New Evidence on the Normality of Market Returns: The Dow Jones Industrial Average Case

Canegrati, Emanuele

Catholic University of the Sacred Heart - Milan

14 December 2008

New Evidence on the Normality of Market Returns: The Dow Jones Industrial Average Case

Emanuele Canegrati
Catholic University of Milan

December 15, 2008

Abstract

In this paper I test the normality of returns of the 30 components of the Dow Jones Industrial Average (DJIA) from January 1st 1990 to December 5th 2008. Results obtained by Kolmogorov - Smirnov, Shapiro - Wilk and Skewness - Kurtosis tests are robust in demonstrating that the hypothesis of normality can always be rejected.

1 Numerical Methods for the Analysis of Normality

Measures of dispersion are able to detect the degree of deviation of returns r_i from their mean $\overline{r}$. The second central moment, *variance*, is

$$s^2 = \frac{\sum (r_i - \overline{r})^2}{n-1} \tag{1}$$

The third moment, *skewness*, measures the symmetry of a probability distribution and it is expressed by

$$sk \triangleq \frac{E\left[(r-\overline{r})^3\right]}{\sigma^3} = \frac{\sum (r_i - \overline{r})^3}{s^3 (n-1)} = \frac{\sqrt{n-1}\sum (r_i - \overline{r})^3}{\left[(r_i - \overline{r})^2\right]^{\frac{3}{2}}} \tag{2}$$

If $sk < 0$, the distributions is said to be "left-skewed", because it has more observations on the right; otherwise it is said to be "right-skewed".

Finally, the fourth moment, *kurtosis*, measures the thinness of tails

$$ku \triangleq \frac{E\left[(r-\overline{r})^4\right]}{\sigma^4} = \frac{\sum (r_i - \overline{r})^4}{s^4 (n-1)} = \frac{(n-1)\sum (r_i - \overline{r})^4}{\left[(r_i - \overline{r})^2\right]^{2}} \tag{3}$$

If $ku > 3$, the distribution is called "thin-tailed", otherwise "fat-thailed". It is important to remark that a normal distribution has $sk = 0$ and $ku = 3$, so that a divergence from these values is an evidence of non-normality.

In this work I use three classic test to detect whether the distribution of market returns is normal: Kolmogorov - Smirnov (K - S) test, Shapiro - Wilk test and Skewness - Kurtosis test.

2 Non-parametric Analysis

2.1 Kernel Density Estimation

Kernel Density Estimation intends to give a shape to returns. Kernel estimators smooth out the contribution of each observed data point over a local neighbourhood of that data point. Data point r_i contributes to the estimate at point r depending on how distant r_i and r are. The extent of this contribution depends on two factors: the shape of the kernel function chosen and its bandwidth. The estimated density may be written as:

$$\widehat{\beta} = \frac{1}{n}\sum_{i=1}^{n} Ke\left(\frac{r - r_i}{j}\right)$$

where Ke is a kernel function, j the bandwidth and r the point where the density is evaluated. The Epanechnikov

$$Ke\,[z] = \begin{cases} \frac{3}{4}\left(1 - \frac{1}{5}z^2\right)/5 & \text{if } |z| < \sqrt{5} \\ 0 & \text{otherwise} \end{cases}$$

is the kernel function I used, since it is the most efficient in minimizing the mean integrated squared error. Notice that the choice of j will decide how many values are included in estimating the density at each point and in this model is determined as

$$m = \min\left(\sqrt{variance_r}, \frac{\text{interquartile range}_r}{1.349}\right)$$

$$j = \frac{0.9m}{n^{\frac{1}{5}}}$$

where r is the variable for which the kernel is estimated and n the number of observations.

2.2 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov tests the equality of the comulative density function obtained by the sample with the Normal comulative density function.

Given our sample of daily returns $\{R_i\}_{i=1,...,N}$ let us define a comulative density functions as

$$F_R^{(N)}(r) = \Pr\left\{R \le r\right\} = \frac{1}{N}\sum_{i=1}^{N} 1_{\{R_i \le r\}} \tag{4}$$

and the distance function between the actual CDF and the Normal CDF as

$$D_N = \sup_{-\infty < r < +\infty} \left| F_R^{(N)}(r) - F_0(r) \right| \tag{5}$$

The null hypothesis we want to test is

$$H_0 : F_R^{(N)}(r) = F_0(r)$$

against the alternative hypothesis

$$H_A : F_R^{(N)}(r) \neq F_0(r)$$

That is, we want to test that the distribution of returns we obtain from our dataset is a normal.

The Shapiro and Wilk test (1965) measures the ratio between the best estimator of the variance and the usual corrected sum of squares estimator of the variance.

$$W = \frac{\left(\sum \omega_i r_{(i)}\right)}{\sum \left(r_i - \overline{r}\right)^2} \tag{6}$$

where $\omega' = (\omega_1, ..., \omega_N) = \xi' \Omega^{-1} \left[\xi' \Omega^{-1} \Omega^{-1} \xi\right]^{-\frac{1}{2}}$, and $\xi' = (\xi_1, ..., \xi_N)$ is the vector of expected value of standard normal statistics, Ω the VCV matrix and $r' = (r_1, ..., r_N)$ is randomly chosen with $r_{(1)} < ... < r_{(N)}$.

