Large Impact Events and Finance

Marcus Davidsson

Nygatan 17, 55316 Jonkoping, Sweden

E-mail: davidsson_marcus@hotmail.com

Received: November 30, 2011 Accepted: December 9, 2011 Published: May 15, 2012

doi:10.5430/afr.v1n1p95 URL: http://dx.doi.org/10.5430/afr.v1n1p95

Abstract

This paper will discuss large impact events in financial markets. Two different datasets are investigated; daily data for the SP500 Index from the period 1950 to 2010 and monthly data from 1997-2010 for 23 global stock market indices. We find that the daily returns for the SP-500 are more volatile than expected. Two normality tests are also run on the global stock market indices dataset. The results are mixed whether the data is normally distributed or not. However a significant negative skew and a high degree of peakness (leptokurtic) are found to be present.

Keywords: Crash, Stock Market, Normality, Returns

1. Introduction and Literature Review

During the latest couple of centuries the traditionally financial models has come in to questioning due to the non normality observed in financial returns. Mandelbrot (1963) has empirically investigated the properties of the return distributions in finance. The author explains that financial returns do not tend to be normally distributed. He finds evidence of that the return distributions in finance are better described by stable Paretian distributions. Fama (1965) also finds empirical support for such a hypothesis. In his doctoral thesis he analyzed daily data for 30 stocks from the Dow-Jones Industrial Average for the period 1957 to 1962. He finds that the return distributions have fat tails and a large amount of kurtosis. However, he notes that changes in expected return do not seem to explain the fat tails of the return distributions. The Gaussian or normal distribution on the other hand does not represent a correct picture of return in financial markets. The author further explains that the mean-variance portfolio optimization framework is known to be highly sensitive to the estimates of the return variance. Fat tail distributions make such return variance estimation even harder hence alternative measures of dispersion might be necessary. This is also supported by authors such as Hagerman (1978). Other author such as Taleb (2008) explains that large impact events are far more likely to occur than predicted by the normal distribution. He also argues in his book *The Black Swan: The Impact of the Highly Improbable* that such events have at least three principle characteristics, they are unpredictable, they will have a large impact and after the fact we try to come up with explanations to make these events appear to be less random. The author argues that the reason why we underestimates their importance is because human tend to concentrate on things that we understand rather than focusing on things that we don't understand. One single observations is sometime enough to invalidate a general hypothesis that have been accepted for centuries. Such notion creates fragility and uncertainty.

Hood, Nofsinger & Small (2009) looked at return data for 716 hedge funds and they found that 72 percent statistically reject normality. Karoglou (2010) try to explain such a fact by noting that structural breaks are a major reason why return distributions in financial markets are non-normal. Kon (1984) looks at daily stock return and observes a significant kurtosis and positive skewness for the return distribution. The author also conducts stationarity tests that reveal that differences in variance can explain the observed kurtosis and differences in expected return can explain the observed skewness. However, other authors such as Zhou & Zhu (2010) argue that large negative returns can be explained by the traditional random walk model. They calculated that the probability of a 50 percent stock market drop is about 90 percent over a 100-year period. Hence, investor should prepare for volatility. Cross correlation can also be a potential explanatory factor why we would see large impact event more frequently than predicted by theory. Longin and Solnik (2001) explain that the relationship between cross correlation and volatility is difficult to test because of the spurious relationship between correlation and volatility. The authors use extreme value theory which finds support that cross correlation is related to expected return but not necessarily to market volatility. Correlation tends to increase in bear markets but not in bull markets. They also find that the negative tail but not the positive tail of the return distribution tend to be multivariate normally distributed. The importance of cross correlation to explain return volatility is supported by authors such as Markowitz (1959) which has noted that positive cross correlation will increase return variance hence when global markets around the world becomes more integrated a shock to the system will spread much faster than if the markets where less positive correlated. Arditti and Levy (1975) argues that the first three moments (expected return, return variance and return skew) all have economics importance but the fourth moment (kurtosis) tend to be less important. A normal distribution tend to be symmetrical i.e. large loses tend to be as frequent as large gains and small loses tend to be as frequent as small gains. A skewed distribution on the other hand tends to be asymmetric. A positive skewed distribution will have many small losses and a few large gains while a negative skewed distribution will have many small gains and a few large losses. Return skew tends to be highly important in for example state lotteries. Lotteries are designed in such a way that the cost of the lottery ticket is higher than then expected gain. The question then becomes why do people participate in such unfavorable games? Garrett and Sobel (1999) argue that price distribution skewness explains why risk averse individuals may play the lottery. People are aware that most lottery tickets will not result in any price but the small chance of getting a lottery tickets with a large payout makes up for such a small loss. Kahneman and Tversky (1979) also argue that people are not consistently risk averse. People tend to be risk-averse when it comes to gains i.e. they are more attracted by a certain small gain than an uncertain large gain and risk loving when it comes to losses i.e. they are more repelled by a certain small loss than an uncertain large loss. Kraus and Litzenberger (1976) explain that the traditional portfolio optimization model introduced by Markowitz assumes that financial returns are normally distributed and can be accurately described by the mean-variance framework. Fung and Hsieh (2001) on the other hand have illustrated that trend following return distributions tend to be asymmetric and have a positive skew. The trend following payoff tend to be option like with a limited loss and an unlimited gain. Such payoff is achieved by using a stop loss. A stop loss distribution will have a positive skew i.e. many small losses and a few large gains. In Mitton and Vorkink (2007) trading model portfolio returns of under diversified investors tend to be more positively skewed than those of diversified investors. Investors also tend to sacrifice mean-variance efficiency for a return distribution with larger skewness.

