World Applied Sciences Journal 3 (2): 254-258, 2008
ISSN 1818-4952

Modelling Extreme Financial Returns of Malaysian Stock Exchange

[1]Chin Wen Cheong, [2]Abu Hassan Shaari Mohd Nor and [3]Zaidi Isa

1Faculty of Information Technology, Multimedia University 63100 Cyberjaya, Selangor, Malaysia
[2]Faculty of Economic and Business, National University Malaysia, 43600 Bangi, Selangor, Malaysia
[3]Faculty of Science and Technology, National University Malaysia, 43600 Bangi, Selangor, Malaysia

Abstract: In this paper, we study the extreme behaviour of Malaysian stock exchange returns. Ten stock indices have been selected to investigate the possible similarities and divergences in their tail properties. Our empirical results evidenced that the tail realizations violated the normality and fitted well with heavy-tailed Pareto distribution with a mixture of positive and negative skewness. These findings provided non-trivial information to the investors who involved in long and short financial positions in Malaysian stock exchange.

Key word: Pareto distribution . goodness-of-fit test . financial time series . extreme value theory

INTRODUCTION

The tail behaviour of asset pricing provides important information for nowadays financial markets. Especially from the point of view in market risk management, extreme swings in the asset returns have major impacts to the derivatives hedging and portfolio management. The drastic asset price movements also give important influences in financial stability as well as monetary polices in general. The Gaussian distribution tail assumption in financial asset returns is unable to portray the fat-tails syndrome that commonly exhibited in most of the worldwide financial markets. Literatures [1-4], among others, reported that empirical studies in financial asset returns such as stock market, foreign exchange, bond, etc. are found to be heavier than a normal distribution. In addition, the IID (identical and independent distributed) property in Gaussian distribution with the assumption of complete randomness normally violated the actual market conditions. Thus, using the Gaussian distribution without adjustment in risk management may underestimate the probability of creating large unpleasant losses.

Besides the heavy-tailed issue, asymmetry distribution also often observed in financial time series. Studies by [5, 6] implemented skewed distributions that allowed upper and lower tails to have dissimilar behaviours. This property is very important in risk analysis to determine the Value-at-Risk where the long and short position investments over a given time period relied heavily on the lower and upper tails behaviours.

In this research, we take into account the heavy-tailed and asymmetry properties by studying separately the upper and lower tails behaviours. The shapes of the tails are estimated using Pareto-type distribution. Our empirical results evidenced the heavy-tailed and asymmetry behaviours in all the selected indices.

DATA SOURCE

This study evaluated the Kuala Lumpur Stock Exchange (KLSE) indices which consisted of Composite Index (CI) and the nine major sectoral indices. All the data are taken from *Datastream* from 25 Oct 1993 until 31 May 2007 with a total of 3569 observations for each series. According to the *Datastream*, all the selected sectoral indices are available during this period of time. This is important for us to investigate the possible similarities and divergences in their returns series. The percentage continuous compounded interday returns can be expressed as:

$$r_t = 100(\ln P_{t,close} - \ln P_{t-1,close})$$

METHODS

Empirical power laws: In this study, we have selected the power laws Pareto distribution as our framework to study the tail behaviour. The Pareto distribution can be obtained from the reparameterized of Generalized Pareto distribution. The Pareto distribution is related to extreme value theory where the Type II Frèchet distribution can be generated from the Pareto

Corresponding Author: Mr. Chin Wen Cheong, Faculty of Information Technology, Multimedia University 63100 Cyberjaya, Selangor, Malaysia

distribution using pseudorandom numbers [7]. The Pareto distribution is parameterized by the location (*k*) and shape (*a*) parameters as follows:

$$f(r_t, k, a) = \frac{(ak^a)}{r_t^{a+1}} \quad \text{and} \quad F(r_t) = 1 - \left(\frac{k}{r_t}\right)^a \tag{1}$$

Where, $0 \le k \le r_t$ and $a > 0$. The parameters derived from the first four moments are bounded by the shape parameter as below:

$$\text{Mean} = \frac{ak}{a-1} \quad \text{for } a > 1; \tag{2}$$

$$\text{Variance} = \frac{k^2 a}{(a-1)^2 (a-2)} \quad \text{for } a > 2; \tag{3}$$

$$\text{Skewness} = \frac{2(1+a)}{a-3}\sqrt{\frac{a-2}{a}} \quad \text{for } a > 3; \tag{4}$$

$$\text{Excess Kurtosis} = \frac{6(a^3 + a^2 - 6a - 2)}{a(a-3)(a-4)} \quad \text{for } a > 4 \tag{5}$$