The Jarque - Bera (Skewness - Kurtosis) tests that the expression

$$n\left[\frac{sk^2}{6} + \frac{(ku-3)^2}{24}\right]$$

is asymptotically distributed as χ^2, with two degrees of freedom.

3 Dataset

The dataset contains the historical series of daily returns for the thirty components of the Dow Jones Industrial Average Index. Appendix 1 reports the list of the components with relative tickers. The source is Yahoo's web site[1]. The historical series starts on January 1st 1990 until December 5th 2008. Returns are calculated starting from the closing prices P, using the following formula:

$$r_t = \frac{P_t}{P_{t-1}} - 1$$

[1] http://finance.yahoo.com

4 Main Findings

Table 1 reports overall statistics for the sample. Notice how the values for Skewness are always negative with respect to the mean (very close to zero), meaning that there are more observations on the right hand side. Secondly, note the very high values for Kurtosis, which are well above the number 3, the value of a Normal Distribution's Skewness. Therefore, the distribution is a "thin-tailed" type. Figure 1 shows the quantile-quantile plot (Q-Q plot) for every component of the DJIA, which compares ordered values of a variable with quintiles of a specific theoretical distribution (i.e., the standardised normal distribution).

[FIGURE 1 HERE]

A straight line made by points representing the singletons and passing through the origin with a slope equal to one is formed if the two distribution match. By inspecting the graphs it is easy to verify how data points are deviated from the straight fitted line. In particular, DJIA components appear to have many outliers both in the upper and lower extremes. Therefore, the graphical analysis shows without any doubt that equities do not follow a normal distribution. Of course we want to obtain also a numerical proof of our results. Table 1 - 3 show the results of the K - S, S - W and S - K tests, respectively. Again, results are very robust in demonstrating how distributions of returns are not normal.

[FIGURE 2 HERE]

The null hypothesis of equality between fitted distribution obtained from the sample and normal distribution is always rejected at the 1% of the confidence interval for all of the three tests.

5 Conlusions and discussion

The empirical evidence supplied in this work should be convincing enough that market returns do not follow a normal distribution. Both the graphical analysis and the numerical tests strongly reject the hypothesis of normality. The DJIA's components have negative skewness and high kurtosis. Which could be the implications of this finding? A very important consequence is related to risk management, in particular to the calculation of the Value-at-Risk. VaR is most of the time calculated according to famous and established tools, such as the Variance-Covariance method proposed by JPMorgan (1996) with its RiskMetrics^{TM} system. As we know, the classic version of the VCV method is based on the assumption that returns on the various assets obey a normal law. If this law is not respected, as our work demonstrates, the VCV method cannot be used and we must turn to other techniques, such as the historical simulation.

References

[1] D'Agostino, R. B. et al. (1990): *A Suggestion for Using Powerful and Informative Tests of Normality*, American Statistician, 44(4): 316-321

[2] Esch, L. et al. (2005): *Asset and Risk Management* - Risk Oriented Finance, John Wiley & Sons Ltd.

[3] J. P. Morgan (1996): *RiskMetrics^{TM} - Technical Document*, 4th edition, Morgan Guaranty Trust Company

[4] Park, H. M. (2008): Univariate Analysis and Normality Test Using SAS, Stata, and SPSS, The Trustees of Indiana University, http://www.indiana.edu/∼statmath

[5] Shapiro, S. S. and Wilk, M. B. (1965): *An Analysis of Variance Test for Normality (Complete Samples)*, Biometrika, 52(3/4): 591-611

6 Appendix

6.1 List of Dow Jones Industrial Average's components (with relative tickers)

AA	ALCOA INC
AXP	AMER EXPRESS INC
BA	BOEING CO
BAC	BK OF AMERICA CP
C	CITIGROUP INC
CAT	CATERPILLAR INC
CVX	CHEVRON CORP
DD	DU PONT E I DE NEM
DIS	WALT DISNEY-DISNEY C
GE	GEN ELECTRIC CO
GM	GEN MOTORS
HD	HOME DEPOT INC
HPQ	HEWLETT PACKARD CO
IBM	INTL BUSINESS MACH
INTC	Intel Corporation
JNJ	JOHNSON AND JOHNS DC
JPM	JP MORGAN CHASE CO
KFT	KRAFT FOODS INC
KO	COCA COLA CO THE
MCD	MCDONALDS CP
MMM	3M COMPANY
MRK	MERCK CO INC
MSFT	Microsoft Corporation
PFE	PFIZER INC
PG	PROCTER GAMBLE CO
T	AT&T INC.
UTX	UNITED TECH
VZ	VERIZON COMMUN
WMT	WAL MART STORES
XOM	EXXON MOBIL CP

