The Fluctuations in Trading Volume Imbalance

: Evidence from KOSPI 200 Index Futures Market

Keunho Hwang, Jangkoo Kang and Doojin Ryu*

Preliminary Version: March 2007

ABSTRACT

Large fluctuations are observed in various financial variables, such as stock price, foreign exchange rate, interest rate, and trading volume. Many empirical evidences show that the Levy or power-law distributions can describe these movements. We elaborately investigate the movement of trading volume imbalances in KOSPI 200 index futures market. The empirical phenomenon, so called two-phase behavior, that the distribution changes from unimodal to bimodal as the variation increases is observed. We also examine the statistical properties of trading volume and number of transactions, which determine the trading volume imbalance. The trading volume and number of transactions within one minute shows power-law property but individual trading volume does not. Contrary to Maria and Yamazaki (2005), the power-law of trading volume is not the necessary condition to generate the two-phase behavior.

JEL classification: C16, D39

* Corresponding author: KAIST Graduate School of Finance 207-43, Cheongryangri 2-Dong, Dongdaemun-Gu, Seoul, 130-012, Korea, Tel: +82-2-958-3693, E-mail: sharpjin@business.kaist.ac.kr

1. Introduction

For many years, the distributions of variables in the financial market have been studied extensively. Many articles show that large fluctuations which cannot be explained by Gaussian distribution are observed and that the power-law or Levy distribution can describe these movements. In the early 1960s, Mandelbrot (1963) and Fama (1965) documented that the distributions of returns can be well approximated by a Levy stable distribution. In recent years, many researchers examined various financial markets and found the power-law distributions in stock price, trading volume, and number of transations. (For example, see Gabaix et. al. (2003), Gopikrishnan et. al. (1999), Plerou et. al. (2000).)

A large price fluctuation is induced by a large shock in the market. If demand and supply for the assets are balanced, the change of price is mild. On the other hand, if a large information shock arrives in the market, the imbalance between demand and supply can result in considerable price movement. Plerou et. al. (2003) examined the distribution of trading volume imbalance within a fixed time interval and they empirically found that a unimodal distribution changed into a bimodal distribution as the variation increased. They interpreted this phenomenon that the market moves between the two phases as "equilibrium" state and "out-of-equilibrium" state. Maria and Yamazaki (2005) performed the numerical simulation where trading volume and number of trades are sampled from power-law distribution and showed that the two-phase behavior is induced when the distribution of trading volume has fat-tail.

In this paper, we examine the distribution of trading volume imbalance in KOSPI 200 index futures market. The trading volume imbalance within one minute explicitly shows the two-phase behavior. We also examine the statistical properties of trading volume and number of transactions, which determine the trading volume imbalance. As in other financial markets, the power-law distributions are observed in KOSPI 200 index futures market. As Martia and Yamazaki pointed out, the two-phase behavior observed in KOSPI 200 index futures market may be a consequence of the statistical property of trading volume and number of trades. Though the trading volume which is accumulated within a fixed time interval has a power-law property, the empirical evidence shows that it is hard to say the individual trading volume of

KOSPI 200 index futures has a power-law distribution. This shows the two-phase behavior of trading volume imbalance is not the consequence of the power-law distribution of trading volume.

The rest of this paper is organized as follows. Section 2 describes the properties of KOSPI 200 index futures market and presents how to construct our sample. Section 3 shows the two-phase behavior of trading volume imbalance in the KOSPI 200 index futures market. Section 4 plots the distributions of other market variables and discusses the statistical properties of those variables. And section 5 suggests a potential explanation for the two-phase behavior and concludes briefly.

2. KOSPI 200 Index Futures Market and Data Sampling

Since the KOSPI 200 index futures were introduced in May 1996, they have grown tremendously. Based on its trading volume, the KOSPI 200 index futures market has become one of the top derivative markets in the world. The average daily trading volume in 2003 was 251, 841 contracts[2].