The tail behaviour can be estimated using non-parametric method proposed by Hill [8]. Hill estimator has the advantage of simplicity (no subsamples required) over the maximum likelihood method. However, the Hill's estimator is most effective when the underlying distribution is Pareto type or approximate to Pareto. In order to verify this, we present some statistical tests (Q-Q Plots and Goodness of fit tests) for the empirical tails against the Pareto distribution. Using the quartile definition of extreme outlier, we included all the empirical observations that fallen more than 3 times the interquartile range above the upper quartile or below the lower quartile.

For upper tail, considered a series, $\{r_t\}$ with k extreme observations with the order statistics $r_{(1)}, \ldots, r_{(k)}$ where $r_{(1)}$ and $r_{(k)}$ are the minimum and maximum returns respectively. On the other hand, the lower tail can be analyzed by a simple sign change: $-r_{(1)}, \ldots, -r_{(k)}$. The log likelihood function for the above observations can be expressed as:

$$l\left(a, r_{(1)}\right) = k\ln a + ka\ln x(1) - (a+1)\sum_{i=1}^{k} \ln r_i \tag{6}$$

Where, $l(a, r_{(1)})$ is monotonically increasing with $r_{(1)}$ with the estimated $r_{(1)}$ equal to minimum $r_{(i)}$. Finally, the estimated a can be obtained by using the partial derivative approach as follow:

$$\hat{a} = \frac{k}{\sum_{i}^{k} \ln x_{(i)} - \ln x_{(l)}} \tag{7}$$

For goodness-of-fit test, we implemented the Cramer-von-Mises (W^2) and Watson (U^2) statistics for normality tests and to check the discrepancy between the estimated tail and the empirically observed tails.

EMPIRICAL RESULTS

Preliminary tail distribution statistics: For graphical illustrations, we compared the kernel density estimates (adjusted histogram) of the probability distribution for standardized returns with a simulated normal distribution in Fig. 1. Five selected empirical series evidenced the high peak property. However, the tail property is not clearly observed.

Table 1 reported the descriptive statistics, test statistics corresponding to skewness, kurtosis and autocorrelation tests for all the standardized returns series ($r_t - \bar{r}/\hat{\sigma}$). The null hypotheses of zero skewness and zero excess kurtosis are both rejected with highly significant values of test statistics. Heavy-tailed clearly indicated by the large values of kurtosis across the market indices with the maximum of 46.198 (CI) and minimum 20.666 (PRP) respectively. Overall, the first order autocorrelations indicated relatively higher values as compare further lags autocorrelation. However, the daily returns especially in emerging market might caused by the infrequent trading behaviour. According to Miller *et al*. [9], this spurious autocorrelation can be adjusted using a first order autoregressive or moving average. The autocorrelations however die out insignificantly after few lags. Both the normality tests, Jacque-Bera statistics and the Cramer-Von-Mises statistics reported non-Gaussianity in the empirical distributions.

The pareto distribution: The approximation numbers of observations are experimented for 0.5%, 1% and extreme outliers for the left and right tails. Using the