Equity	Obs	Mean	Variance	Skewness	Kurtosis
AA	4773	-0,000081	0,0006826	-4,341065	96,70331
AXP	4772	0,000235	0,0005929	-4,065344	115,9066
BA	4770	0,0001414	0,0004531	-3,197039	73,35262
BAC	4771	0,0001034	0,0006313	-3,171425	84,21513
C	4769	0,0001203	0,0008221	-0,0360531	63,28877
CAT	4772	0,002279	0,0005708	-5,803072	126,8031
CVX	4773	0,0002231	0,0003598	-7,305726	207,6356
DD	4772	-0,0000536	0,0004666	-8,885345	262,9335
DIS	4771	0,0000536	0,0006044	-9,581025	284,0721
GE	4768	0	0,0004975	-10,1463	274,2428
GM	4770	-0,00019	0,0006623	0,536797	26,1319
HD	4771	0,0002608	0,0006549	-3,796777	59,45416
HPQ	4769	0,0003295	0,0007932	-2,838647	56,46649
IBM	4772	0,0002231	0,0004757	-4,980018	124,6709
INTC	4770	0,0003192	0,0009756	-4,149777	69,17098
JNJ	4767	0,0002109	0,0003877	-10,58064	275,518
JPM	4767	0,000336	0,0006767	-1,644472	45,26433
KFT	1877	0,0000493	0,0002031	-0,4265233	10,92126
KO	4770	0,0001145	0,0004005	-9,311633	236,2209
MCD	4771	0,0003465	0,0003982	-6,804407	179,5469
MMM	4769	0,0001203	0,0003287	-8,809903	250,8141
MRK	4771	0,0000756	0,0004907	-8,821152	245,5308
MSFT	4772	0,0001736	0,0008242	-6,359202	109,0631
PFE	4772	0,0000698	0,0005901	-9,441603	222,3284
PG	4772	0,0002189	0,0004015	-10,52393	259,2541
T	4770	0,0000538	0,0004259	-5,21559	131,7076
UTX	4772	0,0002568	0,0004782	-8,081657	190,0385
VZ	4773	0,0000249	0,0004114	-6,018771	158,411
WMT	4771	0,0003099	0,0004887	-6,203852	140,6709
XOM	4770	0,0002787	0,0003423	-7,518671	211,2723

Figure 1: Overall Statistics; source: *Yahoo*

Equity	D	P-value	Corrected	Normality?
AA	0,0952	0,000	0,000	NO
AXP	0,0798	0,000	0,000	NO
BA	0,0762	0,000	0,000	NO
BAC	0,1136	0,000	0,000	NO
C	0,114	0,000	0,000	NO
CAT	0,0874	0,000	0,000	NO
CVX	0,0918	0,000	0,000	NO
DD	0,0986	0,000	0,000	NO
DIS	0,1048	0,000	0,000	NO
GE	0,1247	0,000	0,000	NO
GM	0,0773	0,000	0,000	NO
HD	0,0938	0,000	0,000	NO
HPQ	0,0818	0,000	0,000	NO
IBM	0,0907	0,000	0,000	NO
INTC	0,0877	0,000	0,000	NO
JNJ	0,1119	0,000	0,000	NO
JPM	0,0967	0,000	0,000	NO
KFT	0,0732	0,000	0,000	NO
KO	0,1121	0,000	0,000	NO
MCD	0,081	0,000	0,000	NO
MMM	0,0982	0,000	0,000	NO
MRK	0,0997	0,000	0,000	NO
MSFT	0,1156	0,000	0,000	NO
PFE	0,1036	0,000	0,000	NO
PG	0,1205	0,000	0,000	NO
T	0,0911	0,000	0,000	NO
UTX	0,1104	0,000	0,000	NO
VZ	0,085	0,000	0,000	NO
WMT	0,0863	0,000	0,000	NO
XOM	0,0914	0,000	0,000	NO

