Research Journal of Applied Sciences 10 (8): 324-333, 2015
ISSN: 1815-932X

Modelling Stock Market Return via Normal Mixture Distribution

[1]Zetty Ain Kamaruzzaman and [2]Zaidi Isa
[1]Department of Statistics, [2]Department of Actuarial Science,
Faculty of Science and Technology, School of Mathematical Sciences,
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

Abstract: Previous studies proved that the distributions of stock returns exhibit fat-tail and skewness. The normal mixture distribution provides a practical extension of the normal distribution for modelling stock market returns with the above mention stylized facts. It has been successfully applied in financial time series modelling and the application is still expending not only in asset-return modelling but in other applied fields. In this study, normal mixture distribution is proposed to accommodate the non-normality and asymmetry characteristics of financial time series data as found in the distribution of returns for Bursa Malaysia index series namely the FTSE Bursa Malaysia EMAS Shariah Index (FBMS) from October 2006 until July 2012. Empirical analysis is conducted across frequencies (monthly, weekly and daily) to demonstrate the proposed method. Firstly, we present the basic definitions, concepts and distribution properties of normal mixture distribution. In support of determining the number of components in the mixture, we use the information criterion for model selection. The goodness-of-fit measures provide supporting evidence in favour of the two-component normal mixture distribution at all frequency levels. For parameter estimation, we use the most commonly used Maximum Likelihood Estimation (MLE) via the EM algorithm to fit the two-component normal mixture distribution. Also, the empirical results indicate that the normal mixture distributions offer a plausible description of the data. It also shown to be more superior compare to the use of other distributions.

Key words: Behaviour of financial time series, stock market return modelling, normal mixture distribution, model selection criteria, EM algorithm

INTRODUCTION

Previous research on financial assets returns is devoted through the approximately normally distributed assumption. The normal distribution is commonly used in the 1700's especially when it has been successfully applied to astronomical data analysis by Karl Gauss in the 1800. However, from the late 1960's, the empirical finance analyses failed to support the normality assumption on estimating the financial assets returns distribution. Mandelbrot (1963) and Fama (1965) have pointed out the existence of excess kurtosis and heavy tails in financial assets returns and established the empirical evidence on the non-normality in returns. Later, strong evidence by numerous empirical finance studies has indicated that most of the financial assets return distributions are non-normal where normality is overwhelmingly rejected in many returns distribution (Esch, 2010; Tan and Chu, 2012).

Thus, the financial assets distributions such as stock market returns are poorly represented by the normal distribution, especially for high-frequency dataset. It is a stylized facts that returns exhibits non-normality and asymmetry characteristics where it have thick tails are skewed and leptokurtic relative to the normal distribution (Eijgenhuijsen and Buckley, 1999; Cont, 2001; Behr and Potter, 1998). This happened because it have more values near the mean and in the extreme tails (Hall *et al*., 1989) and dramatic falls and spectacular jumps appear with higher frequency than predicted (Frances and van Dijk, 2000). Stylized facts as defined by Cont (2001) are statistical properties of financial time series, common across a wide range of instruments, markets and time periods.

Therefore, assumption regarding the returns distribution plays a very vital role in both financial modelling as well as in its applications. For example, one may underestimate the occurrence of extreme financial

Corresponding Author: Zetty Ain Kamaruzzaman, Department of Statistics, Faculty of Science and Technology, School of Mathematical Sciences, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

events such as market crashes if ally with a wrong distributional assumption. It is very important to find an accurate distribution that empirically fits the observed returns. The form of the returns distribution is a very crucial assumption.

One way to accommodate the above-mentioned stylized fact is to introduce a more flexible distribution model. Normal mixture distribution has gain increasing attention in various disciplines of knowledge. The earliest recorded application of normal mixture distribution was undertaken by Simon Newcomb in his study in Astronomy in 1886 followed by Karl Pearson in his classic work on Method of Moments in 1894. In empirical finance applications, the use of normal mixture distribution to handle fat tails was first considered by Newcomb (1886). Gridgeman (1970) proves that a normal mixture distribution is leptokurtic, when all regimes have the same mean. There exists a long history of modelling asset returns with normal mixture distribution (Peters, 1991; Praetz, 1972; Clark, 1973; Blattberg and Gonedes, 1974). Kon (1984) examined the daily returns from 30 different stocks in the Dow-Jones industrial average, estimated normal mixture distribution with two up to four components which were found to fit appropriately and showed that the normal mixture distribution has more descriptive validity than Student t distribution. Others are Hall *et al.* (1989), Tucker (1992) and Tran (1998). And the lists of literatures continue to emerge until at this moment.

Some attractive property of normal mixture distribution is that it is flexible to accommodate various shapes of continuous distributions and able to capture leptokurtic, skewed and multimodal characteristics of financial time series data. Also it is believed that normal mixture distribution is appropriate in order to accommodate certain discontinuities in stock returns such as the 'weekend effect', the 'turn-of-the month effect' and the 'January effect' (Klar and Meintanis, 2005).

A good introduction to the theory and applications of normal mixture distribution can be found by Everitt and Hand, Titterington, McLachlan and Basford, Lindsay and McLachlan and Peel. Meanwhile, various applications of normal mixture distribution in empirical finance are documented by Fruhwirth-Schnatter and Alexander.

Fitting mixture distributions can be handled by a wide variety of techniques such as graphical methods, method of moments, minimum-distance methods, maximum likelihood and Bayesian approaches for an exhaustive review of these methods. Considerable advances have been made in the fitting of mixture models especially via the maximum likelihood method. The maximum likelihood method has focused many attentions and by far has been the most commonly used approach to fitting the mixture distributions mainly due to the existence of an associated statistical theory and since the advent of the EM algorithm. The key property of the EM algorithm has been established by Dempster *et al.* (1977). The EM algorithm is a popular tool for simplifying maximum likelihood problems in the context of a mixture model. The EM algorithm has become the method of choice for estimating the parameters of a mixture model, since its formulation leads to straightforward estimators.