Aside from its sheer trading intensity, the KOSPI 200 index futures market has a peculiar characteristic compared to other derivative markets of developed countries. While the trading activities of institutional investors dominate those of individual investors in most derivative markets of developed countries, the domestic individual investor's trading volume is two times more than the sum of domestic and foreign institutional investors'. Individual investors are generally considered unsophisticated traders, (i.e. uninformed or noise traders), who do not trade based on true information. In the emerging market dominated by individual traders, such as the KOSPI 200 index futures market, there may be more room for sophisticated institutional investors to exploit less sophisticated individual investors using their information and trading

[2] To more specific, 128,058 contracts in 2001, 175,689 contracts in 2002, 223,329 contracts in 2004 and 176,099 contracts in 2005.

skill. So, we may observe some peculiar phenomenon unique to KOSPI 200 index futures market.

Our sample period is from September 2001, to December 2004[3]. There are four kinds of index futures that have different maturities. But only the nearest month's future is actively traded while other months' futures are barely traded at any given time[4]. So we use only the nearest month future data.

One futures contract size is 500,000 Korean Won (KRW, hereafter) times the current level of the KOSPI 200 index value. Tick size is 0.05 points. (i.e. 25,000 KRW) Using the historical Trade and Quote (TAQ) database of the KOSPI 200 index futures provided by Korea Stock Exchange (KRX), we organize a limit order book. We calculate bid-ask price, quoted spreads and effective spreads just before every transaction. Since all orders and trades are ten milliseconds time-stamped, we can easily identify who initiates trade (buyer or seller).

We use all trades and quotes recorded during continuous trading secession of each day and we discard the call trading secession data. The trading secession of KOSPI 200 index futures market consists of pre-opening call market (8:00~9:00), closing call market (15:05~15:15) and continuous double auction market (9:00~15:05).

3. Two Phase Phenomenon

As in Plerou et al (2003), to quantify demand, we define the volume imbalance, $\Omega(t)$, as the volume weighted sum of trade indicators within Δt.

$$\Omega(t) = Q_B - Q_S = \sum_{i=1}^{N(t)} q_i a_i \tag{1}$$

[3] TAQ data is not available from December 19, 2002 to March 31, 2003. Because we found some errors (omission of raw data) on several days such as December 8, 2004 or and December 20, 2004, we discard those days. So, total sample size is about (little more than) three years. (36 months, 760 trading days)
[4] Even on the day of maturity of the nearest month's future, the transaction volume of that future is about five times more than that of the second nearest month's future. 'Roll Over' does not exist.

N denotes the number of transaction within time interval Δt.

q_i denotes the i-th transaction volume.

$a_i = \begin{cases} 1 \\ -1 \end{cases}$ a_i, trade indicator, is 1 if i-th transaction is buyer-initiated trade, and -1 if i-th transaction if seller-initiated trade. Unlike a quote-driven market or a hybrid market such as NASDAQ or NYSE, there is no mid-quote transaction in a pure order-driven market. So we let a_i have only two values depending on who initiates the trade.

The local noise density, $\Sigma(t)$, is defined as the first centered moment of volume weighted trade indicators.

$$\Sigma(t) = \left\langle \left| q_i a_i - \langle q_i a_i \rangle \right| \right\rangle, \qquad (2)$$

where <> means local expectation value.

Ω and Σ are normalized so that they have zero mean and first unite centered moment.

We draw probability density functions conditional on Σ. Figure 1 shows how $P(\Omega|\Sigma)$ change as Σ changes. (Normalized) Local deviation interval has following six intervals.

(1) $\Sigma \le -1$ (2) $-1 < \Sigma \le 0$ (3) $0 < \Sigma \le 1$ (4) $1 < \Sigma \le 2$ (5) $2 < \Sigma \le 3$ (6) $3 < \Sigma$

We find explicit two-phase behavior when order imbalance fluctuates. PDF shows single-peaked shape for small Σ values and double peaked shape for large Σ values.

[Figure 1 is about here]

As in Plerou et al (2003), to investigate wherein the phase shift phenomenon occurss, we split local deviation to a tighter interval and calculate order parameter $\Psi(\Sigma)$. The order parameter $\Psi = \Psi(\Sigma)$ is defined as the value of $\Omega_{\pm}$ where $P(\Omega|\Sigma)$ has its maximum value.