2. Empirical Analysis

We can start by looking at daily data for the SP500 Index from the period 1950 to 2010. I have in the exhibit-1 plotted the percentage return distribution, the theoretical normal distribution with the same mean and standard deviation as the empirical dataset, the frequency table, the theoretical Cumulative Distribution Function (CDF) for a normal distribution with the same mean and standard deviation as our empirical distribution and the CDF of these extreme returns. The first thing we notice is that the data does not appear to be normally distributed. We can observe both infrequent positive and negative outliers. The expected percentage return for our dataset is -0.022719 and the standard deviation is equal to 0.98475. We note that 35 observations (0.22 percent) are smaller than the expected percentage return - 4* standard deviations and 50 observations (0.35 percent) are larger than the expected percentage return + 4* standard deviations. The total number of daily percentage return observations are 15 546. This means that approximately 0.54 percent of the observations are outside the range expected return+/- 4*standard deviations [3.91, -3.96].

The CDF for a theoretical normal distribution with the same mean and standard deviation as the empirical distribution can also be seen in exhibit-1. The CDF will give you the probability of observing an observation equal to or smaller than x for a normal distribution. It turns out that 0.00003167 percent of the observations are smaller than the expected percentage return - 4* standard deviations and 0.00003167 are larger than the expected percentage return + 4* standard deviations. This means that 0.00006364 percent of the observations are outside the range expected return+/- 4*standard deviations [3.91, -3.96]. Another thing to note is that the expected percentage return for the largest 100 positive returns was 4.7 while the expected percentage return for the largest 100 negative returns was -4.05. Hence, we would expect to make more money than we lose on the 100 largest returns.

We can now introduce another dataset which consists of monthly data from 1997-2010 for 23 global stock market indices. The Shapiro and Wilk's W-test is used to test for normality. We start by simulating one Gumble(0,1) return series and one Normal(0,1) return series to test the power of the statistical test. The output can be seen in exhibit-2. We can see that the test appears to work i.e. it can differentiate between a gumble and a normal distribution. If the calculated p-value is smaller (larger) then 0.05 then we will reject (accept) the null hypothesis that the data is drawn from a normal distribution. For the simulated data we test if the sample is drawn from a normal distribution with zero expected return and standard deviation of one.

We now run the test on our empirical data set as seen in appendix-1. We can see that for 12 out of the 23 markets the hypothesis that the return distribution is normal is rejected. This means that for 11 markets there does not exist any evidence against the hypothesis that the return distribution is normal. What this has shown is that we would expect approximately half the markets to have returns that are normally distributed while the other half non-normal distributed returns. We can also run a Chi-Square suitable model test to test for normality. We start again by simulating one

Gumble(0,1) return series and one Normal(0,1) return series to test the power of the statistical test. The output can be seen in exhibit-3. We can see that the test appears to work i.e. it can differentiate between a gumble and a normal distribution.

We now run the test on the same empirical data set as previously. The expected return and standard deviation of return for the test statistics are calculated from the original dataset. We can see in appendix-2 that for 6 out of the 23 markets the hypothesis that the return distribution is normal is rejected. This means that for 17 markets there does not exist any evidence against the hypothesis that the return distribution is normal. What this has shown is that we would expect approximately seventy percent of the markets to have returns that are normally distributed while the thirty percent non-normal distributed returns. We have also in appendix-3 plotted the probability plot to detect normality for our global stock market indices dataset. Again the results are mixed whether the data is normally distributed or not. We can also calculate the skewness for the dataset. The skewness coefficient for the dataset is calculated as seen below. The expected skewness coefficient, which measures the symmetry of the distribution, for a normal distribution is equal to zero as seen in exhibit-4. The expected skew for the dataset is -0.37 hence our data exhibits negative skewness compared to a normal distribution.

$$Skewness(A) = \frac{N * CentralMoment(A,3)}{(N-1) * S\tan dardDeviation(A)^3}$$

We can also calculate the kurtosis for the dataset. The Kurtosis coefficient for the dataset is calculated as seen below. The expected kurtosis coefficient, which measures the peakness of the distribution, for a normal distribution is equal to 3 as seen in exhibit-5. The expected kurtosis for the dataset is 4.73 hence our data exhibits a high peakness (leptokurtic) compared to a normal distribution.

$$Kurtosis(A) = \frac{N * CentralMoment(A,4)}{(N-1) * Variance(A)^2}$$

3. Conclusion

We have in this paper discussed large impact events and normality in financial markets. Two different datasets are investigated; daily data for the SP500 Index from the period 1950 to 2010 and monthly data from 1997-2010 for 23 global stock market indices. We started out by looking at the daily SP500 dataset. We calculated the expected return and standard deviation of return for the dataset. We then calculated how many percent of the observations fell outside the mean +/- 4*standard deviation range. We compared that number to the number that we would expect from a normally distributed dataset with the same mean and standard deviation as the empirical dataset. Several conclusions can be drawn from such an investigation. i) The daily returns for the SP-500 are more volatile than expected i.e. we would expect less observations to fall outside our expected specified range. ii) Both large negative and positive returns are far more frequent than predicted by the normal distribution. iii) Large positive returns are more frequent than large negative returns. iiii) The expected percentage return for the largest 100 positive returns was larger than the 100 negative returns. We then looked at the monthly dataset for 23 global stock market indices. We ran the Shapiro and Wilk's W-test and the Chi-Square suitable model test for normality. The results were mixed whether financial markets are normally distributed or not. Since our analysis gave us mixed results as to normality of financial markets we also extended our analysis to skew and kurtosis for the dataset in order to get a more accurate description of what was going on. The expected skew for the dataset was -0.37 hence our data exhibits negative skewness compared to a normal distribution. The expected kurtosis for the dataset was 4.73 hence our data exhibits a high peakness (leptokurtic) compared to a normal distribution. Third (skew) and fourth (kurtosis) moments are important in order to try to understand the non-normality that certain financial markets exhibits.