Determining the number of components, k is a major issue in mixture modelling. Two commonly employed techniques in determining the number of components, k are the information criterion and parametric bootstrapping of the likelihood ratio test statistic values (McLachlan, 1987). Majority of the estimation techniques assume that the number of components, k in the mixture is known at a priori where it is known before the estimation of parameters is attempted.

This study paves the way for an easy applied estimation of returns distribution using normal mixture distribution through its flexibility. The significance of this study lays in the accurate distributional assumption of returns for modelling purpose and its application. Also, the purpose of this study is to provide evidence on the descriptive validity of normal mixture distribution as a statistical model for stock returns.

Our motivation is that it is vital to understand the pattern and distributional properties of Malaysia stock market indices as there are tremendous growth and increasing interest of investors towards investment in Malaysia stock market. Our goal is to find a best model in describing the Malaysia stock market indices based on its time series patterns. We investigate FTSE Bursa Malaysia EMAS shariah index across frequencies, i.e., monthly, weekly and daily series. These stock market indices encompass of stock market crises such as Asian financial crisis and subprime crisis. In this study, we focus both on the statistical and financial properties of the normal mixture distribution.

STYLIZED FACTS OF MALAYSIA STOCK MARKET RETURNS

In this study, we provide stylized facts and descriptive evidence on the distribution of the FTSE Bursa Malaysia EMAS shariah index (FBMS) monthly, weekly and daily returns.

The FTSE Bursa Malaysia index series was officially launched on 26 June, 2006 (http://www.bursamalaysia.com). The data set used in this study is the closing price covers a 6 years period from October 2006 to July 2012 for Malaysia stock market index namely the FTSE Bursa

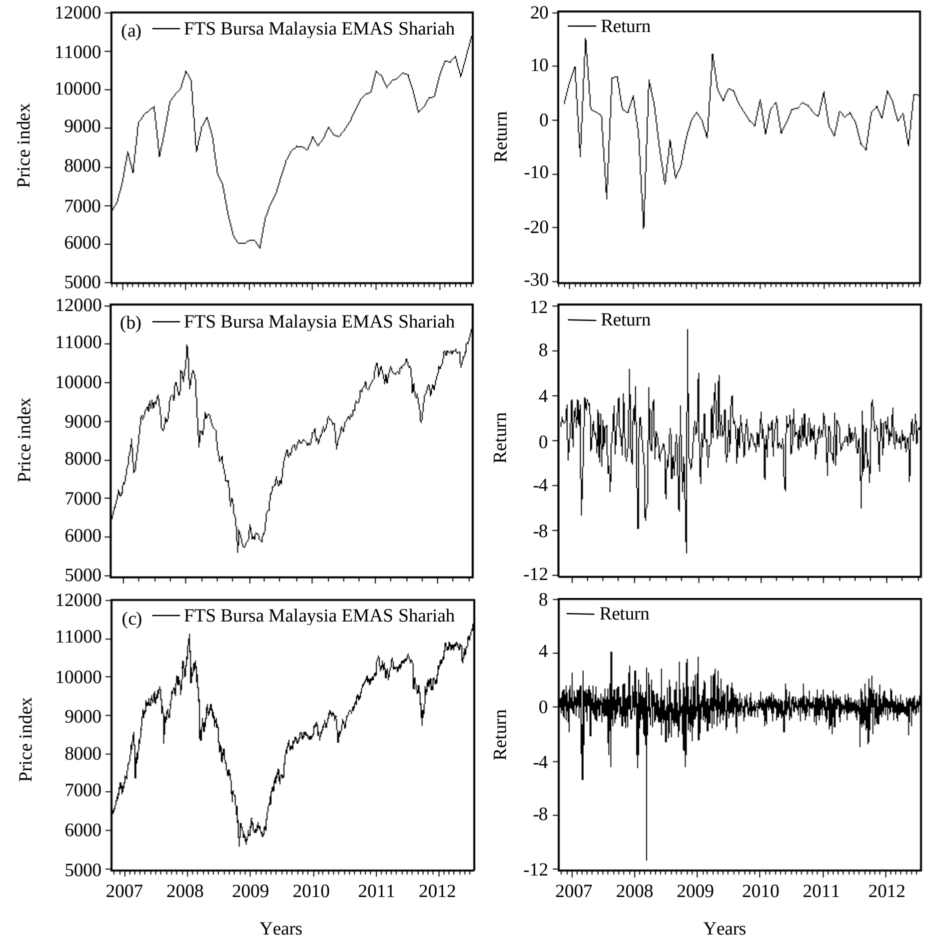

Fig. 1: Time series and returns plot of FTSE Bursa Malaysia EMAS shariah index; a) monthly series; b) weekly series and c) daily series

Malaysia EMAS shariah index (FBMS) as obtained from DataStream. It is a benchmark index based on the Main Market. This index comprises the Shariah-compliant constituents of the FTSE Bursa Malaysia EMAS index that meet the screening requirements of the Securities Commission's Shariah Advisory Council (SAC). The series is denominated in Malaysian Ringgit (MYR).

