

## **RISK MEASUREMENT PERFORMANCE OF ALTERNATIVE DISTRIBUTION FUNCTIONS**

Turan G. Bali  
Panayiotis Theodossiou

### **ABSTRACT**

This paper evaluates the performance of three extreme value distributions, i.e., generalized Pareto distribution (GPD), generalized extreme value distribution (GEV), and Box-Cox-GEV, and four skewed fat-tailed distributions, i.e., skewed generalized error distribution (SGED), skewed generalized  $t$  (SGT), exponential generalized beta of the second kind (EGB2), and inverse hyperbolic sign (IHS) in estimating conditional and unconditional value at risk (VaR) thresholds. The results provide strong evidence that the SGT, EGB2, and IHS distributions perform as well as the more specialized extreme value distributions in modeling the tail behavior of portfolio returns. All three distributions produce similar VaR thresholds and perform better than the SGED and the normal distribution in approximating the extreme tails of the return distribution. The conditional coverage and the out-of-sample performance tests show that the actual VaR thresholds are time varying to a degree not captured by unconditional VaR measures. In light of the fact that VaR type measures are employed in many different types of financial and insurance applications including the determination of capital requirements, capital reserves, the setting of insurance deductibles, the setting of reinsurance cedance

---

Turan G. Bali is the David Krell Chair Professor of Finance at the Department of Economics & Finance, Baruch College, Zicklin School of Business, City University of New York, One Bernard Baruch Way, New York, NY 10010, and the Professor of Finance at the College of Administrative Sciences and Economics at Koc University, Istanbul, Turkey, 80910. Panayiotis Theodossiou is the Professor of Finance at School of Business, Rutgers University, 227 Penn Street, Camden, NJ 08102. The authors can be contacted via e-mail: [Turan.Bali@baruch.cuny.edu](mailto:Turan.Bali@baruch.cuny.edu) and [theodoss@camden.rutgers.edu](mailto:theodoss@camden.rutgers.edu). We thank Linda Allen, Peter Christoffersen, George Constantinides, Armen Hovakimian, Jim McDonald, John Merrick, Salih Neftci, Robert Schwartz, Lenos Trigeorgis, Jack Worral, Liuren Wu, and Stavros Zenios for their extremely helpful comments and suggestions. An earlier version of this article was presented at the Aristotle's University of Thessaloniki, Baruch College and the Graduate School and University Center of the City University of New York, and the University of Cyprus. The authors gratefully acknowledge the financial support from the PSC-CUNY research foundation of the City University of New York, the Research Council fund of Rutgers University and the Whitcomb Center for Research in Financial Services for providing research support through use of the WRDS system.

levels, as well as the estimation of expected claims and expected losses, these results are important to financial managers, actuaries, and insurance practitioners.

## INTRODUCTION

Since the early 1990s, there has been an increased focus on the measurement of risk and the determination of capital requirements for financial institutions, such as banks, insurance companies, and funds, to meet catastrophic market risks. This increased focus has led to the development of various risk measures based on quantiles and loss functions, such as value at risk (VaR), which gives the maximum expected loss on a portfolio of assets over a certain holding period at a given confidence level (probability), expected shortfall (ES or tail conditional VaR) and tail risk. The VaR and ES measures have been among the most popular measures used in the financial risk management area.<sup>1</sup>

Actuaries and insurance practitioners have shown considerable interest in the measurement of financial risk for a long period of time. The aforementioned risk measures are employed in many different types of insurance applications, including the determination of capital reserves, the setting of insurance deductibles, the setting of reinsurance cedance levels, as well as the estimation of expected claims and expected losses. Interest in these measures has been triggered to a large extent from the wide spread use of the VaR-type of measures in the financial risk management area in conjunction with the increasing securitization of insurance-related risks and the use of risk measures in regulatory capital and solvency requirements of financial companies including insurance companies.<sup>2</sup>

Earlier VaR measures are based on the assumption that returns of financial assets follow a multivariate normal distribution. The multivariate normal VaR approach is easy to implement because the VaR for a portfolio of financial assets can be computed from a simple quadratic formula based on the variances and covariances of returns of individual assets. The major drawback of this approach arises from the fact that returns for individual or set of assets exhibit skewness and significant excess kurtosis (fat-tails and peakness). Consequently, the size of actual losses is much higher than that predicted by the variance–covariance VaR thresholds (tail bias).