We find a phase shift arise when Σ is around -0.2. In other words, the critical point of Σ (Σ_c) is where the behavior of system suffers qualitative change, at about -0.2, and there exists only one peak in conditional PDF $P(\Omega|\Sigma)$ if Σ is smaller than Σ_c and there exists double peaks in conditional PDF $P(\Omega|\Sigma)$ if Σ is larger than Σ_c.

Figure 2 shows that

$\Psi(\Sigma)=0$ where $\Sigma < \Sigma_c$ and $\Psi(\Sigma)=(\Sigma-\Sigma_c)/3$ where $\Sigma >> \Sigma_c$.

[Figure 2 is about here]

4. The Characteristics of Market Variables

The two-phase behavior that we reported in the last section can be relate to the properties of variables that compose Ω and Σ as pervious works pointed.

Using simulation, Matia and Yamasaki (2005) finds that two-phase behavior arises when the market variable, q_i, follows power-law distribution. They insist that the fat tail property of power law distribution relates the two phase behavior of the volume imbalance (Ω) when it is volatile. Their results show that if each transaction volume(q_i) generated by power-law distribution, one can observe a two-phase behavior irrespective of other variables' distributions such as the number of transactions(N) or the trade indicator(a_i). They explain since the volume imbalance, Ω, is function of q_i, then Ω also has a functional form of power-law, and this directly relates to bifurcation of conditional PDF, $P(\Omega|\Sigma)$.

In the case of KOSPI 200 index futures market, the large portion of q_i is just 1. This is because the individual trader (who are mostly regarded as noisy traders) dominates other type of traders in our market.

And since institutional traders often summit very large orders, q_i can sometimes have a very large value. So we can find any fat tailed property of empirical PDF of q_i and we are sure that q_i doesn't follow power-law distribution. Matia and Yamasaki (2005) insist that the only important factor that generates the two phase behavior is the power-law distribution of q_i and the distribution of N or a_i doesn't matter. But we suspect their argument because q_i

doesn't follow as power law distribution of fat-tail distribution.

We investigate how market variables are empirically distributed.

Table 1 shows the rough distribution of q_i. Nearly half of the transaction volume is just one contract and 1.16(%) of the transaction volume is more than one hundred contracts.

[Table 1 is about here]

Figure 3 shows the PDF of log transformation of the following variables[5]. Comparing Gaussian distribution, they are skewed and fat-tailed.

Market variables

(1) N : number of transactions during Δt

(2) $q(t)$: Average transaction volume during Δt, $q(t) = \frac{1}{N(t)} \sum_{i=1}^{N(t)} q_i$,

(3) $Q(t)$: Summation of whole trans action volume during Δt, $Q(t) = \sum_{i=1}^{N(t)} q_i$

All market variables are calculated during time interval, Δt.

Basic time interval is one-minute. Δt =1min.

[Figure 3 is about here]

Figure 4 shows the CDF of above variables. When compared to Gaussian distribution, they are more similar to power-law distribution which has fat tail.

[Figure 4 is about here]

[5] It is usually known, the market microstructure variables, proxies for liquidity (number of transaction, trading volume, bid-ask spread) show intra day U-shape or downward sloping patterns. To eliminate this trend, we divide each variable by the mean value which is estimated by the values included in the same intraday interval of whole sample day. For example, the variables calculated from 9:00 to 9:01 of September 2, 2001 should be divided by the mean value of the variables calculated from 9:00 to 9:01 of whole sample day.

And we use the method of detrended fluctuation to measure power-law time correlations in each market variable if it follows power-law distribution. The procedure of detrended fluctuation analysis[6] follows. For normalized market variables which eliminate intraday pattern, we make integrated process.

Figure 5 shows detrended fluctuation of each market variable.

[Figure 5 is about here]

Each point is calculated as follows.

First, we divide the whole sequence of each market variable into N/t non-overlapping interval. In each time interval, the number of data points is t. We make integrated process and use the linear least square method to get the locally detrended residual (detrended fluctuation function) in each interval.