Prior to analysis, the series is analyzed in return which is the first difference of natural algorithms multiplied by 100 over the whole period. This is done to express things in percentage terms. Let P_{it} be the observed closing price of market index i on day t, i = 1, ..., n and t = 1, ..., T. The rates of return is defined as the percentage rate of return by:

$$y_{it} = 100\times\log\left(\frac{P_{it}}{P_{i,t-1}}\right) \quad (1)$$

Figure 1 depicts the time series plot of stock market index of Bursa Malaysia together with the returns of Bursa Malaysia stock market index plot for the time span 10/2006 to 7/2012. From Fig. 1, it can be seen that the price rise and fall over time. There are periods of quiet and periods of wild variation in the returns. The period analyzed can be characterized as a period of market instability as it reflects the upturn and downturn of Malaysia stock market. It can also be seen that the growth pattern in Malaysia has not been smooth at all times. Malaysia undergoes sequence of upward and downward economical episodes due to global crisis. Just to mention a few; in 1997, Malaysia experienced a reduction in economic growth due to the Asian financial crisis where the rapid growth in Asian economies had come to a halt; The September 11 attacks had a significant economic impact on world market; bird flu epidemic especially in Asia in 2003; subprime mortgage crisis in 2006; the price of petroleum spiked in 2008 as well as the rapid increase in food price on the same year and H1N1 attacks in 2009. Those can be identified as shocks in the stock market of Malaysia (Cheng, 2003).

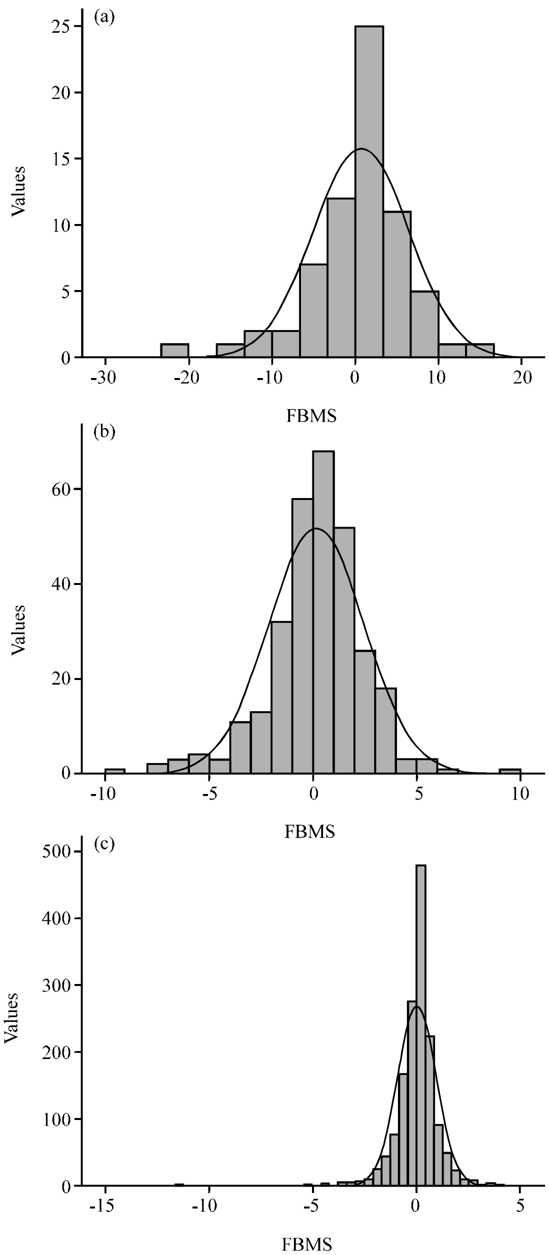

Fig. 2: Histogram with corresponding normal curve; a) monthly series; b) weekly series and c) daily series

Figure 2 depicts the histogram of the returns with the corresponding normal curve with same mean and standard deviation. From Fig. 2, it can easily be seen that the empirical distribution is higher peak and has heavier tails than the normal distribution. Also note that the return distributions with thicker tails have a thinner and higher peak in the center compared to normal distribution. The fit of the normal curve to the histogram is poor.

Table 1 reports the descriptive statistics and tests for the stock prices over the whole sample period. First, the means of the series in general not significantly different from zero. Second, there is evidence of negative skewness. Third, it has been found that stock returns in financial market have excess kurtosis. And the Jarque-Bera test rejects the null hypothesis of normality for the Malaysia stock market returns at all frequency level.

Table 1: Statistical properties of FTSE Bursa Malaysia EMAS shariah index

Statistics	Monthly	Weekly	Daily
N	68	299	1415
Mean	0.7081	0.1897	0.0401
Median	1.5734	0.3583	0.0897
Maximum	16.8037	9.8535	4.0747
Minimum	-19.6753	-9.9771	-11.3205
Std. Dev.	5.6573	2.3031	0.9556
Skewness	-0.7471	-0.5133	-1.5840
Kurtosis	5.7956	5.7382	19.6562
Jarque-Bera	26.7939	106.5375	16948.3800
Probability	0.0000	0.0000	0.0000

Thus, the returns of FTSE Bursa Malaysia EMAS shariah index (FBMS) are poorly described by normal distribution not only at high-frequency level (daily series), but also at medium-frequency (weekly series) and low-frequency level (monthly series).

DISTRIBUTIONAL FITTING

In this study, we present the distributional fitting for FTSE Bursa Malaysia EMAS shariah index (FBMS) monthly, weekly and daily returns.

Using our historical data, assuming that historical patterns hold and that history tends to repeat itself then it can be used to find the best-fitting distribution with their relevant parameters (Mun in 2010). We perform the distributional fitting (30 continuous distributions) using Risk Simulator Software (www.risksimulator.com).

The statistical ranking method used in distributional fitting routines is Kolmogorov-Smirnov test. Kolmogorov-Smirnov is a non-parametric test for the equality of continuous probability distributions and can be used to compare a sample with a reference probability distribution, making it useful for testing abnormally shaped distributions and non-normal distributions.

A hypothesis test coupled with the maximum likelihood procedure with an internal optimization routine is used to find the best-fitting parameters on each distribution tested and the results are ranked from the best fit to the worst fit (Mun in 2010). The null hypothesis is the fitted distribution is the same distribution as the population from which the sample data to be fitted come. The higher p-value, the better the distribution fits the data.