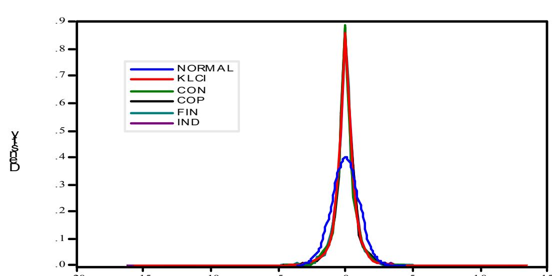

Fig. 1: Density kernel distribution

Table 1: Descriptive statistics and test statistics

Statistics	KLCI	CON	COP	FIN	IND	INP	PLN	PRP	TIN	TAS
Max	13.345	11.585	12.926	12.894	12.150	12.414	10.026	11.217	15.666	13.380
Min	-15.498	-11.034	-13.233	-11.739	-16.013	-16.203	-11.003	-10.128	-12.660	-12.621
$\tilde{S}$	0.535*	0.831*	0.200*	1.194*	-0.092*	-0.599*	-0.288*	0.572*	0.705*	0.816*
	(13.04)	(20.28)	(4.87)	(29.11)	(-2.25)	(-14.60)	(-7.03)	(13.94)	(17.20)	(19.89)
$\tilde{K}$	46.198*	27.642*	40.300*	30.795*	45.696*	41.596*	26.110*	20.666*	44.890*	33.298*
	(526.78)	(300.50)	(454.86)	(338.95)	(520.65)	(470.66)	(281.81)	(215.43)	(510.83)	(369.47)
Autocorrelation										
lag 1	0.056	0.112	0.110	0.141	0.036	0.023	0.072	0.086	0.089	0.058
lag 2	0.036	0.094	0.005	0.079	0.000	0.031	0.003	0.055	0.040	0.050
lag 3	0.030	0.006	0.006	0.069	-0.001	0.053	0.035	0.046	0.041	0.003
lag 4	-0.092	-0.023	-0.019	-0.018	-0.038	-0.021	0.021	0.031	-0.018	-0.087
lag 5	0.056	0.059	0.054	0.038	0.063	0.095	0.107	0.060	0.028	0.029
lag 6	-0.054	-0.049	-0.044	-0.055	-0.047	-0.069	-0.020	-0.052	-0.089	-0.051
lag 7	-0.018	-0.014	-0.019	-0.045	0.008	-0.025	-0.050	-0.016	-0.067	-0.029
lag 8	-0.004	-0.004	0.006	-0.004	0.003	0.018	0.034	0.021	-0.028	-0.008
Normality test										
JB	277675*	90713*	206918*	115738*	271092*	221733*	79468*	46606*	261250*	136906*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
W^2	27.680*	24.710*	25.798*	24.104*	24.562*	26.391*	25.089*	24.376*	32.272*	25.306*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

1: *t-test* for Gaussian skewness and kurtosis. The standard error for Gaussian skewness and kurtosis are $\sqrt{6/T}$ =0.041 and $\sqrt{24/T}$ =0.082. The parentheses indicate the *t*-statistics. The 5% and 1% critical values are 1.96 and 2.58 respectively. The null hypothesis indicates $\tilde{S}=0$ and $\tilde{K}=3$ respectively, 2: Cramer-von Mises(W^2) Empirical distribution test(EDT): The parentheses indicate the p-values, H_0: The return series follows a Gaussian distribution, H_1: The return series does not follow a Gaussian distribution, 3: * denotes 5% level of significance

Table 2: Hill's estimator for upper tail

	0.5 % observations		1.0 % observations			Extreme outlier (%)			
Index	a	k	a	k	> Q3+3IQR	a	k	W^2	U^2
KLCI	3.6557	2.1257	2.6886	2.2543	1.29	2.4692	2.3641	0.0746	0.0575
	-0.1007	-0.6174	-0.0339	-0.4155		-0.0231	-0.3769	-0.4910	-0.5080
CON	3.9391	2.5635	2.7859	2.2696	1.40	2.5125	2.4720	0.0622	0.0562
	-0.0892	-0.7446	-0.0349	-0.4183		-0.0206	-0.3756	-0.5770	-0.5200
COP	3.2638	2.2077	2.5402	2.5566	1.06	2.5012	2.5957	0.1587	0.0845
	-0.0864	-0.6412	-0.0282	-0.4712		-0.0258	-0.4631	-0.1250	-0.2810
FIN	4.0379	2.6238	2.9261	2.3785	1.29	2.5590	2.2317	0.0481	0.0348
	-0.0892	-0.7621	-0.0349	-0.4384		-0.0254	-0.3558	-0.7480	-0.8360
IND	3.4446	2.2910	2.5421	2.1868	0.98	2.6261	2.2838	0.1062	0.1055
	-0.0877	-0.6654	-0.0331	-0.4030		-0.0336	-0.4282	-0.2900	-0.1700
INP	3.5813	2.3740	2.7816	2.6167	1.40	2.4092	2.5524	0.0379	0.0328
	-0.0878	-0.6896	-0.0301	-0.4823		-0.0191	-0.3878	-0.8610	-0.8610
PLN	3.6785	2.9042	2.7476	2.6690	1.29	2.4753	2.5931	0.0291	0.0286
	-0.0731	-0.8436	-0.0292	-0.4919		-0.0211	-0.4134	-0.9354	-0.9060
PRP	4.1273	3.6061	3.1320	2.9818	1.76	2.4573	2.6824	0.0479	0.0362
	-0.0655	-1.0474	-0.0297	-0.5496		-0.0147	-0.3576	-0.7510	-0.8170
TIN	3.7298	2.1713	2.9097	2.5231	1.79	2.2378	2.3471	0.0563	0.0541
	-0.1005	-0.6307	-0.0327	-0.4650		-0.0150	-0.3102	-0.6450	-0.5430
TAS	3.7383	2.3894	2.7444	2.2726	1.37	2.5073	2.4957	0.0553	0.0474
	-0.0911	-0.6940	-0.0343	-0.4188		-0.0208	-0.3836	-0.6570	-0.6400