The detrended fluctuation function is calculated from

$$F^2(t) = \frac{1}{t} \sum_{n=(k-1)t+1}^{kt} \left| y(n) - z(n) \right|^2 \quad k = 1,2,3,\ldots,N/t$$

$y(n)$ is market variable.

$z(n) = an + b$ is local trend

For example, in the case of $N(n)$ for the t =4,

The 4 points N_1, $N_1 + N_2$, $N_1 + N_2 + N_3$, $N_1 + N_2 + N_3 + N_4$ are regressed local time points 1, 2, 3, 4. And we get the residual and calculate $\frac{1}{4}\sum_{i=1}^{4} e^2$ for each interval. Finally we average the time series of $\frac{1}{4}\sum_{i=1}^{4} e^2$ and this average value is the one point in figure corresponding t =4.

---

[6] For more details on this procedure, refer Vandewalle et. al. "Detrended fluctuation analysis of the foreign exchange market"

Second we average detrended fluctuation function, $F(t)$, over the N/t intervals and get a function which depends on the interval size t. We set time size t as 4, 8, 16, 32, 64, 128, 256, 512, 1024 and 2048 minutes. If the $F(t)$ follows power-law distribution ($\langle F \rangle \sim t^a$), the log-log plot of detrended fluctuation function versus interval size t is straight line.

5. The Possible Cause of Two Phase Behavior and Concluding Remarks

The two phase behavior may be generated by power-law distribution property of market variables as Matia and Yamasaki (2005) insist. But according to their argument, even in the case when N is generated from power-law distribution, if q_i does not follow power-law distribution, they can't observe a two-phase behavior. But in our market, although three market variables (N, $q(t)$, $Q(t)$) seem to follow power-law distribution and have fat tailed property, q_i does not. It still may be important for $q(t)$ or $Q(t)$ to have fat tailed property for two-phase behavior. Or the autocorrelation of a_i is an important factor for two-phase behavior. In Matia and Yamasaki (2005), they ignore serial correlation of each variable.

Without information shock in the market, there is no reason to occur statistically net demand or net supply. Namely, the probability that a transaction is initiated by buyer is equal to the probability that a transaction is initiated by seller. This state corresponds to '$\Sigma < \Sigma_c$' market phase. We can observe symmetric single modal distribution of the volume imbalance if the net demand stays in the equilibrium state.

If the information shock arrives at the market, however, the transactions are derived by one side, ask or bid side, depending on the sign of information shock (positive or negative information). Furthermore, if the information shock arrives at the liquidity-abundant and efficient derivative market, such as KOSPI 200 index futures market, it is swiftly reflected to price and volume imbalance. So if this happens, the fluctuation and local noise intensity of volume imbalance increase.

This is why we can observe a two-phase behavior when the local noise intensity is large enough. This state corresponds to '$\Sigma > \Sigma_c$' market phase. We can observe symmetric bimodal distribution of the volume imbalance. When information shock arrives, the autocorrelation of trader indicator a_i is not zero. It will show that price continuation and this reflects that the transactions are biased on one side.

This may be the important reason we can still find the two-phase behavior even when q_i doesn't follow power-law distribution. The two-phase behavior is not just an echo of q_i's power-law distribution as Matia and Yamasaki(2005) insist, rather it relates to the information effect.

References

Fama, Eugene F., 1965, The behavior of stock-market prices, *The Journal of Business*, vol. 38, no. 1, pp. 34-105.

Gabaix X., Gopikrishnan P., Plerou V. and Stanley H.E., 2002, A theory of power-law distributions in financial market fluctuations, *Nature*, vol. 423, 267–270.

Gopikrishnan P., Meyer M., Amaral L.A.N. and Stanley H.E., 1998, Inverse cubic law for the distribution of stock price variations, *European Physical Journal B* 3, pp. 139–140.

Gopikrishnan P., Plerou V., Amaral L.A.N., Meyer M. and Stanley H.E., 1999, Scaling of the distributions of fluctuations of financial market indices, *Physics Review E* 60, 5305–5316.