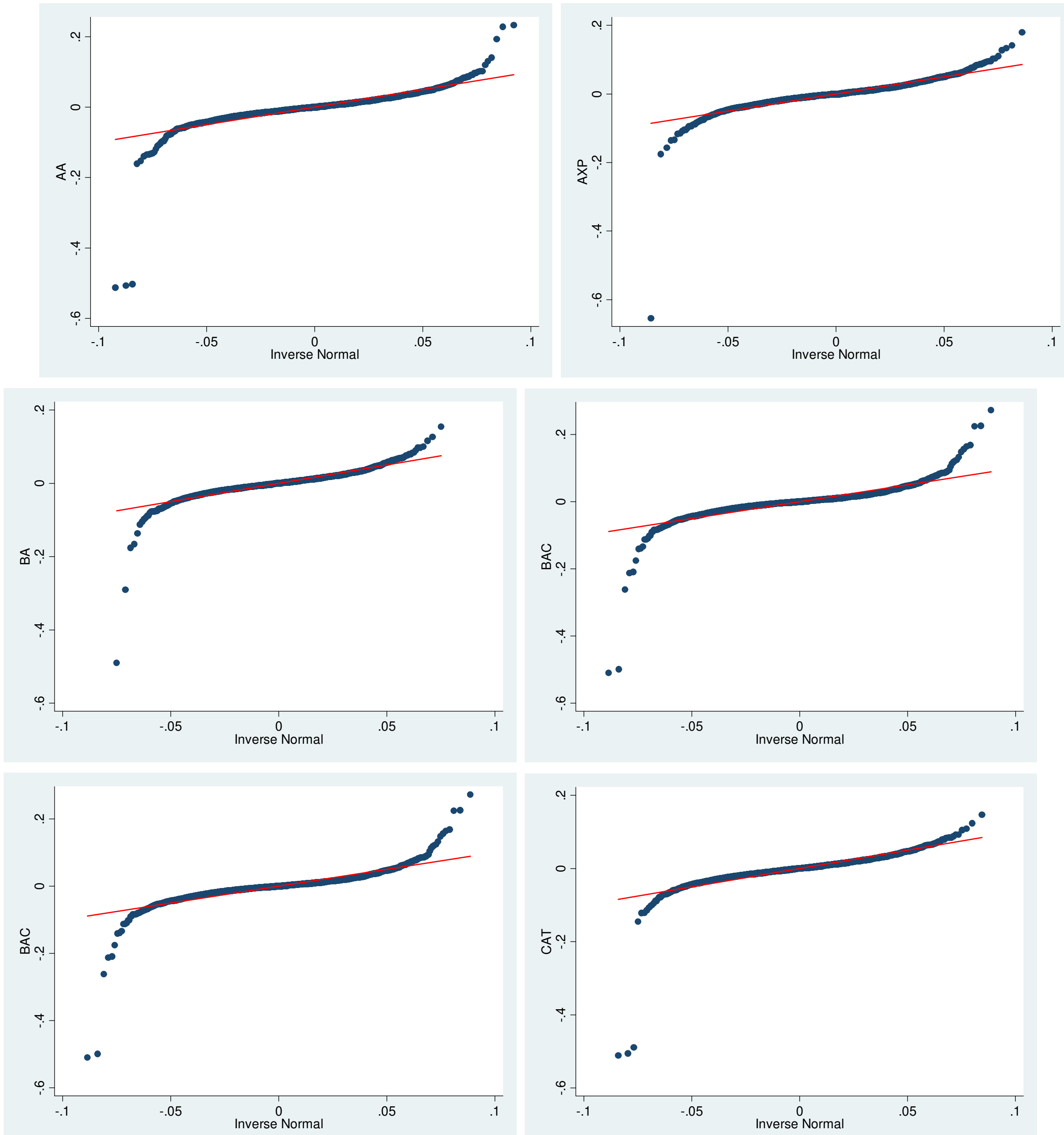
Figure 2: Kolmogorov - Smirnov Test

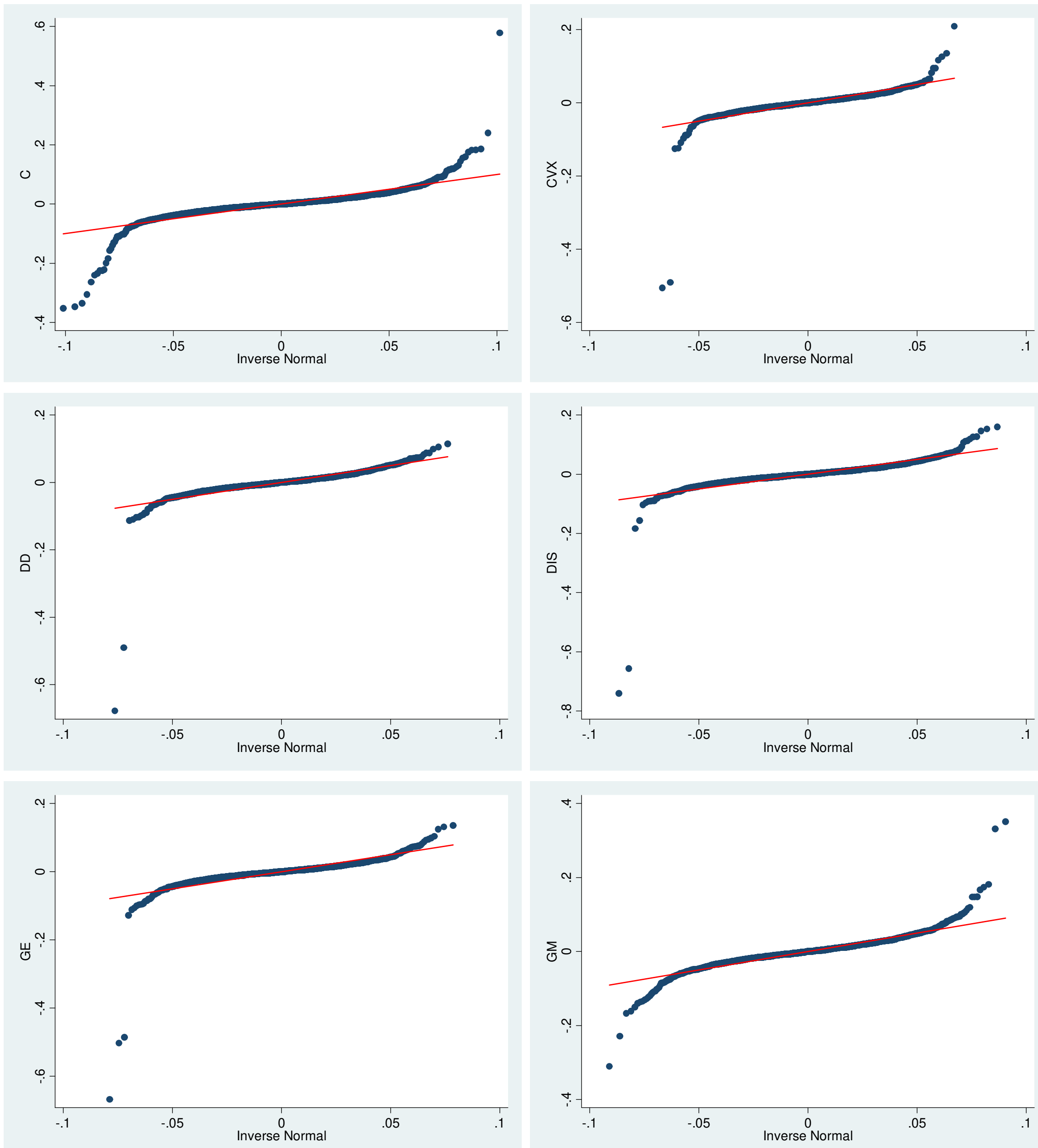
Equity	W	V	z	Prob>z	Normality?
AA	0,77224	592,087	16,726	0,00000	NO
AXP	0,83723	423,049	15,845	0,00000	NO
BA	0,84519	402,213	15,712	0,00000	NO
BAC	0,75628	633,351	16,902	0,00000	NO
C	0,75726	630,555	16,89	0,00000	NO
CAT	0,76547	609,574	16,802	0,00000	NO
CVX	0,72355	718,655	17,233	0,00000	NO
DD	0,72038	726,769	17,262	0,00000	NO
DIS	0,68409	820,931	17,582	0,00000	NO
GE	0,64169	930,592	17,91	0,00000	NO
GM	0,87483	325,21	15,155	0,00000	NO
HD	0,79242	539,434	16,481	0,00000	NO
HPQ	0,83144	437,865	155,935	0,00000	NO
IBM	0,78155	567,777	16,616	0,00000	NO
INTC	0,79361	536,238	16,466	0,00000	NO
JNJ	0,64453	92,035	17,888	0,00000	NO
JPM	0,84984	389,93	15,631	0,00000	NO
KFT	0,92907	79,401	11,102	0,00000	NO
KO	0,66624	867,164	17,725	0,00000	NO
MCD	0,76393	613,46	16,818	0,00000	NO
MMM	0,70721	760,562	17,381	0,00000	NO
MRK	0,70711	761,104	17,383	0,00000	NO
MSFT	0,6851	818,477	17,574	0,00000	NO
PFE	0,66257	877,03	17,755	0,00000	NO
PG	0,62557	973,182	18,027	0,00000	NO
T	0,79095	543,148	16,499	0,00000	NO
UTX	0,69362	796,309	17,502	0,00000	NO
VZ	0,77208	592,208	16,727	0,00000	NO
WMT	0,75891	626,503	16,873	0,00000	NO
XOM	0,72064	725,822	17,259	0,00000	NO