1: Extreme outliers are observations exceed Q3+3IQR where Q3 and IQR are the upper quartile and interquartile range respectively, 2: *a* and *k* represent the shape and location parameter for Pareto distribution. Value in the parenthesis denotes the standard error, 3: The goodness-of-fit tests follow the null and alternative hypotheses as follows: H_0: Both the distributions are identical; H_1: H_0 is not true, Value in the parenthesis denotes the p-value

Table 3: Hill's estimator for lower tail

	0.5 % observations		1.0 % observations			Extreme outlier (%)			
Index	a	k	a	k	< Q1-3IQR	a	k	W^2	U^2
KLCI	3.4662	3.1593	2.7410	3.1258	1.32	2.4651	3.0023	0.0286	0.0250
	-0.0631	-0.9177	-0.0247	-0.5761		-0.0177	-0.4727	-0.9390	-0.9420
CON	3.6595	3.2001	2.4613	2.2119	0.98	2.5193	2.2637	0.1541	0.1022
	-0.0658	-0.9295	-0.0317	-0.4076		-0.0326	-0.4244	-0.1340	-0.1840
COP	3.6830	2.5345	2.8165	2.5897	1.29	2.4962	2.4628	0.0344	0.0337
	-0.0844	-0.7362	-0.0308	-0.4773		-0.0224	-0.3926	-0.8910	-0.8500
FIN	3.5898	4.0449	2.7567	3.2372	1.42	2.5264	3.4003	0.0553	0.0553
	-0.0506	-1.1749	-0.0240	-0.5966		-0.0147	-0.5108	-0.6570	-0.5300
IND	3.5454	3.1604	2.6446	2.7062	1.09	2.6007	2.8015	0.0517	0.0512
	-0.0645	-0.9180	-0.0277	-0.4988		-0.0242	-0.4921	-0.7020	-0.5800
INP	3.7843	2.7875	2.6941	2.5104	1.37	2.4030	2.5661	0.0911	0.0792
	-0.0785	-0.8097	-0.0304	-0.4627		-0.0194	-0.3944	-0.3720	-0.3190
PLN	3.9887	3.2150	2.8108	2.5105	1.29	2.5225	2.4600	0.0438	0.0414
	-0.0713	-0.9338	-0.0318	-0.4627		-0.0226	-0.3922	-0.8000	-0.7390
PRP	4.1879	4.4098	2.8071	2.6446	1.15	2.5233	2.3142	0.1996	0.1262
	-0.0541	-1.2809	-0.0301	-0.4874		-0.0271	-0.3946	-0.0680	-0.1050
TIN	3.3922	2.8981	2.6429	2.8773	1.34	2.3315	2.7371	0.0494	0.0462
	-0.0676	-0.8418	-0.0260	-0.5303		-0.0180	-0.4257	-0.7320	-0.6600
TAS	3.4598	2.9168	2.7096	2.8574	1.29	2.5177	2.9657	0.0336	0.0313
	-0.0684	-0.8472	-0.0268	-0.5266		-0.0187	-0.4728	-0.8970	-0.8780

1: Extreme outliers are observations less than Q1-3IQR where Q1 and IQR are the lower quartile and interquartile range respectively, 2: *a* and *k* represent the shape and location parameter for Pareto distribution. Value in the parenthesis denotes the standard error, 3: The goodness-of-fit tests follow the null and alternative hypotheses as follows: H_0: Both the distributions are identical; H_1: H_0 is not true, Value in the parenthesis denotes the p-value

