model still misses the skew terms. How to derive and validate the skew terms from the fundamental stochastic calculus is still under research. It is unknown whether the skew terms have any long-term effect.

3 Discussion

I have shown that the log-returns of a large market follows a lognormal cascade distribution under the following hypotheses:

1. The population distribution of the market constituents is Gaussian.
2. The influence of the constituents are proportional to the exponential of the coordinate.
3. The variance of the population distribution approaches the logarithm of the population. This is called optimal volatility equilibrium.
4. The fluctuation of the variance around the optimal volatility equilibrium is the source of the lognormal cascade, which causes the market's log-returns to exhibit fat tails.

[2] The simplest model for $\psi(t)$ is an Ornstein-Uhlenbeck process, but it is beyond the scope of this paper.

This model is very different from that of P. K. Clark's lognormal subordinated Levy process (Clark, 1973). But surprisingly, the mathematical form of the pdf is the same (Lihn, 2010). In my model, the market's out-sized fluctuations do not come from the changes of trading time, which are linked to some point-in-time market quantities, such as trading volumes, transaction sizes, etc. Nevertheless, the market's out-sized fluctuations come from the changing shape (expansion) of the population distribution of the underlying market constituents.

The tendency that the market always approaches the level of optimal volatility equilibrium seems to nullify the "benefit" of diversification by investing in a large market. When the market's population distribution reaches a point where the few large constituents are exceedingly more influential than the rest of population, there is no way to avoid the mischief resulted from them by being "more diversified". At this point, it seems that the only way to avoid investment loss is to not be in the market at all unless the market readjusted itself to a less volatile and more diverse distribution.

One good example of such extreme volatility is the financial crisis of 2008 that inflicted the entire world. It was brought about by the mischief of less than 10 large companies. In a further distant memory, at the peak of the Internet bubble (March, 2000), Cisco was about to reach the obscene market capitalization of $1 trillion in a few more month. Its dominance eventually evaporated in 2001 and brought the market down with it. In 1998, the market crash was caused by a single hedge fund (LTCM) sponsored by the handful of largest banks on Wall Street. History has shown again and again the result of optimal volatility equilibrium, and disapproved the belief that a larger capital market can reduce market volatility.

References

[1] Clark, Peter K., 1973, A Subordinated Stochastic Process Model With Finite Variance For Speculative Prices, Econometrica, Vol. 41, pp. 135-155.

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

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

[4] Lihn, Stephen H., 2009, The Theory of Higher Order Lognormal Cascade Distribution and the Origin of Fat Tails in the Fluctuations of Stock Market Index. Available at SSRN: http://ssrn.com/abstract=1414392.

[5] Lihn, Stephen H., 2010, Comment on P. K. Clark's Distribution of Lognormal-Normal Increments And the Lognormal Cascade Distribution. Available at SSRN: http://ssrn.com/abstract=1537686.