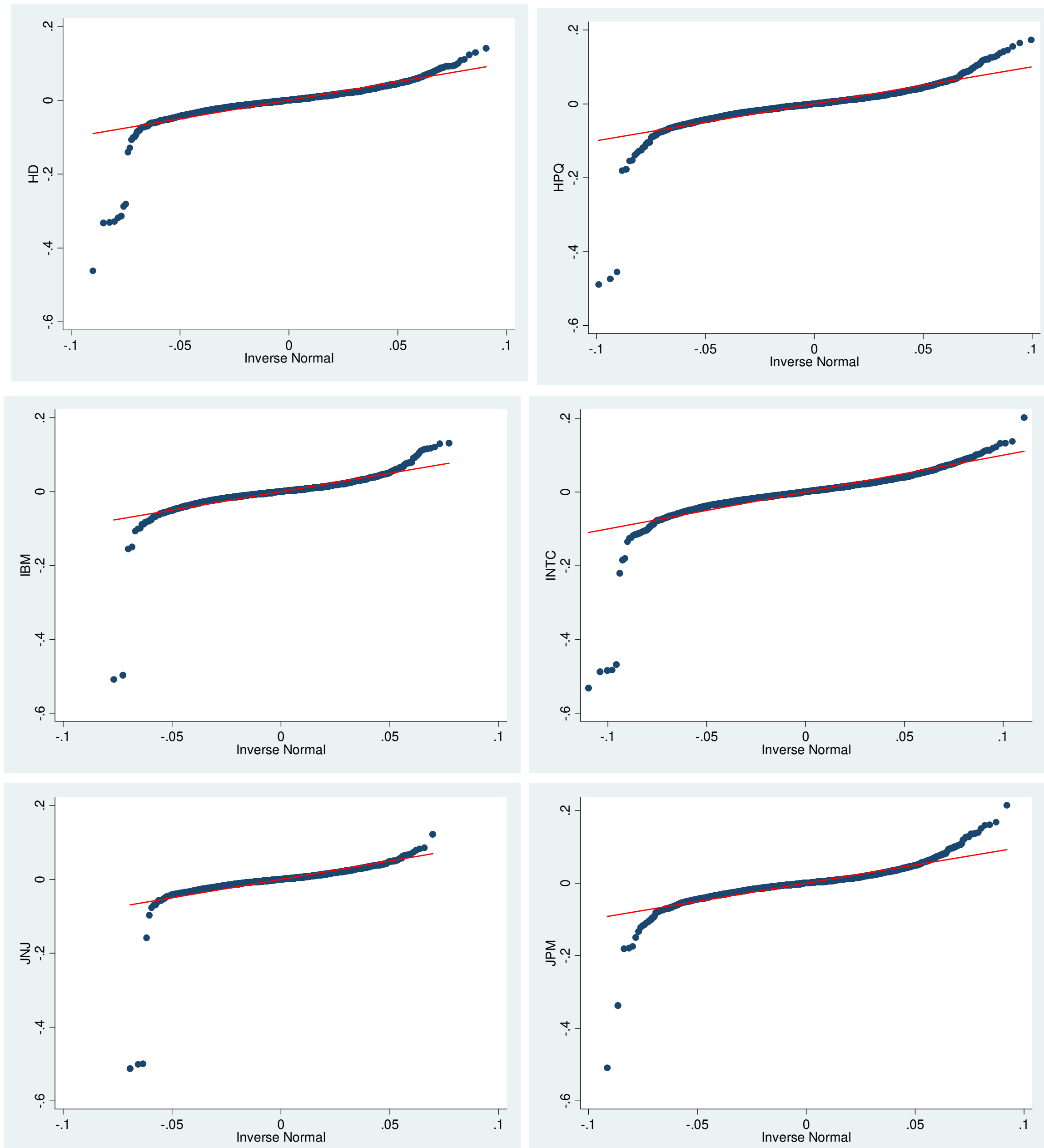
Figure 3: Shapiro - Wilk test for normality

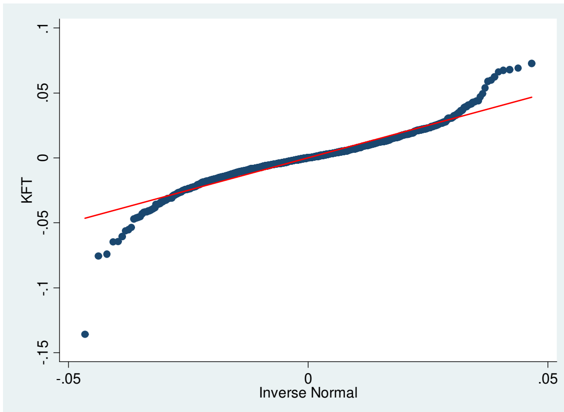
Eqiuty	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	Prob>chi2	Normality?
AA	0,000	0,000	-	-	NO
AXP	0,000	0,000	-	-	NO
BA	0,000	0,000	-	-	NO
BAC	0,000	0,000	-	-	NO
C	0,031	0,000	-	-	NO
CAT	0,000	0,000	-	-	NO
CVX	0,000	0,000	-	-	NO
DD	0,000	0,000	-	-	NO
DIS	0,000	0,000	-	-	NO
GE	0,000	0,000	-	-	NO
GM	0,000	0,000	-	-	NO
HD	0,000	0,000	-	-	NO
HPQ	0,000	0,000	-	-	NO
IBM	0,000	0,000	-	-	NO
INTC	0,000	0,000	-	-	NO
JNJ	0,000	0,000	-	-	NO
JPM	0,000	0,000	-	-	NO
KFT	0,000	0,000	-	-	NO
KO	0,000	0,000	-	-	NO
MCD	0,000	0,000	-	-	NO
MMM	0,000	0,000	-	-	NO
MRK	0,000	0,000	-	-	NO
MSFT	0,000	0,000	-	-	NO
PFE	0,000	0,000	-	-	NO
PG	0,000	0,000	-	-	NO
T	0,000	0,000	-	-	NO
UTX	0,000	0,000	-	-	NO
VZ	0,000	0,000	-	-	NO
WMT	0,000	0,000	-	-	NO
XOM	0,000	0,000	-	-	NO

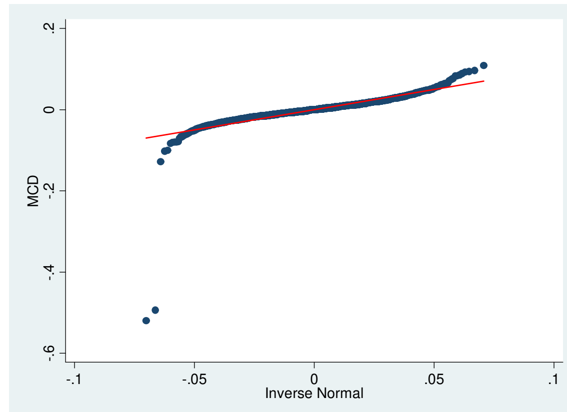
Figure 4: Skewness - Kurtosis test for normality