References

Arditti, F and Levy, H. (1975). Portfolio Efficiency Analysis in Three Moments: The Multiperiod Case, *The Journal of Finance*, vol 30, no 3, pp 797-809. http://dx.doi.org/10.2307/2326859

Fama, E. (1965). The behavior of stock market prices, *Journal of Business*, vol 38, no 1, pp 34-105. http://dx.doi.org/10.1086/294743

Fung, W and Hsieh, D. (2001). The Risks in Hedge Fund Strategies: Theory and Evidence From Trend Followers, *Review of Financial Studies*, vol 14, no 2, pp 313-341. http://dx.doi.org/10.1093/rfs/14.2.313

Garrett, T and Sobel, R. (1999). Gamblers favor skewness, not risk: Further evidence from United States' lottery games, *Economics Letters*, volume 63, issue 1, Pages 85-90. http://dx.doi.org/10.1016/S0165-1765(99)00012-9

Hagerman, R. (1978). More Evidence on the Distribution of Security Returns. *The Journal of Finance,* vol. 33, no. 4, pp. 1213-1221. http://dx.doi.org/10.2307/2326950

Kahneman, D and Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk, *Econometrica,* vol 47, pp 263-291. http://dx.doi.org/10.2307/1914185

Karoglou, M. (2010). Breaking down the non-normality of stock returns, *The European journal of finance*, vol 16, issue 1, pp 79-95. http://dx.doi.org/10.1080/13518470902872343

Kon, S. (1984). Models of Stock Returns--A Comparison. *The Journal of Finance, v*ol. 39, no. 1, pp. 147-165. http://dx.doi.org/10.2307/2327673

Kraus,A and Litzenberger, R. (1976). Skewness Preference and the Valuation of Risk Assets. *The Journal of Finance*, vol. 31, no. 4, pp. 1085-1100. http://dx.doi.org/10.2307/2326275

Longin, F and Solnik, B. (2001). Extreme Correlation of International Equity Markets. The *Journal of Finance*, vol 56, issue 2, pp 649–676. http://dx.doi.org/10.1111/0022-1082.00340

Mandelbrot, B. (1963). The variation of certain speculative prices, *Journal of Business*, XXXVI, pp. 392–417.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*. New York: John Wiley & Sons

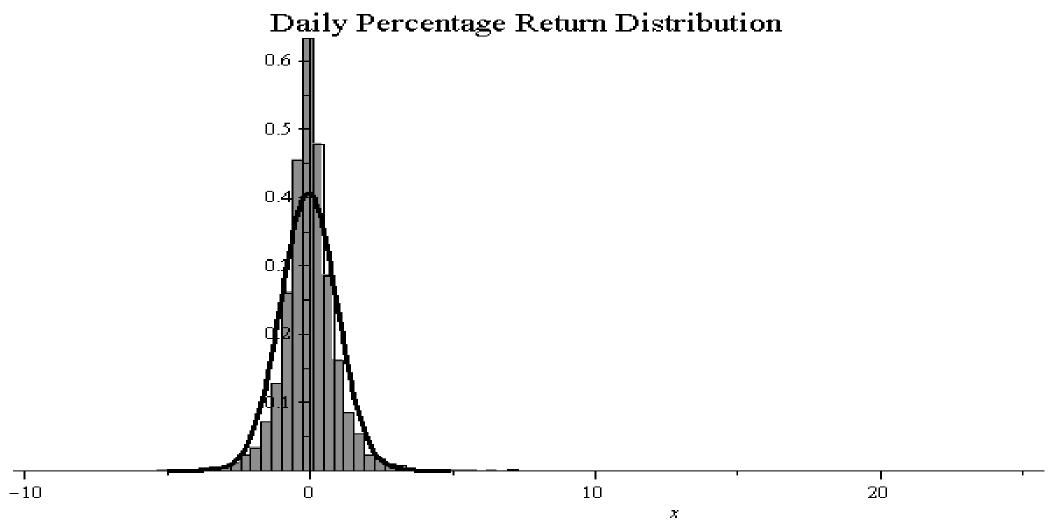
Mitton, T and Vorkink, K. (2007). Equilibrium Underdiversification and the Preference for Skewness, *Review of Financial Studies*, vol 20, issue 4, pp 1255-1288. http://dx.doi.org/10.1093/revfin/hhm011

Hood, M, Nofsinger, J & Small, K. (2009). Identifying the non-normality premium of an investment", *Managerial Finance*, Vol. 35 Iss: 4, pp.385 – 403. http://dx.doi.org/10.1108/03074350910935858