In light of the above, an alternative approach that approximates the tails of the return distribution asymptotically is more appropriate than imposing a symmetric thin-tailed

---

<sup>1</sup> For example, a 1 percent daily VaR threshold is an estimate of the decline in the portfolio value that could occur at the 1 percent confidence level in each trading day. The latter VaR also implies that the probability of a portfolio loss less than the VaR threshold is 1 percent. On the other hand, the conditional 1 percent VaR or expected shortfall measure is simply the conditional expected value of the loss exceeding the 1 percent VaR threshold (continuous case).

<sup>2</sup> Santomero and Babbel (1997) provide reasons why firms manage risk and why managers are concerned about both expected profit and the distribution of firm returns. Sprecher and Pertl (1983) show that large losses can have very negative impact on shareholders' wealth in an efficient stock market; also see Cummins (1976) and Cho (1988) for the impact of risk management decisions of a firm on its value. Dowd and Blake (2006) provide an extensive discussion on the use of VaR measures in the insurance sector.

functional form like the normal distribution. Although VaR measures based on the normal distribution may provide acceptable estimates of the maximum likely loss under normal market conditions, they fail to account for extremely volatile periods corresponding to financial crises. Longin (2000), McNeil and Frey (2000), and Bali (2003a) show that VaR measures based on the distribution of extreme returns, instead of the distribution of all returns, provide good predictions of catastrophic market risks during extraordinary periods.

As indicated by a number of excellent papers that have developed models of insurance claim severity distributions, the estimated frequency, severity, and total claims distributions are useful in numerous practical applications.<sup>3</sup> Cummins et al. (1990), Cummins, Lewis, and Phillips (1999), and Cummins, McDonald, and Merrill (2004) show that the selection of distributional forms can have a significant impact on estimates of insurance premiums, tail quintiles, and other important statistics.

An important contribution of this article's assessment of VaR type measures is based on several flexible probability distribution functions. These distributions are the inverse hyperbolic sine (IHS) of Johnson (1949), the exponential generalized beta of the second kind (EGB2) of McDonald and Xu (1995),<sup>4</sup> the skewed generalized-*t* (SGT) of Theodossiou (1998), and the SGED of Theodossiou (2004). The latter three distributions are quite new in the VaR literature.

This article evaluates the relative VaR performance of the aforementioned distributions using the unconditional and conditional coverage tests of Kupiec (1995) and Christoffersen (1998). VaR measures based on the unconditional extreme value and skewed fat-tailed distributions do not account for systematic time-varying changes in the distribution of returns. The conditional coverage test and out-of-sample performance results indicate that the actual thresholds are time varying to a degree not captured by the unconditional density functions. In this respect, using the unconditional mean and volatility of risk factors can be a major drawback in VaR measures.<sup>5</sup> This article extends the VaR methodology by taking into account the dynamic behavior of the conditional mean and volatility of financial returns. The article then evaluates the in-sample and out-of-sample performance of conditional distribution functions in calculating VaRs.

The results indicate that the conditional extreme value and skewed fat-tailed distributions perform much better than the conditional normal density in capturing both

<sup>3</sup> Hogg and Klugman (1984) and Klugman, Panjer, and Willmot (1998) discuss a wide range of alternative models for loss distributions. Paulson and Faris (1985) applied the stable family of distributions, and Aiuppa (1988) considered the Pearson family as models for insurance losses. Ramlau-Hansen (1988) modeled fire, windstorm, and glass claims using the log-gamma and log-normal distributions. Cummins et al. (1990) considered the four-parameter generalized beta of the second kind (GB2) distribution as a model for insured fire losses, and Cummins, Lewis, and Phillips (1999) use the log-normal, Burr 12, and GB2 distributions to model the severity of insured hurricane and earthquake losses.

<sup>4</sup> GB2 (logarithmic transformation of EGB2) is used by Cummins et al. (1990) and Cummins, McDonald, and Merrill (2004) to model insurance losses.