Figure 3 depicts the distributional fitting plot. Table 2 reports the distributional fitting results using

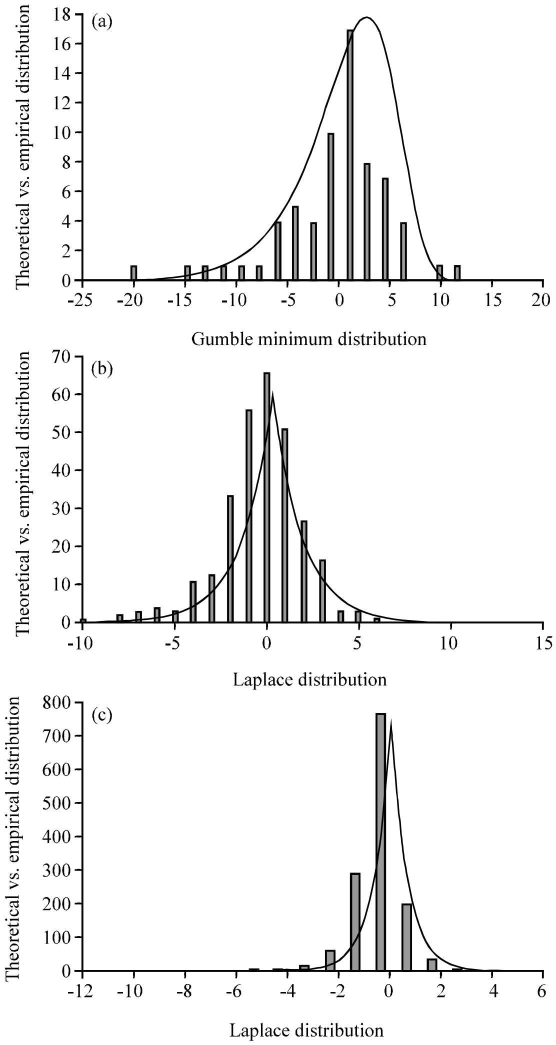

Fig. 3: Distributional fitting for FTSE Bursa Malaysia EMAS shariah index; a) monthly series; b) weekly series and c) daily series

Table 2: Distributional fitting results for FTSE Bursa Malaysia EMAS shariah index

Frequency	Distribution	KS test statistic	p-values
Monthly	Gumbel minimum	0.06	0.9656
Weekly	Laplace	0.03	0.9408
Daily	Laplace	0.02	0.3577

Table 3: Parameter fitting results for FTSE Bursa Malaysia EMAS shariah index

Frequency	Distribution	Parameters	Results
Monthly	Gumbel Minimum	Mode α	2.83
		Scale β	3.71
Weekly	Laplace	Mean α	0.33
		Scale β	1.63
Daily	Laplace	Mean α	0.07
		Scale β	0.66

Probability density function (pdf) of: Gumbel Minimum distribution: $f(x) = (1/\beta) e^{\frac{x-\alpha}{\beta}} e^{-\frac{x-\alpha}{\beta}}$ and Laplace distribution: $f(x) = (1/2\beta)\exp(-|x-\alpha|/\beta)$

Risk Simulator. The results also show the Kolmogorov-Smirnov test statistics and p-values of the best-fitting distribution of FTSE Bursa Malaysia EMAS shariah index (FBMS). From Table 2, Gumbel minimum is the best-fitting distribution for monthly return, Laplace is the best-fitting distribution for weekly and daily returns. Meanwhile, Table 3 reports the parameter fitting results.

NORMAL MIXTURE DISTRIBUTION

In this study, first, we introduce the normal mixture distribution. Next, we present the model selection criteria to determine the number of components k in the normal mixture distribution. We then present the Maximum Likelihood Estimation (MLE) via the EM algorithm to fit the mixture and study the case of k-component univariate normal mixtures.

As discussed earlier, we know that the assumption of normally distributed returns is not valid. Two features that account for non-normality in finance and economic series are; one is the presence of big shocks or outlying observations or rare events, another is abrupt regime changes over different sub-periods. The worst part of the assumption of a normal distribution is particularly with respect to the tail behaviour of the series where the tails of a normal distribution taper very rapidly, hence, the normal assumption will exclude the possibility of extreme returns where such events are frequently seen in financial markets. Moreover, according to Bidarkota, normal distribution has tails that are too thin to accommodate shocks in financial markets.

Most financial markets returns are both skewed and leptokurtic. Based on the above analysis, the FTSE Bursa Malaysia EMAS index is no exception; the monthly, weekly and daily log return is far from being normally distributed. Hence, a number of alternatives skewed and leptokurtic distributions have been applied. The normal mixture distribution is by far the most extensively applied and the simplest case is a mixture of two univariate normal distributions may be considered as the most widely applied. A flexible and tractable alternative of departures from normality is a mixture of two normal distributions. A mixture of two log normal distributions fit financial data better than a single normal distribution. Fama (1965) claims that a mixture of several normal distributions with same mean but different variances are the most popular approach to describe long-tailed distribution of price changes.

One of the most appealing features of the normal mixture distribution for modelling assets returns is that it has the flexibility to approximate various shapes of continuous distributions by adjusting its component weights, means and variances (Tan and Chu, 2012). Other advantages of using normal mixture distribution are they maintain the tractability of normal have finite higher order moments, plus can capture excess kurtosis (Tsay, 2005). Besides, normal mixture distribution can capture the structural change both in the mean and variance and it can be asymmetric. Also the normal mixture distribution are easy to interpret if the asset returns are viewed as generated from different information distributions where the mixture proportion can accommodate parameter cyclical shifts or switches among a finite number of regimes (Xu and Wirjanto, 2010). A mixture of two normal densities is defined by:

$$g(x) = \pi\varphi_1(x)+(1-\pi)\varphi_2(x),\ 0<\pi<1 \quad (2)$$

where, φ_1 and φ_2 are two normal densities with different expectations and variances. In general, the cumulative distribution function (cdf) of a mixture of K normal random variable X can be represented by:

$$F(x) = \sum_{i=1}^{K}\pi_i\Phi\left(\frac{x-\mu_i}{\sigma_i}\right) \quad (3)$$

where, Φ is the cdf of N(0, 1). Therefore, its probability density function (pdf) is:

$$f(x) = \sum_{i=1}^{K}\pi_i\phi(x;\mu_i,\sigma_i) \quad (4)$$

where, for i = 1, ..., K:

$$\phi(x;\mu_i,\sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i}e^{\frac{-(x-\mu_i)^2}{2\sigma_i^2}}$$

$$\sum_{i=1}^{K}\pi_i = 1 \text{ and } 0\leq\pi_i\leq 1$$

If X is a mixture of K normals with pdf in (Eq. 3), then its mean, variance, skewness and kurtosis are:

$$\begin{aligned}
\mu &= \sum_{i=1}^{K}\pi_i\mu_i \\
\sigma^2 &= \sum_{i=1}^{K}\pi_i\left(\sigma_i^2+\mu_i^2\right)-\mu^2 \\
\tau &= \frac{1}{\sigma^3}\sum_{i=1}^{K}\pi_i(\mu_i-\mu)\left[3\sigma_i^3+(\mu_i-\mu)^2\right] \\
\kappa &= \frac{1}{\sigma^4}\sum_{i=1}^{K}\pi_i\left[3\sigma_i^4+6(\mu_i-\mu)^2\sigma_i^2+(\mu_i-\mu)^4\right]
\end{aligned} \quad (5)$$

Alexander defines finite mixture distribution as a probability-weighted sum of other distribution functions where the density function of a mixture distribution is the same probability-weighted sum of the component density function. In normal mixture distribution, the return distribution is approximated by a mixture of normal each has unique mean μ_1 standard deviation σ_i and weight (sometimes also called as probability or mixing parameter) π_i.

The mixture setting, according to Alexander is design to capture different market regimes. The typical interpretation of a mixture of two normal distributions is that there are two regimes for returns. One where the return has mean μ_1 and variance σ_1^2 another where the return has mean μ_2 and variance σ_2^2. The weight is the probability π which the first regime occurs while the second regime occurs with probability $1-\pi$. For example, in a two-component mixture (k = 2), the first component, with a relatively high mean and small variance may be interpreted as the bull market regime, occurring with probability π_1 whereas the second regime with a lower expected return and a greater variance, represent the bear market.

Wang showed an illustration of a normal distribution model that is quite inappropriate for fitting market data since its density does not take into accounts the fat tails and skewness. Dias (2007) discuss how skewness and excess kurtosis in financial time series can be deal using finite normal mixture distribution? and illustrate why mixtures of normal models provide a flexible way of dealing with skewness and kurtosis?

Thus, the five parameters (π, μ_1, μ_2, σ_1^2, σ_2^2) of normal mixture distribution allow a very flexible definition of departures from symmetry and normality. Therefore, normal mixture distribution has the ability to deal with skewness and kurtosis in analyzing financial time series. By using normal mixture distribution, we can obtain densities with higher peaks and heavier tails than normal distribution.

MODEL SELECTION

Determining the number of components k is a major issue in mixture modelling. Two commonly employed techniques in determining the numbers of components k are the information criterion and parametric bootstrapping of the likelihood ratio test statistic values (McLachlan, 1987). Majority of the estimation techniques assume that the number of components, k, in the mixture is known at a priori where it is known before the estimation of parameters is attempted.

First, we did the calibration checking for the normal mixture distribution. Figure 4 depicts the calibration plot

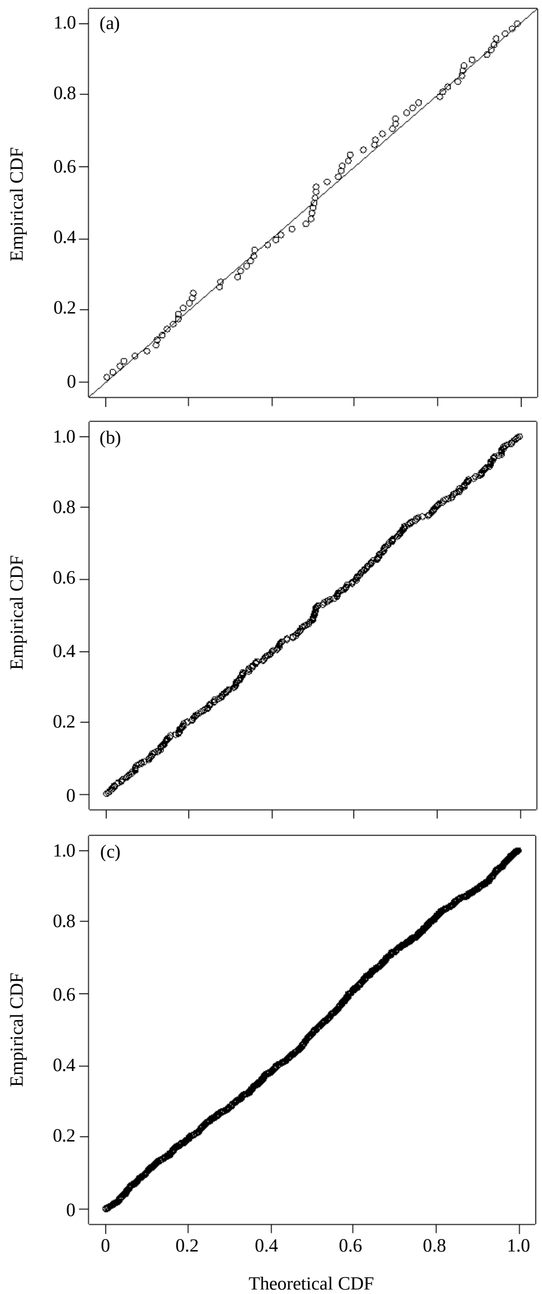

Fig. 4: Calibration plot of two-component normal mixture distribution model; a) monthly series; b) weekly series and c) daily series