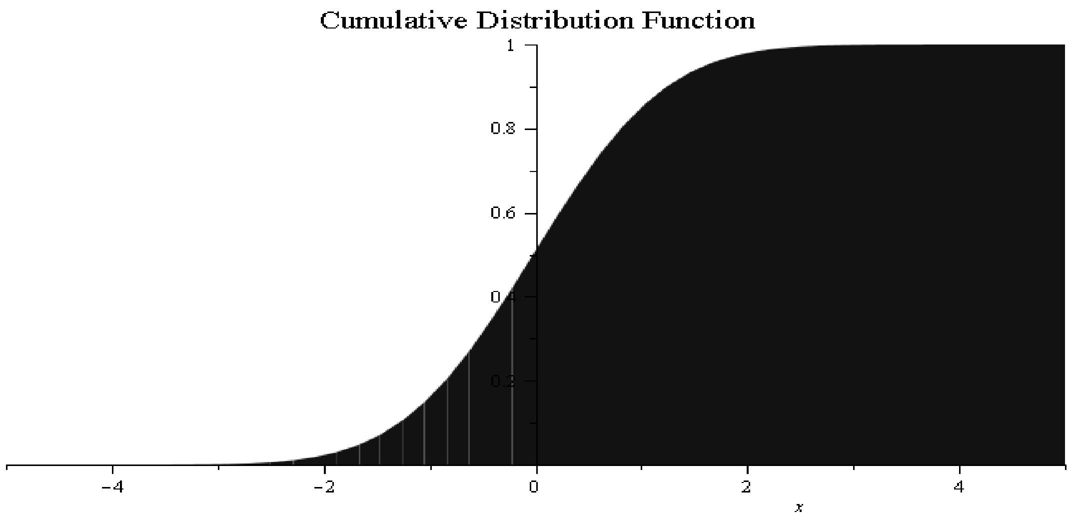
Taleb, N. (2008). *The Black Swan: The Impact of the Highly Improbable*,
Penguin Books, Limited (UK).

Zhou, G & Zhu, Y. (2010). Is the Recent Financial Crisis Really a "Once-in-a-Century" Event?, *Financial Analysts Journal*, vol. 66, No. 1, pp 3.

Exhibit-1 The Empirical SP-500 Dataset



Range	Abs Frequency	Percentage	Cum Frequency	Cum Percentage
-10.37823292 .. -6.767024109	3	0.019298	3	0.0193
-6.767024109 .. -3.155815298	85	0.546764	88	0.5661
-3.155815298 .. .4553935130	11776	75.749389	11864	76.3155
.4553935130 .. 4.066602324	3638	23.401518	15502	99.7170
4.066602324 .. 7.677811135	38	0.244436	15540	99.9614
7.677811135 .. 11.28901995	5	0.032163	15545	99.9936
11.28901995 .. 14.90022876	0	0.000000	15545	99.9936
14.90022876 .. 18.51143757	0	0.000000	15545	99.9936
18.51143757 .. 22.12264638	0	0.000000	15545	99.9936
22.12264638 .. 25.73385519	1	0.006433	15546	100.0000



Largest + Return	1-CDF Probability	Largest - Return	CDF Probability
25.73385519	0	-10.37823292	0
9.932366937	0	-9.738333457	0
9.805074674	0	-8.340428826	0
9.657273007	0	-6.608175764	0
9.026222164	0	-6.473241229	0
8.244680851	0	-6.078820394	4.00E-10
7.371805836	0	-5.985269594	7.00E-10
7.297760321	0	-5.94852693	9.00E-10
7.259654889	0	-5.421908161	2.10E-08
7.195258094	0	-5.139065126	1.02E-07
7.153153153	0	-5.130372875	1.07E-07
7.139156379	0	-5.062703205	1.54E-07
7.087538137	0	-4.8851267	3.95E-07
6.515809981	0	-4.866301459	4.36E-07
6.513922983	0	-4.843374436	4.91E-07
6.497691742	0	-4.782190463	6.72E-07
6.188447249	2.00E-10	-4.770845083	7.12E-07
6.088955362	2.00E-10	-4.694202721	1.05E-06
5.687465489	3.40E-09	-4.551451187	2.12E-06
5.576115844	6.50E-09	-4.547916652	2.16E-06
5.560628483	7.20E-09	-4.539618489	2.25E-06
5.47342485	1.20E-08	-4.526110206	2.40E-06
5.44039618	1.44E-08	-4.519613739	2.48E-06
5.292414464	3.38E-08	-4.44214876	3.60E-06
5.176314295	6.48E-08	-4.424205212	3.92E-06
5.165868756	6.86E-08	-4.393984075	4.52E-06
5.051449953	1.28E-07	-4.296587282	7.12E-06
5.022207038	1.50E-07	-4.212187319	1.05E-05
4.947292868	2.24E-07	-4.185194192	1.18E-05
4.946759453	2.25E-07	-4.167856803	1.28E-05
4.889986016	3.04E-07	-4.154130509	1.36E-05
4.773369489	5.57E-07	-4.077427808	1.92E-05
4.667514137	9.54E-07	-4.068413063	1.99E-05
4.61918698	1.22E-06	-4.018475751	2.48E-05
4.554379211	1.68E-06	-4.005766165	2.62E-05
4.512947397	2.05E-06	-3.922445936	3.75E-05
4.470260335	2.53E-06	-3.920625376	3.77E-05
4.442165409	2.89E-06	-3.913472041	3.89E-05
4.359531046	4.29E-06	-3.869872837	4.68E-05
4.351360945	4.46E-06	-3.855705306	4.96E-05
4.333614268	4.85E-06	-3.850503469	5.07E-05

Exhibit-2 Shapiro and Wilk's W-test for Simulated Data

Shapiro and Wilk's W-Test for Normality

Null Hypothesis:

Sample drawn from a population that follows a normal distribution

Alt. Hypothesis:

Sample drawn from population that does not follow a normal distribution

	Gumble(0,1)	Normal(0,1)
Sample size:	100	100
Computed statistic:	0.941602	0.97849
Computed pvalue:	0.000364451	0.449561
Result:	Rejected]	[Accepted]

Exhibit-3 Chi-Square Suitable Model Test for Simulated Data

Chi-Square Test for Suitable Probability Model

Null Hypothesis:

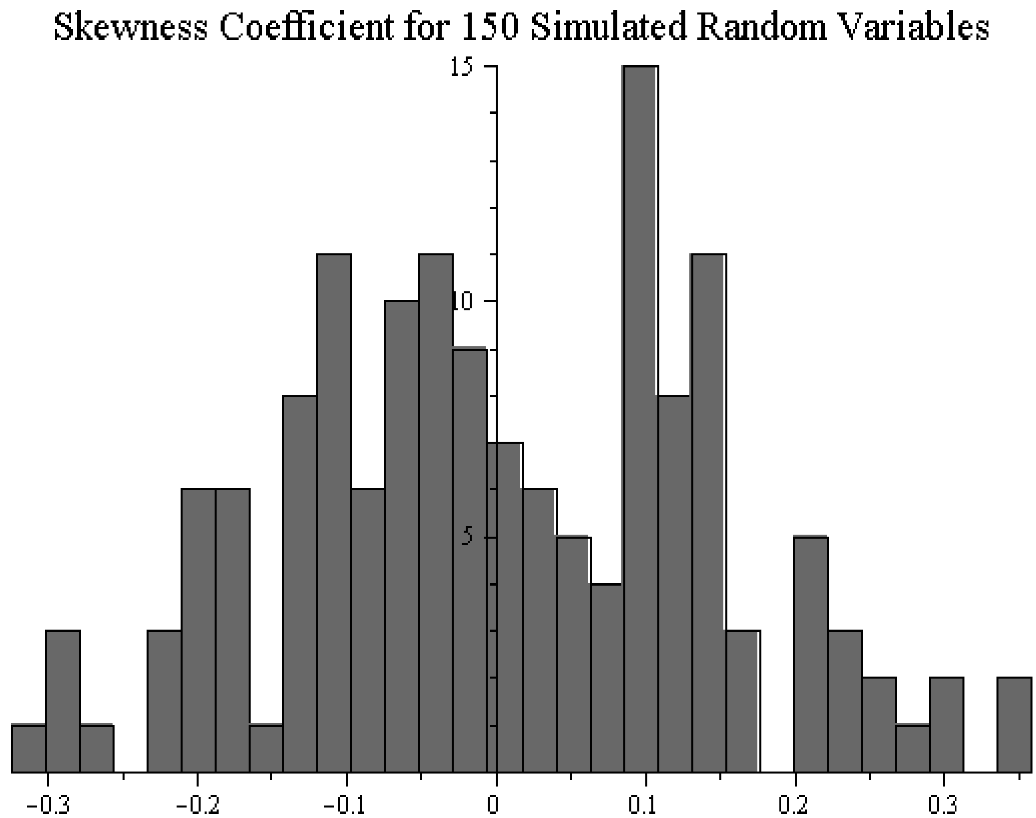
Sample was drawn from specified probability distribution

Alt. Hypothesis:

Sample was not drawn from specified probability distribution

	Gumble(0,1)	Normal(0,1)
Bins:	10	10
Distribution:	ChiSquare(9)	ChiSquare(9)
Computed statistic:	41.6532	3.61974
Computed pvalue:	3.8026e-06	0.934614
Critical value:	16.91897745	16.91897745
Result:	[Rejected]	[Accepted]