Gopikrishnan P., Bernd Rosenow, Vasiliki Plerou, and H. Eugene Stanley, 2001, Quantifying and interpreting collective behavior in financial markets, *Phys. Rev.* vol. 64.

Mandelbrot, B.,1963, The variation of certain speculative prices, *The Journal of Business*, vol. 36, no. 4. , pp. 394-419.

Matia, K., Yamasaki, K., 2005. Statistical properties of demand fluctuation in the financial market, *Quantitative Finance*, Vol. 5, No.6, 513-517.

Mitzenmacher, M., A brief history of generative models for power law and lognormal distributions, *Internet Mathematics* , vol 1, no. 2, 226-251.

Vasiliki Plerou, Parameswaran Gopikrishnan, H.Eugene Stanley , 2003, Two-phase behavior of financial markets, *Nature*, Vol 421, 130.

Vasiliki Plerou, Parameswaran Gopikrishnan, Luis A, Nunes Amaral, Xavier Gabaix, and H. Eugene Statnley, 2000, Economic fluctuations and anomalous diffusion, *Physical review*, Vol 62, number 3

Vandewalle, N., Ausloos, M., Boveroux, Ph., *working paper* , Detrended fluctuation analysis of the foreign exchange market.

Table 1

q_i	(%)
q=1	47.534(%)
1<q<=5	29.612
5<q<=10	8.858
10<q<=50	10.651
50<q<=100	2.1834
100<q	1.1617

Figure 1

Figure 1 shows conditional density of the volume imbalance, $P(\Omega|\Sigma)$ for varying Σ interval.

(1) $\Sigma \le -1$ (2) $-1 < \Sigma \le 0$ (3) $0 < \Sigma \le 1$ (4) $1 < \Sigma \le 2$ (5) $2 < \Sigma \le 3$ (6) $3 < \Sigma$

Ω is the volume imbalance, $\Omega(t) = Q_B - Q_S = \sum_{i=1}^{N(t)} q_i a_i$ and Σ is the local noise intensity, $\Sigma(t) = \langle |q_i a_i - \langle q_i a_i \rangle| \rangle$. Ω and Σ are normalized so that they have zero mean and first unite centered moment.

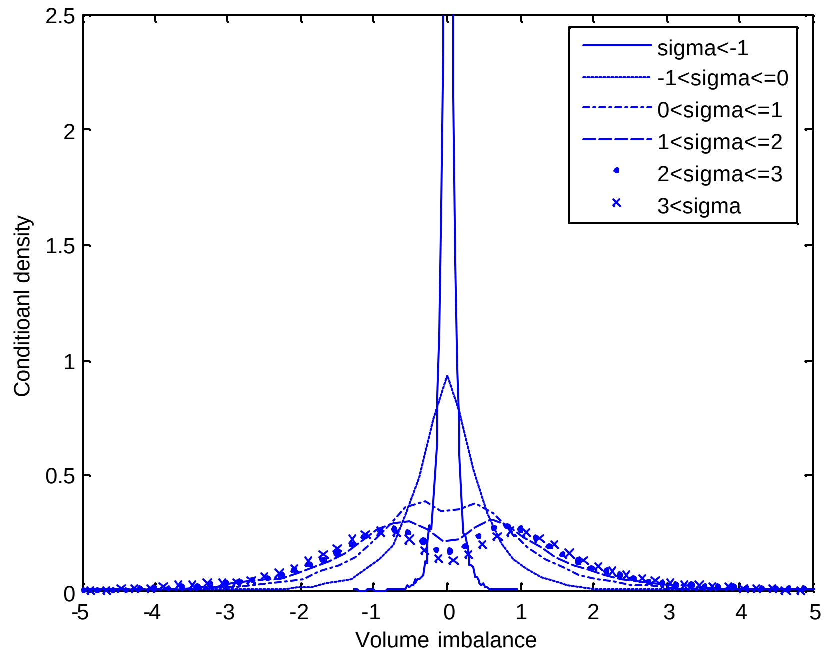


Figure 2

Order parameter $\Psi = \Psi(\Sigma)$ means, given Σ, where the maximum value of the conditional density of the volume imbalance. We split local deviation to 0.1 interval to calculate order parameter $\Psi(\Sigma)$.