<sup>5</sup> Note that unconditional VaR models can be appropriate in long-run contexts.

the rate of occurrence and the extent of extreme events in financial markets. The SGT, EGB2, and IHS (henceforth, referred to as the skewed fat-tailed distributions) produce precise VaR measures and compare favorably to the extreme value distributions. The latter is attributed to the fact that the latter distributions provide an excellent fit to the tails of the empirical return distribution. Note that these findings have significant implications for actuaries and the insurance sector. In addition to VaR measurement, the skewed fat-tailed and extreme value distributions can be used to investigate the effects of model selection on the tails of the severity distributions, excess of loss reinsurance premiums, and total claims distributions.

This article is organized as follows. The first section contains a short discussion of the extreme value and flexible distributions. The second section provides the unconditional VaR models based on alternative distribution functions. The third section describes the data. The fourth section presents the estimation results. The fifth section compares the risk measurement performance of alternative VaR models. The sixth section proposes a conditional VaR approach. The final section concludes.

### Extreme Value Distributions

We investigate the fluctuations of the sample maxima (minima) of a sequence of random variables  $\{X_1, X_2, \dots, X_n\}$  with common cumulative distribution function (cdf)  $F(x)$ . The three standard extreme value distributions are defined as:

$$\text{Frechet: } H_{\max, \xi}(x) = \begin{cases} \exp(-x^{-1/\xi}) & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$\text{Weibull: } H_{\max, \xi}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ \exp(-(-x)^{-1/\xi}) & \text{otherwise} \end{cases} \quad (2)$$

$$\text{Gumbel: } H_{\max, 0}(x) = \exp[-\exp(-x)] \quad -\infty < x < +\infty \quad (3)$$

Jenkinson (1955) proposes a generalized extreme value (GEV) distribution, which includes the three limit distributions in (1) to (3), distinguished by Gnedenko (1943):

$$H_{\max, \xi}(M; \mu, \sigma) = \exp \left\{ - \left[ 1 + \xi \left( \frac{M - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}; \quad 1 + \xi \left( \frac{M - \mu}{\sigma} \right) \geq 0, \quad (4)$$

where  $\xi$  is a shape parameter, which is also called the tail index that reflects the fatness of the distribution (i.e., the weight of the tails), whereas the parameters of *scale*  $\sigma$  and of *location*  $\mu$  represent the dispersion and average of the extremes, respectively.<sup>6</sup>

<sup>6</sup> In Equation (4), for  $\xi > 0$ ,  $\xi < 0$ , and  $\xi = 0$  we obtain the Frechet, Weibull, and Gumbel families, respectively. The Frechet distribution is fat tailed as its tail is slowly decreasing; the Weibull distribution has no tail—after a certain point there are no extremes; the Gumbel distribution is thin tailed as its tail is rapidly decreasing.

An alternative approach to determine the type of asymptotic distribution for extremes can be based on the concept of generalized Pareto distribution (GPD). Excesses over high thresholds can be modeled by the GPD, which can be derived from the GEV distribution. The GPD of the standardized maxima denoted by  $G_{\max}(x)$  is given by  $G_{\max}(x) = 1 + \ln [H_{\max}(x)]$ , where  $H_{\max}(x)$  is the GEV distribution:

$$G_{\max,\xi}(M; \mu, \sigma) = 1 - \left[ 1 + \xi \left( \frac{M - \mu}{\sigma} \right) \right]^{-1/\xi}. \quad (5)$$

Notice that the GPD presented in Equation (5) nests the standard *Pareto* distribution, the *uniform* distribution on  $[-1, 0]$ , and the standard *exponential* distribution:<sup>7</sup>

$$\text{Pareto} : G_{\max,\xi}(x) = 1 - x^{-1/\xi} \quad \text{for } x \geq 1, \quad (6)$$

$$\text{Uniform} : G_{\max,\xi}(x) = 1 - (-x)^{1/\xi} \quad \text{for } x \in [-1, 0], \quad (7)$$

$$\text{Exponential} : G_{\max,0}(x) = 1 - \exp(-x) \quad \text{for } x \geq 0. \quad (8)$$

To determine whether the GPD or the GEV distribution yields a more accurate characterization of extreme movements in financial markets, Bali (2003b) proposes a more general extreme value distribution using the Box and Cox (1964) transformation:

$$F_{\max,\xi}(M; \mu, \sigma, \phi) = \left( \frac{\left[ \exp \left\{ - \left[ 1 + \xi \left( \frac{M - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \right]^\phi - 1}{\phi} \right) + 1. \quad (9)$$