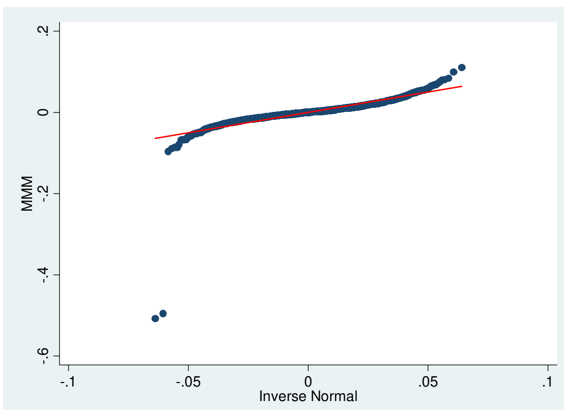

AA
AXP
BA
BAC
BAC
CAT
Inverse Normal
-.1
-.05
0
.05
.1
.2
0
-.2
-.4
-.6

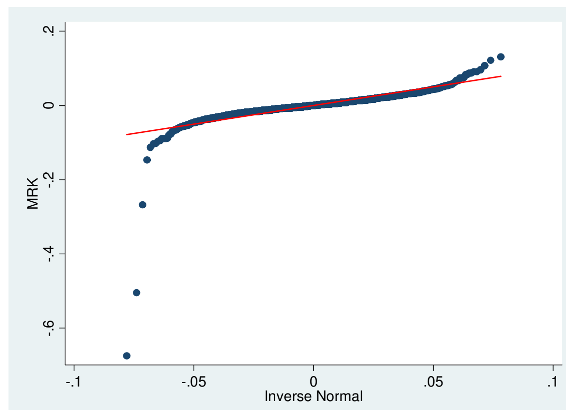

C
Inverse Normal
CVX
Inverse Normal
DD
Inverse Normal
DIS
Inverse Normal
GE
Inverse Normal
GM
Inverse Normal