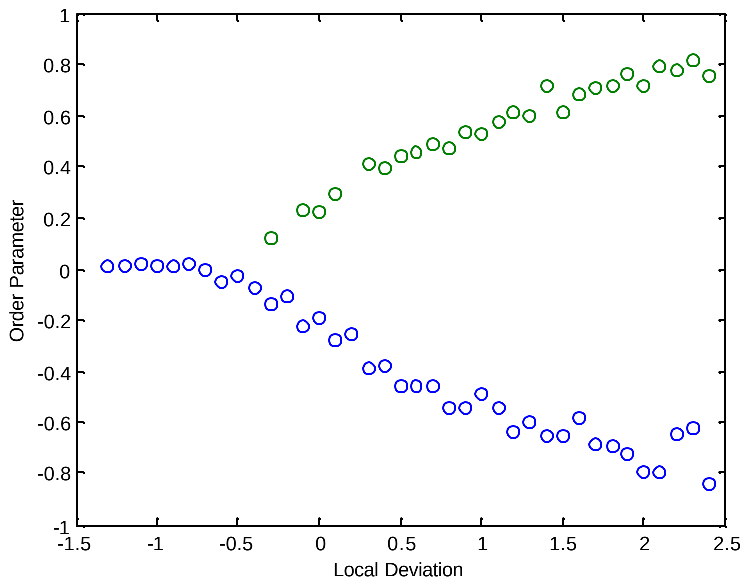
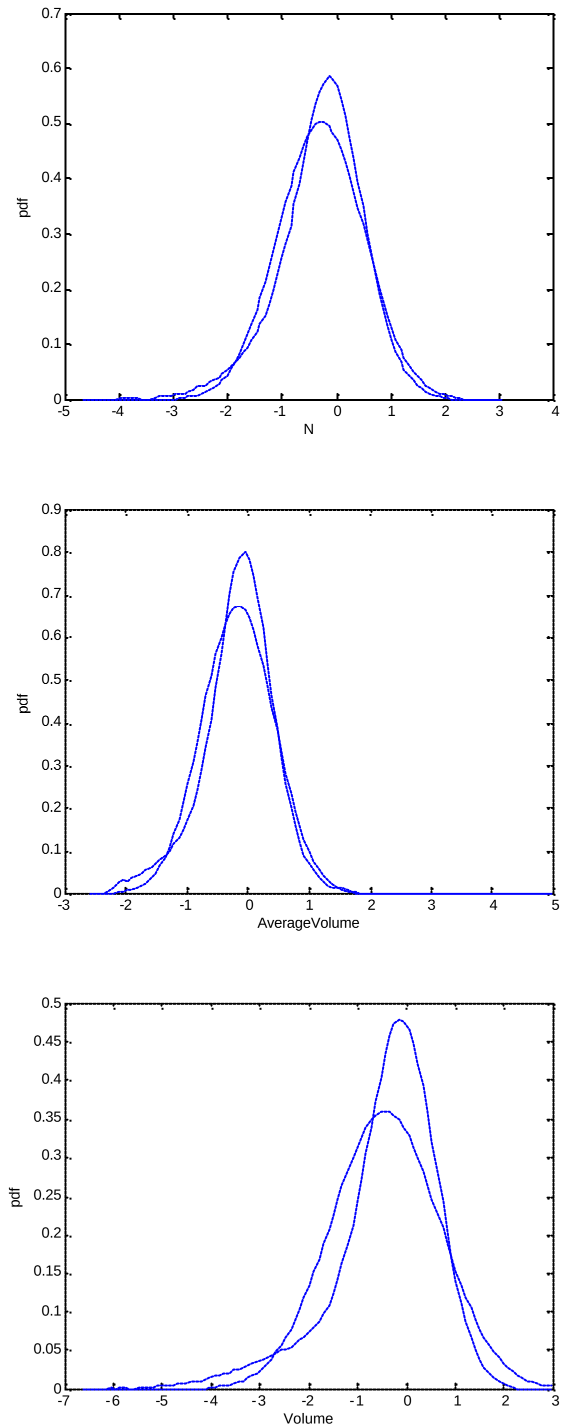