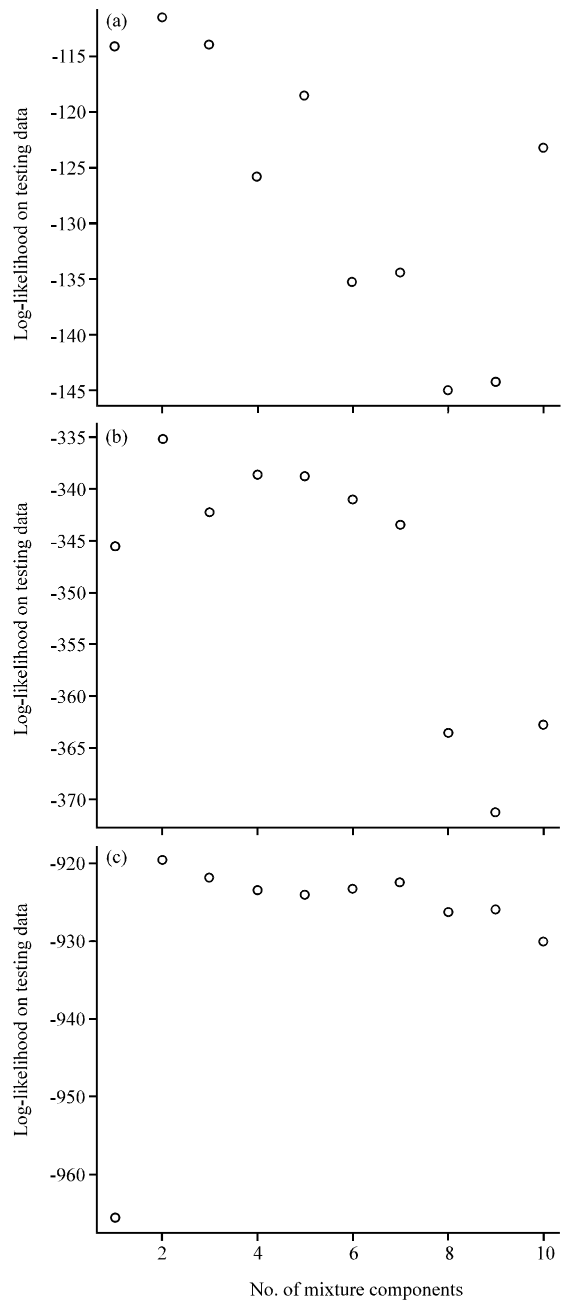

Fig. 5: Log-likelihoods of different sizes of mixture models, fit to a random half of the data for training and evaluated on the other half of the data for testing; a) monthly series; b) weekly series and c) daily series

for the two-component normal mixture distribution. Examining the two-component normal mixture distribution plot, it does look satisfactory.

Next, we do cross-validation to confirm the selection of number of the components for the mixture model. We do simple data-set splitting, where a randomly-selected half of the data is used to fit the model and half to test.

The basic idea is to split a data set into train and test. We fit the model using the training points and then calculate the log-likelihood of the test points under the model. We pick the number of component which maximizes the likelihood of the data. Table 4 and Fig. 5

Table 4: Log-likelihoods of the 10 fold cross-validation for FTSE Bursa Malaysia EMAS shariah index

k	Monthly	Weekly	Daily
1	-114.0325	-345.6042	-965.5971
2	**-111.3669**	**-335.1962**	**-917.7421**
3	-113.8726	-342.3016	-921.1646
4	-125.7174	-338.6399	-922.7361
5	-118.4525	-338.8286	-923.3435
6	-135.2373	-341.0202	-922.5892
7	-134.3250	-343.5147	-921.7902
8	-144.9777	-363.5299	-925.6190
9	-144.2491	-371.1937	-925.2779
10	-123.1648	-362.7534	-929.4505

depict the log-likelihoods of the 10 fold cross-validation. The boldface entry confirms the two components to the mixture model.

Since, a two-component normal mixture distribution seems good, we should not consider using more components as by going to three, four, etc., components, we improve the in-sample likelihood but we could expose ourselves to the danger of over-fitting. Besides, having so many parameters is not always desirable. It can lead to estimation problems and over-fitting the data can lead to specification problems.

PARAMETER ESTIMATION

Fitting the parameters of the normal mixture distribution is one of the oldest estimation problems in the statistical literature. A variety of approaches have been used to estimate the normal mixture distribution as discussed earlier. The Maximum Likelihood Method (MLE) is the most commonly preferred method for the estimation problem of normal mixture distribution. Unfortunately, the MLEs have no closed forms, hence, they have to be computed iteratively. However, the computation becomes straightforward using the Expectation-Maximization (EM) algorithm.

The EM algorithm is widely used as it is an easy and implementable method as well as a popular tool for simplifying difficult maximum likelihood problems plus has shown great performance in practice where it has the ability to deal with missing data, unobserved variables and mixture density problems. The EM algorithm will find the expected value as well as the current parameter estimates at the E step and maximizes the expectation at the M step. By repeating the E and M step, the algorithm will converge to a local maximum of the likelihood function. Various EM-type algorithms can be found in the literature.

Denote θ the parameters of the function to be optimized. The algorithm consists of iterating between two steps, the E-step and the M-step. In the Expectation (E) step, the current estimates of the parameters are used to assign responsibilities according to the relative density of the training points under each model. Next, in the Maximization (M) step, these responsibilities are used in weighted maximum-likelihood fits to update the estimates of the parameters. The E-step is repeated, updated with a new value as the current value of θ and then the M-step again provides a further updated value for θ. Thus, the algorithm proceeds, iterating between the E-step and the M-step until convergence is achieved.