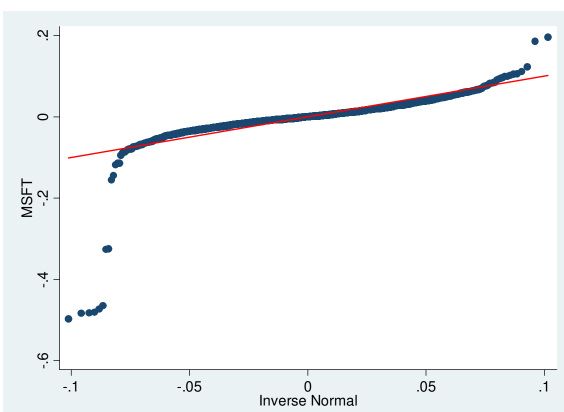

HD
HPQ
IBM
INTC
JNJ
JPM
Inverse Normal
-.6 -.4 -.2 0 .2
-.1 -.05 0 .05 .1

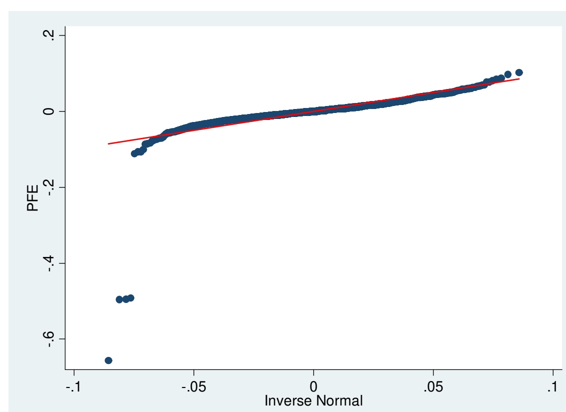

KFT
-.15
-.1
-.05
0
.05
.1
-.05
0
.05
Inverse Normal

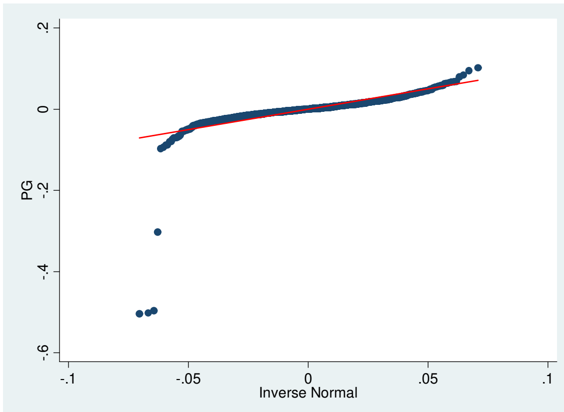

MCD
-.6
-.4
-.2
0
.2
-.1
-.05
0
.05
.1
Inverse Normal

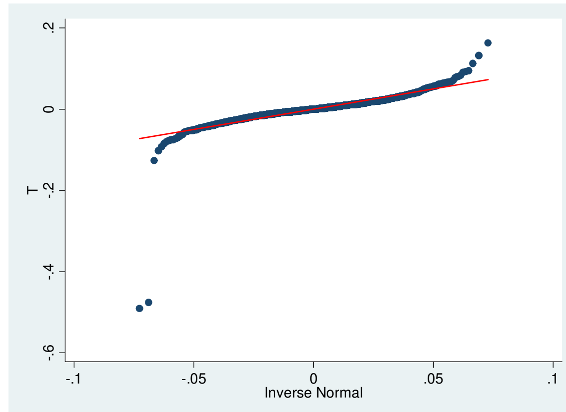

MMM
-.6
-.4
-.2
0
.2
-.1
-.05
0
.05
.1
Inverse Normal

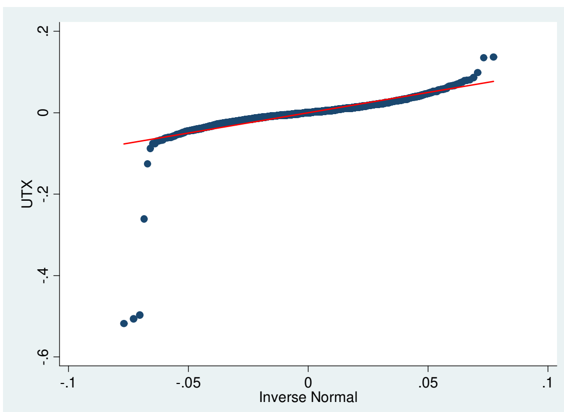

MRK
-.6
-.4
-.2
0
.2
-.1
-.05
0
.05
.1
Inverse Normal

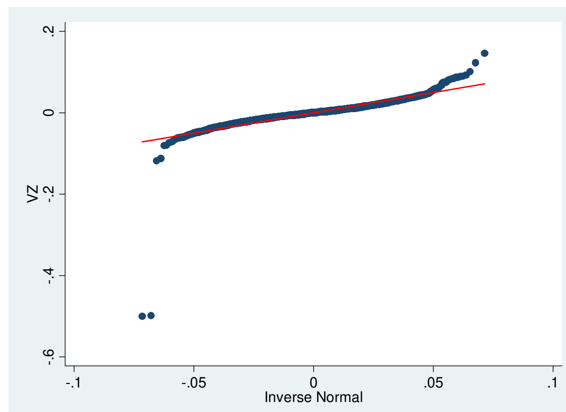

MSFT
-.6
-.4
-.2
0
.2
-.1
-.05
0
.05
.1
Inverse Normal

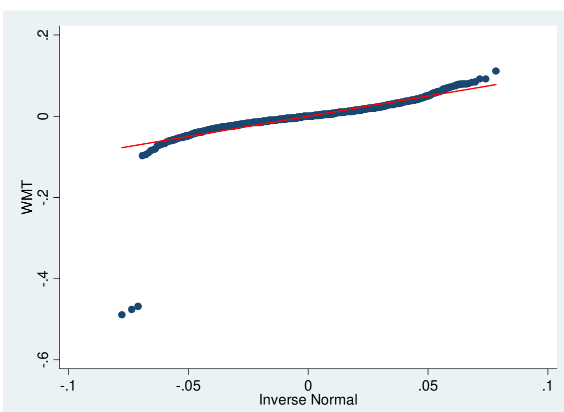

PFE
-.6
-.4
-.2
0
.2
-.1
-.05
0
.05
.1
Inverse Normal

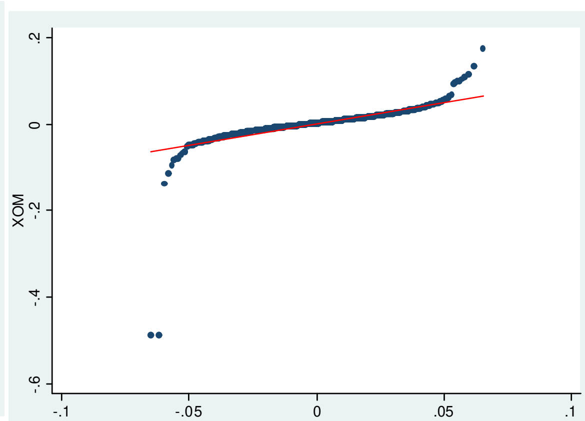

PG
.2
0
-.2
-.4
-.6
-.1
-.05
0
.05
.1
Inverse Normal


T
.2
0
-.2
-.4
-.6
-.1
-.05
0
.05
.1
Inverse Normal


UTX
.2
0
-.2
-.4
-.6
-.1
-.05
0
.05
.1
Inverse Normal


VZ
.2
0
-.2
-.4
-.6
-.1
-.05
0
.05
.1
Inverse Normal


WMT
.2
0
-.2
-.4
-.6
-.1
-.05
0
.05
.1
Inverse Normal


XOM
.2
0
-.2
-.4
-.6
-.1
-.05
0
.05
.1