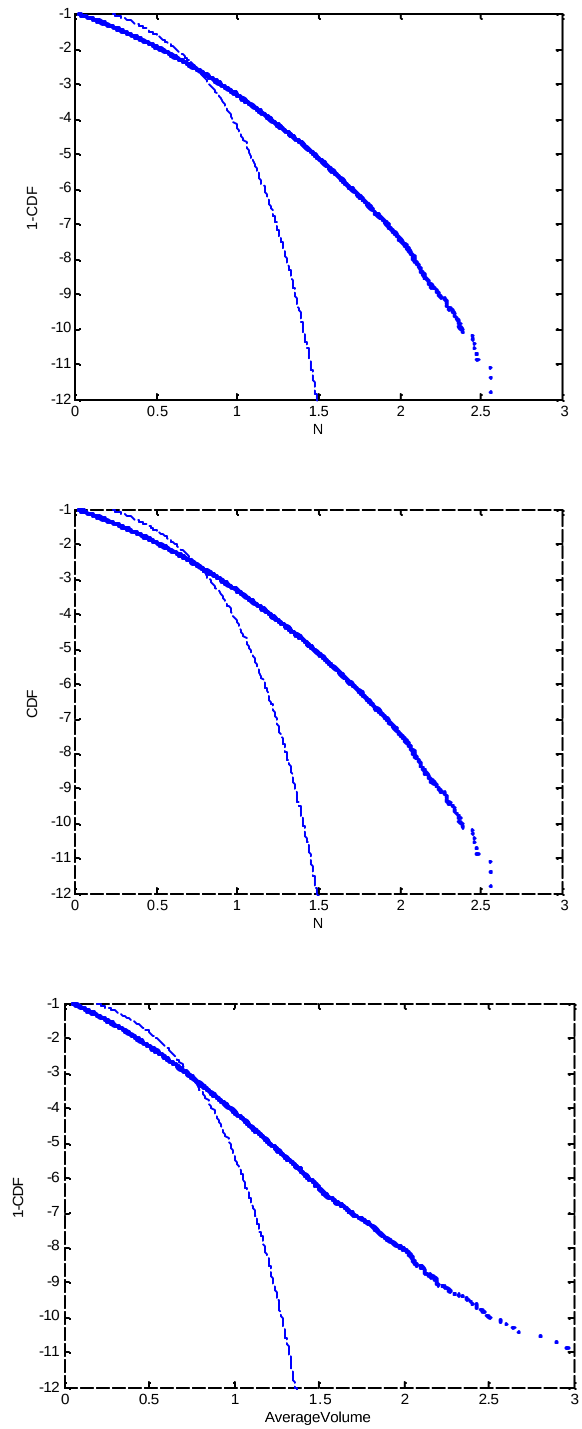


