Expert Systems With Applications 54 (2016) 170–192

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Empirical distributions of daily equity index returns: A comparison

Canan G. Corlu[a], Melike Meterelliyoz[b,*], Murat Tiniç[c]

[a] Metropolitan College, Boston University, Boston, MA 02215 USA
[b] Department of Business Administration, TOBB University of Economics and Technology, Ankara 06560 Turkey
[c] Department of Management, Bilkent University, Ankara 06800 Turkey

ARTICLE INFO

Keywords:
Index returns
Generalized lambda
Johnson translation system
Skewed-*t*
Normal inverse Gaussian
g-and-h

ABSTRACT

The normality assumption concerning the distribution of equity returns has long been challenged both empirically and theoretically. Alternative distributions have been proposed to better capture the characteristics of equity return data. This paper investigates the ability of five alternative distributions to represent the behavior of daily equity index returns over the period 1979–2014: the skewed Student-*t* distribution, the generalized lambda distribution, the Johnson system of distributions, the normal inverse Gaussian distribution, and the g-and-h distribution. We find that the generalized lambda distribution is a prominent alternative for modeling the behavior of daily equity index returns.

1. Introduction

The assumption that stock price changes follow a stable distribution forms the basis for major asset pricing and option pricing models. Early models by Bachelier (1900) take normality as a fundamental assumption for modeling stock price movements. In line with this assumption, Osborne (1959) shows that logarithms of the changes in the stock prices are mutually independent with a common probability distribution (i.e., they conform to a random walk). He then suggests that stock price changes must follow a normal distribution. However, these findings have been challenged both theoretically and empirically.[1]

An early work by Mandelbrot (1967) proposes that stock price returns belong to the family of stable Paretian distributions because they have fatter tails. Fama (1963; 1965) provides empirical evidence that supports this claim and demonstrates that stock price changes indeed have fatter tails and have higher peaks than the normal distribution. More recently, Rachev, Stoyanov, Biglova, and Fabozzi (2005) compared the stable Paretian distribution to the normal distribution using 382 US stock returns over the period 1992–2003. The authors investigated the daily returns using two probability models: the homoskedastic independent and identically distributed model and the conditional heteroskedastic ARMA-GARCH model. Normality was rejected for both models. However, Officer (1972) found that normality holds for monthly returns and that the standard deviation of the returns is inconsistent with the stable hypothesis. To support this argument, Praetz (1972) then suggested the Student-*t* distribution as an alternative to the stable Paretian because the stable Paretian distribution has an infinite variance property and the density function of the stable Paretian is unknown. Over an eight-year period, Praetz (1972) examined weekly data from Sydney Stock Exchange and showed that the Student-*t* distribution can be used as an alternative to explain the stock price behavior.

The Student-*t* distribution was also compared with the normal distribution and the Cauchy distribution by Blattberg and Gonedes (1974). Contrary to Praetz (1972), they used both daily and weekly returns of stocks of the Dow Jones Industrial (DJI), and they used the maximum likelihood estimation method for estimating the parameters of the distributions. Blattberg and Gonedes (1974) showed that the Student-*t* distribution performs better than the normal distribution on daily returns. However, normality is not rejected for monthly return data. Hagerman (1978) tested the normality hypothesis on both individual stocks of the American and New York Stock Exchanges on portfolios that contain these stocks, and found that they do not behave in line with the normal distribution. Hagerman (1978) proposed that the mixture of normal distributions and the Student-*t* distribution can be an alternative to representing the characteristics of stock return data. However, the performance of these two distributions against each other was not investigated in Hagerman's work.

Kon (1984) compared the discrete mixture of normal distributions and the Student-*t* distribution over a period of almost 19 years, examining daily returns of 30 stocks from DJI and Standard & Poor's (S&P) value- and equal-weighted stock market indexes. A

* Corresponding author. Tel.: +90 312 2924214.
E-mail addresses: canan@bu.edu (C.G. Corlu), mkuyzu@etu.edu.tr (M. Meterelliyoz), tinic@bilkent.edu.tr (M. Tiniç).

[1] A similar line of argument holds for exchange rate returns.

http://dx.doi.org/10.1016/j.eswa.2015.12.048
0957-4174/

discrete mixture of normal distributions is shown to have greater validity than the Student-*t* distribution in modeling the data. Similar to Blattberg and Gonedes (1974) and Akgiray and Booth (1987); Officer (1972) stated that the monthly returns of stock prices can be assumed to be normally distributed. However, for the daily data they found that the mixed diffusion process and the mixture of normal distributions perform better than the stable distributions.

Bookstaber and McDonald (1987) proposed the generalized beta (GB2) distribution to explain the behavior of stock returns. This was chosen because the GB2 is a flexible distribution and acknowledges various distributions as special cases. They found that the GB2 distribution is significantly better than the lognormal distribution, especially in relation to short time intervals. Badrinath and Chatterjee (1988) examined the Center for Research in Security Prices (CRSP) value-weighted market index returns between 1962 and 1985 and concluded that returns of stock prices follow a skewed g-and-h distribution.[2] Similarly, Mills (1995) found that the g-and-h distribution accurately fits a dataset that consists of three London Stock Exchange indices: FTSE 100, Mid 250, and FTSE 350.