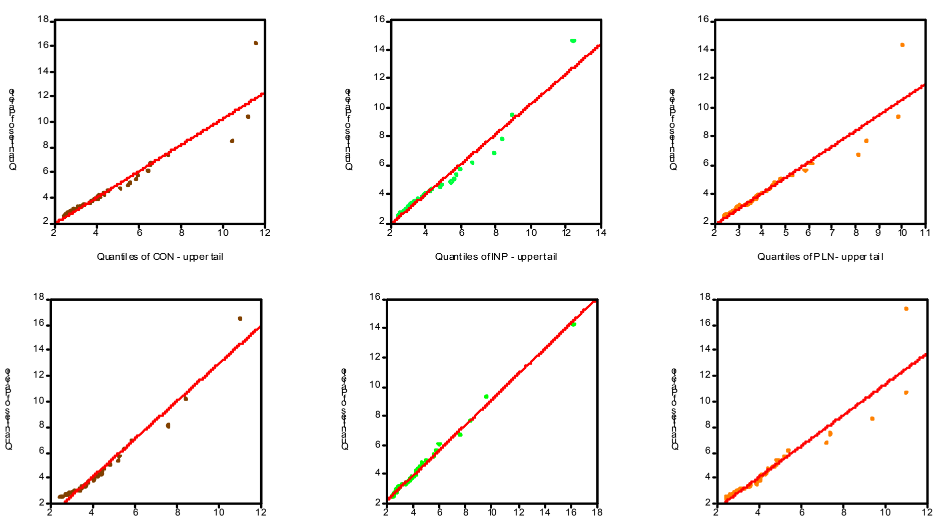

Fig. 2: Q-Q-Plots for empirical and pareto distributions

quartile definition of extreme outlier, the empirical results indicated that the percentages of extreme values are around 1.0%. Based on the extreme outliers observations, Table 2 and 3 shown that the estimated shape parameters (*a*s) are all exceeded 2 for both the upper and lower tails and indicated the presence of finite means and variances whereas moments of order higher than 2 are unbounded. According to Loretan and Phillips [10], not necessary all the moment higher than 2, such as kurtosis is finite. The positive estimated *a* implied that the tails on both tails of the innovation distributions are heavy. For thickness comparison of upper and lower tails, five indices (KLCI, COP, FIN, IND and INP) indicated slightly heavier tails at the lower tails. Whereas, the remaining five indices (CON, PLN, PRP, TIN and TAS) shown opposite results with thicker upper tails. This asymmetry property is further verified by the rejection of skewness test at 5% level in Table 1. These findings suggested that the long trading (lower tail) might encounter higher risk as compare to short trading (upper tail) investments and vice versa.

Figure 2 reported that the Q-Q-plots fitted reasonably well between the empirical and estimated Pareto distributions. In Table 2 and Table 3, the formal discrepancy tests also failed to reject the null hypothesis of no discrepancy between the two tails distributions at 5% significant level.

CONCLUSION

This paper investigates the tail behaviours of the innovation distributions for ten Malaysian stock indices. We estimated the upper and lower tails separately by the Pareto distribution of the data. The positive estimated shape parameters indicated heavy-tailed for all the indices. However, the indices evidenced a mixture of positive and negative skewness. These findings provide non-trivial information to the investors who involve in long and short financial positions.

ACKNOWLEDGEMENT

The authors would like to gratefully acknowledge the financial support from the Fundamental Research Grant Scheme (FGRS0067-2006).

REFERENCES

1. Bollerslev, T., R. Chou and K. Kroner, 1992. ARCH modeling in finance: A review of the theory and empirical evidence. J. Econ., 52: 5-59.
2. Cotter, J., 2006. Modelling catastrophic risk in international equity markets: An extreme value approach. Applied Financial Econ. Lett., 2: 13-17.
3. Mandelbrot, B.B., 1963. The variation of certain speculative prices. J. Business, 36: 392-417.
4. Pagan, A., 1996. The econometrics of financial markets. J. Emp. Finance, 3: 15-102.
5. Barndorff-Nielsen, O.E., 1997. Normal inverse Gaussian distributions and the modelling of stock returns. Scandinavian J. Stat., 24: 1-13.
6. Giot, P. and S. Laurent, 2004. Modelling daily Value-at-Risk using realized volatility and ARCH type models. J. Emp. Finan., 11: 379-398.
7. Samuel, K. and N. Saralees, 2000. Extreme value distributions: Theory and applications, Imperial College Press, London.
8. Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. Ann. Stat., 3: 1163-1173.
9. Miller, M.H., J. Muthuswamy and R.E. Whaley, 1994. Mean reversion of Standard and Poor 500 index basis changes: Arbitrage-induced or statistical illusion? J. Finan., 49: 479-513.
10. Loretan, M. and P.C.B. Phillips, 1994. Testing the covariance stationarity of heavy-tailed time series. J. Emp. Finan., 1: 211-248.