Exhibit-4 Skewness for Global Stockmarket Indices



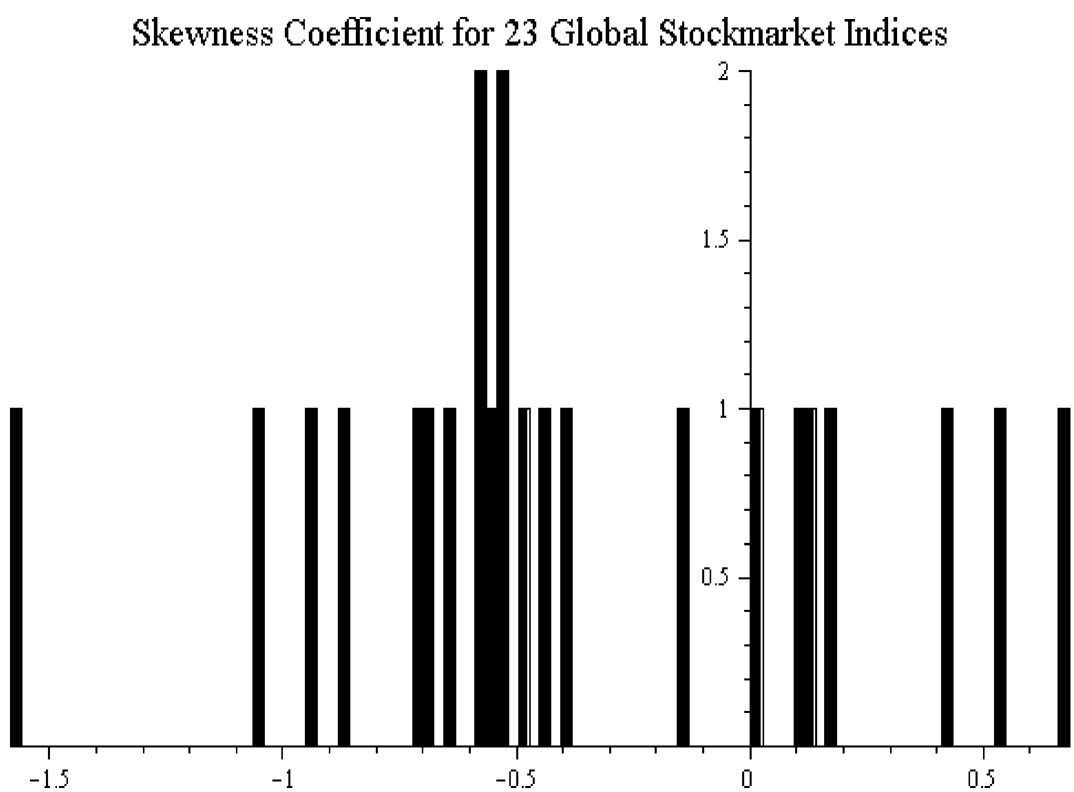
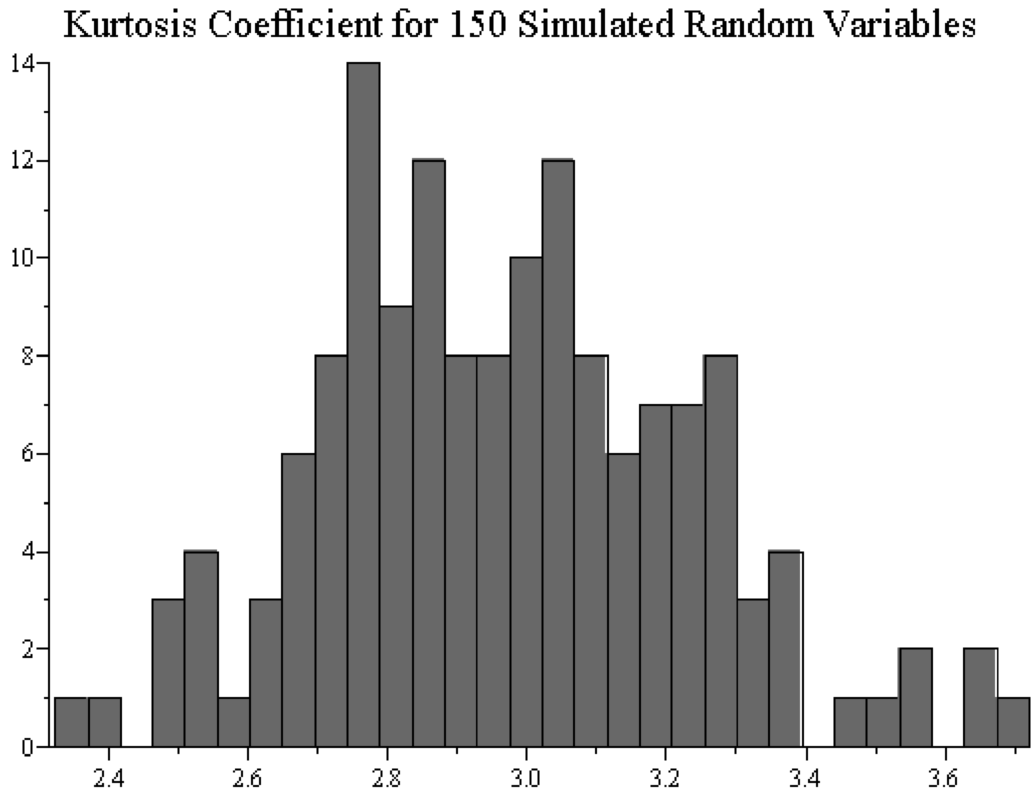
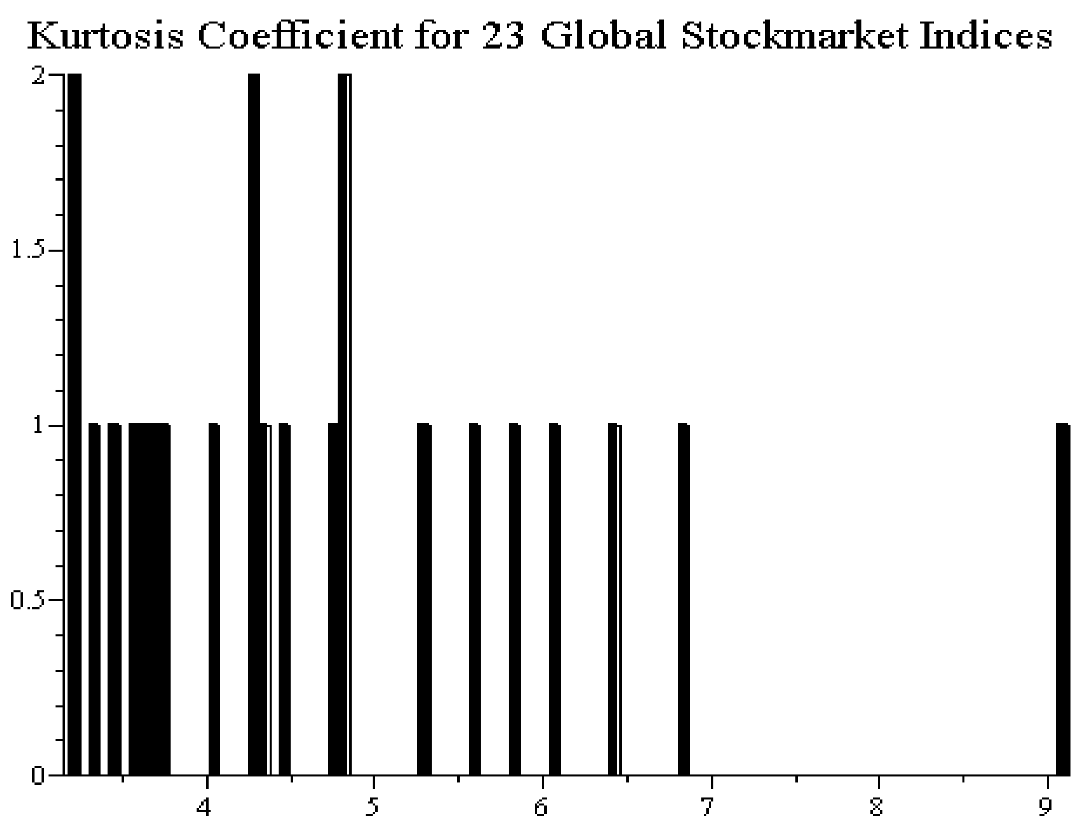