Hastie *et al.* (2005) introduce a simple procedure of the EM algorithm for the special case of normal mixture distribution:

- Expectation (E) step:

$$\gamma_t = \frac{\pi\phi\left(x_t;\, \mu_2,\, \sigma_2^2\right)}{\pi\phi_1\left(x_t;\, \mu_1,\, \sigma_1^2\right)+(1-\pi)\phi_2\left(x_t;\, \mu_2,\, \sigma_2^2\right)} \text{ for } t = 1, ..., T \tag{6}$$

- Maximization (M) step:

$$\begin{aligned} \mu_1 &= \frac{\sum_{t=1}^{T}(1-\gamma_t)x_t}{\sum_{t=1}^{T}(1-\gamma_t)},\ \sigma_1^2 = \frac{\sum_{t=1}^{T}(1-\gamma_t)(x_t-\mu_1)^2}{\sum_{t=1}^{T}(1-\gamma_t)}, \\ \mu_2 &= \frac{\sum_{t=1}^{T}\gamma_t x_t}{\sum_{t=1}^{T}\gamma_t},\ \sigma_2^2 = \frac{\sum_{t=1}^{T}\gamma_t(x_t-\mu_2)^2}{\sum_{t=1}^{T}\gamma_t}, \\ \pi &= \sum_{t=1}^{T}\frac{\gamma_t}{T} \end{aligned} \tag{7}$$

As an illustration, we apply the maximum likelihood method via EM algorithm to fit the normal mixture distribution with two components to the Bursa Malaysia return. Table 5 depicts the summary of two components normal mixture using the EM algorithm. There are two components with two weights (π), two means (μ), two standard deviations (σ) and the overall log-likelihood (logL).

Several important observations may be drawn from Table 5. First, in general the low-variance component has a higher probability. The second component has a lower variance. The high-variance component has the smaller probability for all series. This indicates that the first normal is a low mean high variance regime and the second normal is a high mean low variance regime. Meanwhile, the weights indicate that the second regime is the more prevalent regime for the FTSE Bursa Malaysia EMAS shariah index.

Normal mixture distribution has an intuitive interpretation when markets display regime-specific behaviour. Markets are stable when the expected return

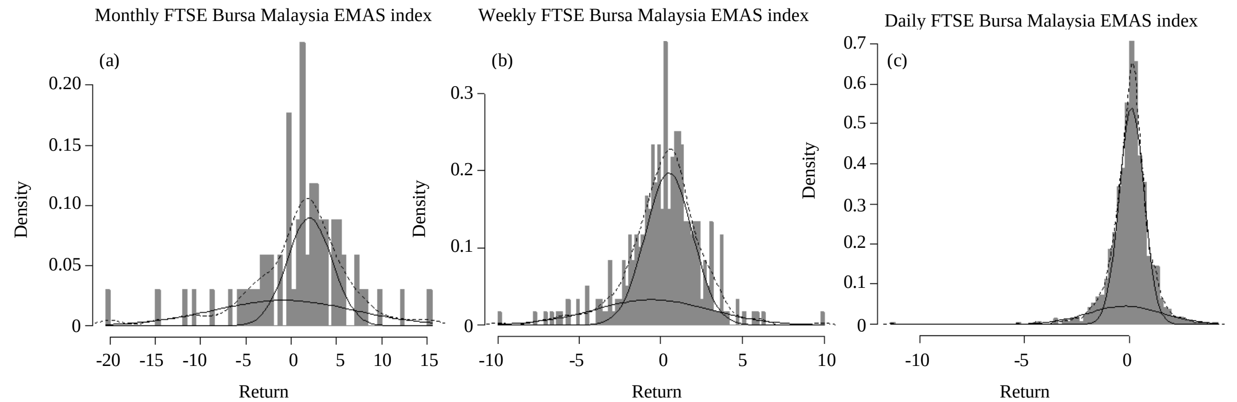

Fig. 6: Histogram, kernel density estimate (dashed line) plus two-component normal mixture distribution model (scaled by the mixing weights); a) monthly series; b) daily series and c) daily series

Table 5: Summary of five parameters two-component normal mixture distributions model using EM algorithm for FTSE Bursa Malaysia EMAS Shariah index

Frequency	MLE	Statistics	π_1	π_2	μ_1	μ_2	σ_1^2	σ_2^2
Monthly	Estimate log-l:	Estimate	0.3959	0.6041	-1.0019	1.8871	65.2078	7.3139
	-207.2483	Std. Err.	10.0445		0.6036	1.7394	0.0527	0.3635
		Ratio	0.0394		-1.6599	1.0849	1237.3397	20.1208
		95%LCL	-0.0684		-5.0319	0.5395	6.0277	-3.5997
		95%UCL	0.8602		3.0281	3.2348	124.3879	18.2275
Weekly	Estimate log-l:	Estimate	0.2602	0.7398	-0.6430	0.4826	13.1709	2.1837
	-651.9003	Std. Err.	21.2023		2.1809	8.1777	0.4326	3.4005
		Ratio	0.0123		-0.2948	0.0590	30.4459	0.6422
		95%LCL	0.0809		-1.7430	0.2218	6.5656	1.3302
		95%UCL	0.4395		0.4569	0.7433	19.7763	3.0372
Daily	Estimate log-l:	Estimate	0.1974	0.8026	-0.2021	0.0997	3.1380	0.1974
	-1775.0290	Std. Err.	55.9009		8.7504	48.2822	3.4368	47.8478
		Ratio	0.0035		-0.0231	0.0021	0.9131	0.0041
		95%LCL	0.1261		-0.4450	0.0573	2.2435	0.2793
		95%UCL	0.2687		0.0407	0.1420	4.0325	0.4149

is relatively small and positive and the volatility is relatively low but market crash as the expected return is relatively large and negative and the returns volatility is larger than when markets are stable.