A more general comparison of distributions with finite variances over equity stocks was conducted by Gray and French (1990). They compared the scaled-*t* distribution, the logistic distribution, the exponential power distribution, and the normal distribution over the log-returns of daily S&P 500 Composite index values for the period 1979–1985. Among four alternatives, the exponential power distribution was found to be the best fit. Lau, Lau, and Wingender (1990) showed that series of returns of stock prices that are taken from the CRSP yield higher kurtosis and skewness than the normal distribution. They proposed the lognormal, beta, Weibull, Pearson Types IV and VI, and Johnson system of distributions as alternatives. A general comparison of the normal distribution to the scaled-*t* distribution and to the mixture of two normal distributions was conducted by Aparicio and Estrada (2001) using the daily returns of 13 different European stock markets. It was found that the scaled-*t* distribution is a significantly better fit for the data, and the partial mixture of two normal distributions also performs well. Normality is rejected in all cases.

Linden (2001) introduced the Laplace mixture distribution, which is derived by conditioning the standard deviation of the normal distribution as an exponentially distributed random variable. Linden (2001) used this distribution to represent the daily, weekly, and monthly returns of the 20 most traded shares and the index of the Helsinki Stock Market. The normality assumption is not always rejected for the weekly and monthly returns. However, for the daily returns, an asymmetric Laplace distribution is found to be a better candidate than the normal distribution.

Harris and Küçüközmen (2001a) and Harris and Küçüközmen (2001b), respectively, examined the skewed generalized-t distribution (SGT) and the exponential generalized beta distribution (EGB) using daily UK, US, and Turkish equity returns. Consequently, they found that the SGT outperforms the EGB. In both studies, the authors rejected the hypothesis that the daily returns are distributed with the Student-*t*, power of exponential, or logistic distribution. In addition, for the daily Turkish returns, the Laplace distribution was also rejected. For the UK returns, the skewed-*t* distribution was preferred, whereas for the US returns, the generalized-t distribution was preferred. More recently, Behr and Pötter (2009) compared the generalized hyperbolic distribution, the generalized logF distribution, and the finite mixture of Gaussians on monthly S&P 500 index returns over the years 1871–2005 and daily returns over the years 2001–2005. For the monthly returns, the two-component Gaussian mixture distribution described the empirical distribution of the returns better than alternative distributions. Although the generalized hyperbolic distribution is the poorest performer for monthly returns, it performs best for daily data. However, as the daily data examined by Behr and Pötter (2009) is almost symmetric, the Laplace distribution, which does not have a parameter to capture the asymmetries, fits as well as the generalized hyperbolic distribution.

Finally, as an alternative to the stable distribution and the Student-*t* distribution, Chalabi, Scott, and Würtz (2010) use the generalized lambda distribution (GLD) for modeling equity returns. Starting with Eberlein and Keller (1995), the normal inverse Gaussian (NIG) distribution is used to model financial returns and particularly for modeling 30 stocks at the German Stock Index. Prause, Zentrum, and Modellbildung (1997) show the applicability of the NIG distribution in modeling German stock and US Stock Index data. Bølviken and Benth (2000) used the NIG distribution to model 8 Norwegian stocks.

In Table 1, we summarize the papers that performed comparison studies to investigate the behavior of stock returns. We find that the outcomes differ and are often conflicting. Based upon this, our goal in this study is to fill this gap in the literature by addressing which distribution is best for modeling daily equity index return data. To this end, we consider the following flexible distributions that are commonly used in finance: the skewed Student-*t* distribution, the GLD, the NIG distribution, the Johnson system of distributions, and the g-and-h distribution. We conduct a comprehensive numerical analysis to compare the overall suitability of these five distributions on the equity index returns of twenty different countries over the period 1979–2014, which is divided into twelve three-year sub-periods. We also include the normal distribution in our experimental design. The overall suitability is initially compared using the Kolmogorov–Smirnov (KS) test statistic (Chakravarti & Laha, 1967) and the Anderson–Darling (AD) test statistic (Anderson & Darling, 1954). Furthermore, we conduct *p*-value tests in order to assess the significance of these KS and AD statistics. In addition, the explanatory power of the models is tested using in-sample Value-at-Risk (VaR) failure rates. Consistent with other studies in the previous research, we find that normality is rejected in all sub-periods for all markets. Our *p*-value tests and the in-sample VaR test suggest that GLD performs best for all markets over all time periods.

The remainder of the paper is organized as follows. Section 2 presents the data. Section 3 presents the distributions along with the fitting methods that are used to estimate the parameters of the distributions. Section 4 discusses our numerical study and Section 5 presents key conclusions.

2. Description of the data

We create a diversified sample from ten developed and ten emerging market indexes. The selected developed stock market indexes are: S&P/ASX 200 Index (Australia), S&P/Toronto Stock Exchange Index (Canada), CAC 40 (France), DAX (Germany), NIKKEI 225 (Japan), the Straits Times Index (Singapore), IBEX 35 (Spain), SMI (Switzerland), FTSE 100 (UK), and S&P 500 (US), while the emerging stock market indexes are the Ibovespa Index (Brazil), IPSA Index (Chile), SHSZ 300 (China), BSE 500 (India), KOSPI Index (Korea), FBMKLCI Index (Malaysia), the Mexican IPC Index (Mexico), MICEX Index (Russia), JALSH Index (South Africa), and BIST 100 (Turkey). The daily closing index levels from January 1979 to August 2014 are collected using the Bloomberg Terminal.

Bloomberg provides index levels for the S&P/ASX 200, CAC 40, DAX and IBEX 35 prior to their establishment date. This can happen due to two reasons. First, the index levels can be adjusted with respect to their ancestor indices. For instance, the DAX

[2] Badrinath and Chatterjee (1988) also provide an excellent review of the literature.