Exhibit-5 Kurtosis for Global Stockmarket Indices





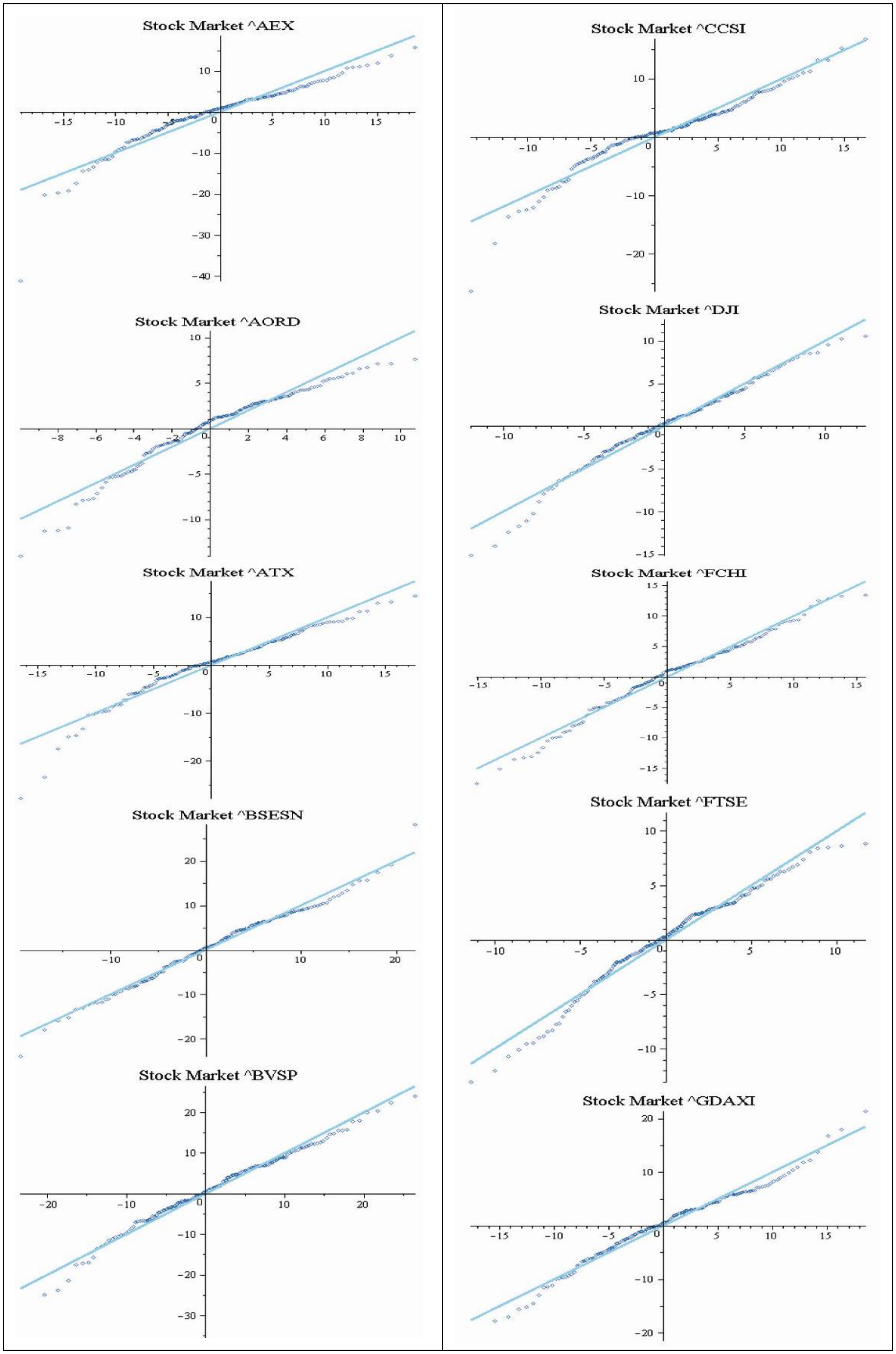
Appendix-1 Shapiro and Wilk's W-test for Empirical Data