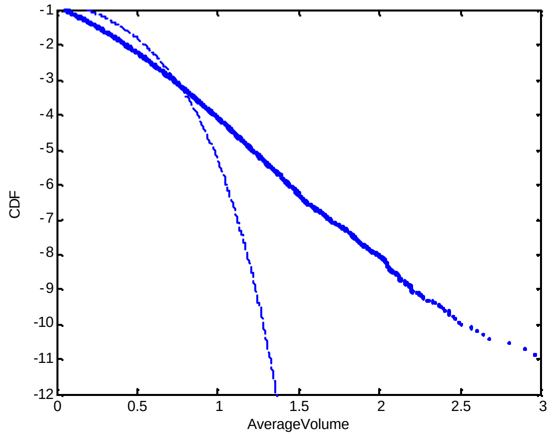
Figure 3



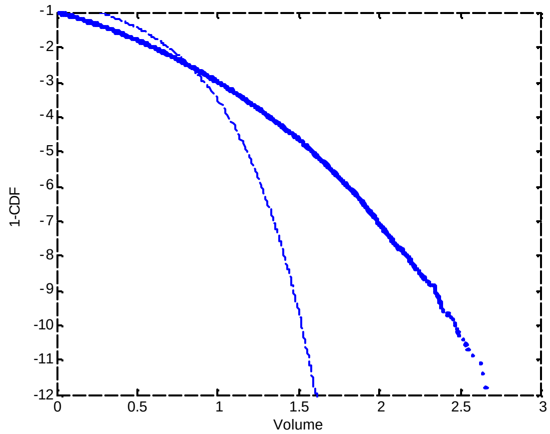
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
pdf
-5 -4 -3 -2 -1 0 1 2 3 4
N
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
pdf
-3 -2 -1 0 1 2 3 4 5
AverageVolume
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
pdf
-7 -6 -5 -4 -3 -2 -1 0 1 2 3
Volume

Figure 4

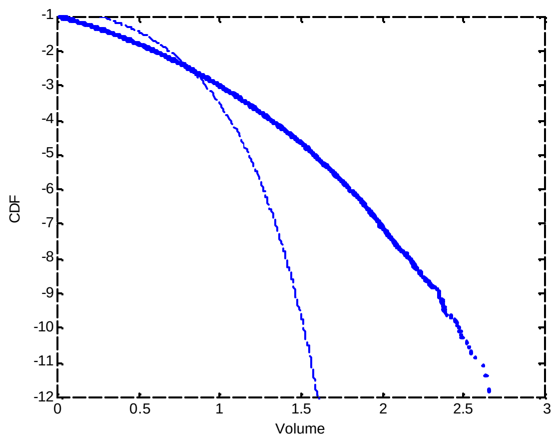




CDF
AverageVolume
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10
-11
-12
0
0.5
1
1.5
2
2.5
3

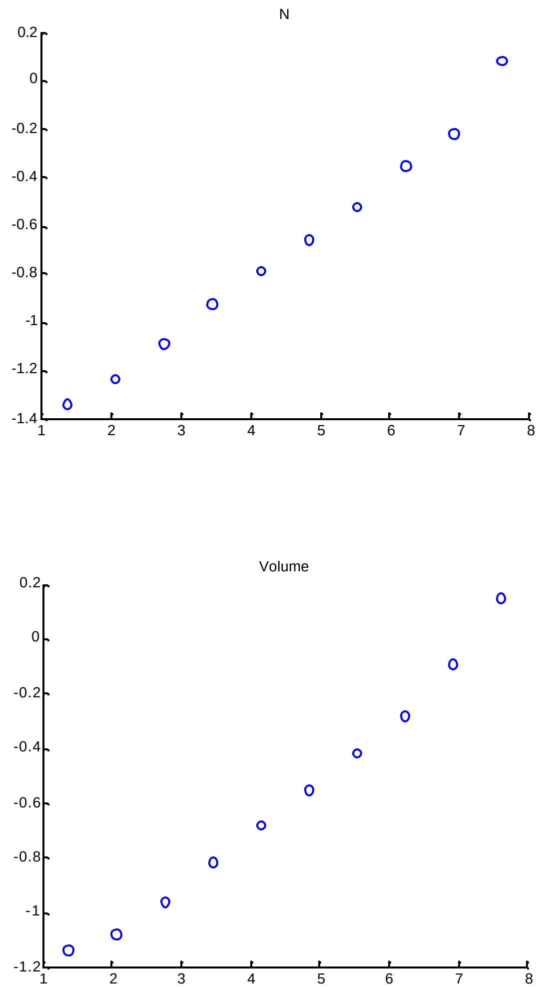


1-CDF
Volume
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10
-11
-12
0
0.5
1
1.5
2
2.5
3



CDF
Volume
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10
-11
-12
0
0.5
1
1.5
2
2.5
3

Figure 5


N
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
1 2 3 4 5 6 7 8
Volume
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
1 2 3 4 5 6 7 8