Figure 6 depicts the plot of two-component normal mixture distribution. We plot the histogram of the data and the non-parametric density estimate. Then, we add the density of a given component to the current plot but scaled by the share it has in the mixture, so that it is visually comparable to the overall density.

MODEL COMPARISON

In this study, we present model comparison between the best-fitting models versus normal mixture distribution model.

Using the same arguments with distributional fitting in study, we employ the Kolmogorov-Smirnov test for model comparison. We compare their test statistics and p-values in order to choose the best model. We choose model with higher p-value.

Table 6: Model comparisons for FTSE Bursa Malaysia EMAS shariah index

Parameters	Monthly	Weekly	Daily
Distribution	Gumbel min	Laplace	Laplace
KS test Stat.	0.0600	0.0300	0.0200
p-value	0.9656	0.9408	0.3577
Distribution	Normal mixture distribution		
KS test Stat.	0.0556	0.0281	0.0235
p-value	0.9770	0.9724	0.4165

Table 6 reports model comparison between the best-fitting models with normal mixture distribution model. The results show that normal mixture distribution model is more superior compared to the best-fitting models at all frequencies for FTSE Bursa Malaysia EMAS shariah index.

CONCLUSION

In this study, we present the empirical evidence on the stock market return of FTSE Bursa Malaysia EMAS shariah index based on its time series patterns. Our goal is to find the best model in describing the series. First, we perform distributional fitting using risk simulator where Gumbel minimum distribution is the best-fitting model for monthly return Laplace distribution is the best-fitting model for weekly and daily returns. We also propose normal mixture distribution which is a flexible family of distribution to accommodate the non-normality and asymmetric characteristics of financial time series data as found in the distribution of monthly, weekly and daily rates of returns for the FTSE Bursa Malaysia EMAS shariah index from October 2006 until July 2012. We define the normal mixture distribution and explore some of its distribution properties. In support of determining number of components in normal mixture distribution, we use two selection criterions where the measures provide supporting evidence in favour of two-component. Then, we fit the two component normal mixture distribution to data sets using the maximum likelihood estimation via EM algorithm. Lastly, for model comparison, we found that Normal Mixture Distribution Model is more superior compared to the best-fitting model (Gumbel minimum distribution and Laplace distribution) for monthly, weekly and daily series. We may conclude from the above analysis that using the two-component normal mixture distribution can fit Malaysia stock market returns well and can captures the stylized facts of non-normality and asymmetry characteristics.

ACKNOWLEDGEMENT

Researchers want to thank anonymous referees whose comments, suggestions and corrections led to a considerably improvement of this study.

REFERENCES

Behr, A. and U. Potter, 1998. Alternatives to the normal model of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models. Ann. Finance, 5: 49-68.

Blattberg, R. and N. Gonedes, 1974. A comparison of the stable and student distributions as statistical models for stock prices. J. Bus., 47: 244-280.

Cheng, M.Y., 2003. Economic fluctuations and growth: An empirical study of the Malaysian economy. J. Bus. Developing Nations, 7: 51-73.

Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative prices. Econometrica, 41: 135-155.

Cont, R., 2001. Empirical properties of asset returns: Stylized facts and statistical issues. Q. Finance, 1: 223-236.

Dempster, A.P., N.M. Laird and D.B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (Methodol.), 39: 1-38.

Eijgenhuijsen, H. and A. Buckley, 1999. An overview of returns in Europe. Eur. J. Finance, 5: 276-297.

Esch, D.N., 2010. Non-normality facts and fallacies. J. Invest. Manage., 8: 49-61.

Fama, E.F., 1965. The behavior of stock-market prices. J. Bus., 38: 34-105.

Gridgeman, N.T., 1970. A comparison of two methods of analysis of mixtures of nomnal distributions. Technometrics, 12: 823-833.

Hall, J.A., B.W. Brorsen and S.H. Irwin, 1989. The distribution of futures prices: A test of the stable paretian and mixture of normals hypotheses. J. Financial Quant. Anal., 24: 105-116.

Hastie, T., R. Tibshirani, J. Friedman and J. Franklin, 2005. The elements of statistical learning: Data mining, inference and prediction. Math. Intelligencer, 27: 83-85.

Klar, B. and S.G. Meintanis, 2005. Tests for normal mixtures based on the empirical characteristic function. Comput. Stat. Data Anal., 49: 227-242.

Kon, S.J., 1984. Models of stock returns: A comparison. J. Finance, 39: 147-165.

Mandelbrot, B.B., 1963. The variation of certain speculative prices. J. Bus., 36: 394-419.

McLachlan, G.J., 1987. On bootstrapping the likelihood ratio test stastistic for the number of components in a normal mixture. Appl. Stat., 36: 318-324.

Newcomb, S., 1886. A generalized theory of the combination of observations so as to obtain the best result.A generalized theory of the combination of observations so as to obtain the best result. J. Mathe. Am., 8: 343-366.

Peters, E., 1991. A compound events model for security prices. J. Bus., 40: 317-335.

Praetz, P., 1972. The distribution of share price changes. J. Business, 45: 49-55.

Tan, K. and M. Chu, 2012. Estimation of portfolio return and value at risk using a class of gaussian mixture distributions. Int. J. Bus. Finance Res., 6: 97-107.

Tran, K.C., 1998. Estimating mixtures of normal distributions via empirical characteristic function. Econo. Rev., 17: 167-183.

Tsay, R.S., 2005. Analysis of Financial Time Series. 2nd Edn., Wiley-Interscience, New York.

Tucker, A.L., 1992. A reexamination of finite and infinite-variance distributions as models of daily stock returns. J. Bus. Econ. Stat., 10: 73-81.

Xu, D. and T.S. Wirjanto, 2010. An empirical characteristic function approach to VaR under a mixture of normal distribution with time-varying volatility. J. Derivatives, 18: 39-58.