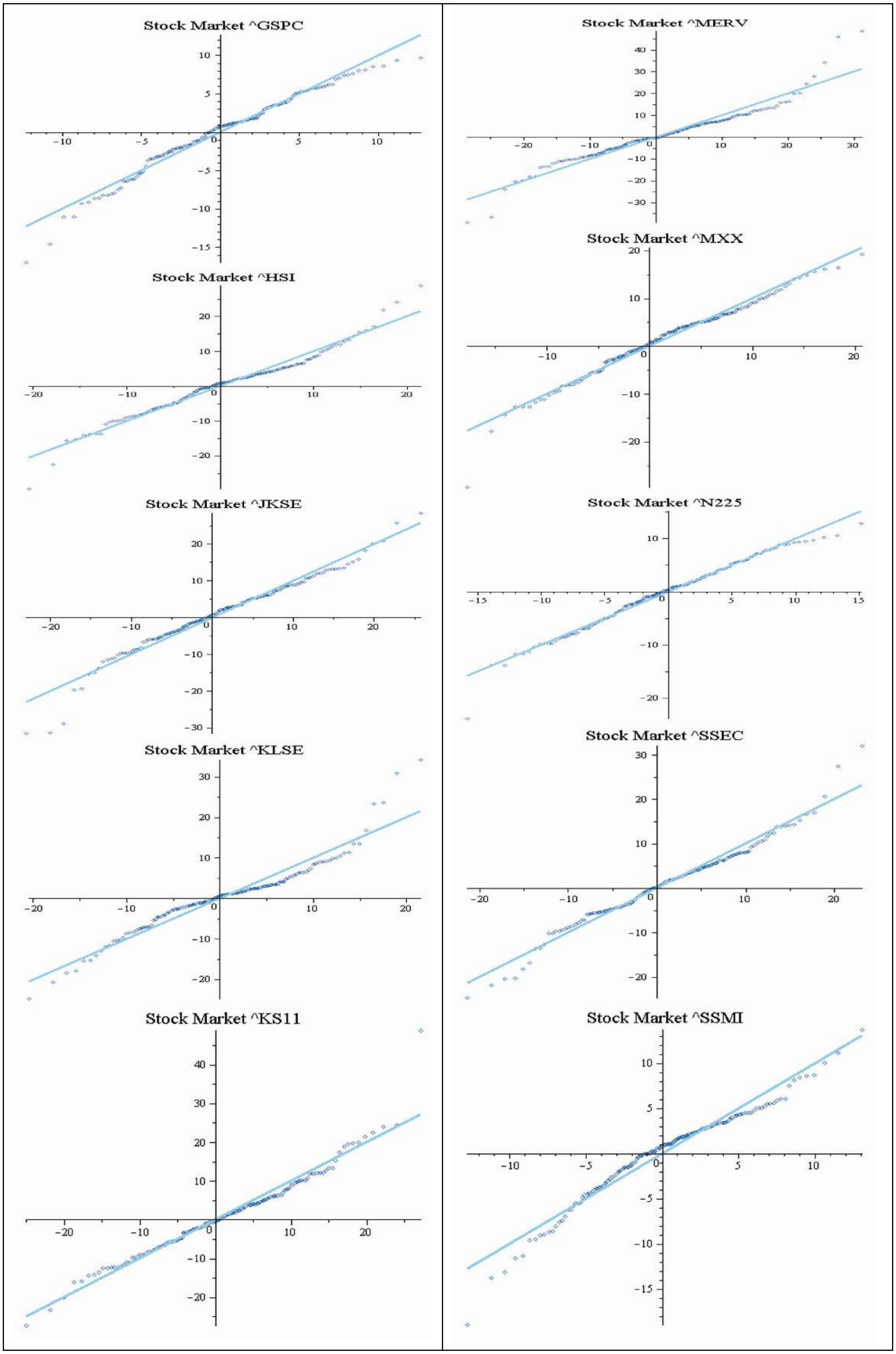
Market	Statistic	Pvalue	Hypothesis
^AEX	0.921297883	0.00000000	false
^AORD	0.939727254	0.00000048	false
^ATX	0.948438081	0.00002253	false
^BSESN	0.991527645	0.97348143	true
^BVSP	0.975217587	0.16359994	true
^CCSI	0.951073776	0.00006719	false
^DJI	0.970223781	0.04909275	false
^FCHI	0.973171096	0.10320194	true
^FTSE	0.961430132	0.00331648	false
^GDAXI	0.979573719	0.36931575	true
^GSPC	0.966042235	0.01479872	false
^HSI	0.981419617	0.48440251	true
^JKSE	0.968096955	0.02721701	false
^KLSE	0.940757083	0.00000077	false
^KS11	0.977891324	0.27736100	true
^MERV	0.948184663	0.00002024	false
^MXX	0.980661321	0.43572669	true
^N225	0.981302291	0.47677293	true
^SSEC	0.980538448	0.42800054	true
^SSMI	0.965883372	0.01409606	false
^STI	0.941595206	0.00000113	false
^TA100	0.974212448	0.13124103	true
^TWII	0.984316337	0.67526698	true

Appendix-2 Chi-Square Suitable Model Test

Market	Statistics	Criticalvalue	Pvalue	Hypothesis
^AEX	27.56821612	16.91897745	0.0011	false
^AORD	24.59057124	16.91897745	0.0035	false
^ATX	7.733540555	16.91897745	0.5612	true
^BSESN	6.275866053	16.91897745	0.7120	true
^BVSP	8.234737884	16.91897745	0.5107	true
^CCSI	17.82389846	16.91897745	0.0373	false
^DJI	6.556378848	16.91897745	0.6832	true
^FCHI	10.14499865	16.91897745	0.3389	true
^FTSE	12.88890166	16.91897745	0.1677	true
^GDAXI	6.113689704	16.91897745	0.7285	true
^GSPC	11.24743039	16.91897745	0.2591	true
^HSI	17.86852909	16.91897745	0.0367	false
^JKSE	10.85092113	16.91897745	0.2861	true
^KLSE	29.89733325	16.91897745	0.0005	false
^KS11	5.650214315	16.91897745	0.7744	true
^MERV	14.32115675	16.91897745	0.1114	true
^MXX	8.168843388	16.91897745	0.5172	true
^N225	7.840963357	16.91897745	0.5502	true
^SSEC	12.11604791	16.91897745	0.2068	true
^SSMI	15.82695292	16.91897745	0.0706	true
^STI	40.21627198	16.91897745	0.0000	false
^TA100	7.107844951	16.91897745	0.6259	true
^TWII	8.230805183	16.91897745	0.5111	true

Appendix-3 Probability Plot for Empirical Data




Stock Market ^GSPC
Stock Market ^MERV
Stock Market ^HSI
Stock Market ^MXX
Stock Market ^JKSE
Stock Market ^N225
Stock Market ^KLSE
Stock Market ^SSEC
Stock Market ^KS11
Stock Market ^SSMI