The Box-Cox-GEV distribution in Equation (9) nests the GPD of Pickands (1975) and the GEV of Jenkinson (1955). More specifically, when  $\phi$  equals one the Box-Cox-GEV reduces to the GEV distribution given in Equation (4): when  $\phi = 1$   $F_{\max,\xi}(x) \rightarrow H_{\max,\xi}(x)$ . When  $\phi$  equals zero, the Box-Cox-GEV converges to the GPD presented in Equation (5): when  $\phi = 0$   $F_{\max,\xi}(x) \rightarrow G_{\max,\xi}(x)$ .<sup>8</sup>

### SKEWED FAT-TAILED DISTRIBUTIONS

This section presents the probability density functions for the SGT, the SGED, the inverse hyperbolic sine (IHS), and the EGB2. As shown in Hansen, McDonald, and Theodossiou (2005), all four densities accommodate diverse distributional characteristics when used for fitting data or as a basis for quasi-maximum likelihood estimation (QMLE) of regression models.

<sup>7</sup> In Equations (6) to (8), the shape parameter,  $\xi$ , determines the tail behavior of the distributions. For  $\xi > 0$ , the distribution has a polynomially decreasing tail (Pareto). For  $\xi = 0$ , the tail decreases exponentially (exponential). For  $\xi < 0$ , the distribution is short tailed (uniform).

<sup>8</sup> See the Appendix for parameter estimation of the Box-Cox-GEV distribution.

The SGT probability density function is

$$SGT(y; \mu, \sigma, n, k, \lambda) = \frac{C}{\sigma} \left( 1 + \frac{1}{((n-2)/k)(1 + \text{sign}(y - \mu + \delta\sigma)\lambda)^k \theta^k \sigma^k} |y - \mu + \delta\sigma|^k \right)^{-(n+1)/k}, \quad (10)$$

where

$$C = k / (2((n-2)/k)^{1/k} \theta B(1/k, n/k)),$$

$$\theta = (k/(n-2))^{1/k} B(1/k, n/k)^{-5} B(3/k, (n-2)/k)^{-5} S(\lambda)^{-1},$$

$$S(\lambda) = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2},$$

$$A = B(2/k, (n-1)/k) B(1/k, n/k)^{-5} B(3/k, (n-2)/k)^{-5},$$

$$\delta = 2\lambda A S(\lambda)^{-1},$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the random variable  $y$ ,  $n$ , and  $k$  are positive kurtosis parameters,  $\lambda$  is a skewness parameter obeying the constraint  $|\lambda| < 1$ ,  $\text{sign}$  is the sign function, and  $B(\cdot)$  is the beta function. In the above density,  $\mu - \delta\sigma$  is the mode and  $\delta = (\mu - \text{mode}(y))/\sigma$  is Pearson's skewness.

The SGT gives several well-known distributions as special cases. Specifically, it gives for  $\lambda = 0$  the generalized- $t$  of McDonald and Newey (1988), for  $k = 2$  the skewed  $t$  of Hansen (1994), for  $n = \infty$  the SGED of Theodossiou (2004), for  $n = \infty$  and  $\lambda = 0$  the generalized error distribution or power exponential distribution of Subbotin (1923) (used by Box and Tiao, 1962; Nelson, 1991), for  $n = \infty$ ,  $\lambda = 0$ ,  $k = 1$  the Laplace or double exponential distribution, for  $n = 1$ ,  $\lambda = 0$ ,  $k = 2$ , the Cauchy distribution, for  $n = \infty$ ,  $\lambda = 0$ ,  $k = 2$  the normal distribution, and for  $n = \infty$ ,  $\lambda = 0$ ,  $k = \infty$  the uniform distribution.

The SGED probability density function is

$$SGED(y; \mu, \sigma, k, \lambda) = \frac{C}{\sigma} \exp \left( -\frac{1}{(1 + \text{sign}(y - \mu + \delta\sigma)\lambda)^k \theta^k \sigma^k} |y - \mu + \delta\sigma|^k \right), \quad (11)$$

where

$$C = k / (2\theta \Gamma(1/k)), \delta = 2\lambda A S(\lambda)^{-1}$$

$$\theta = \Gamma(1/k)^{-5} \Gamma(3/k)^{-5} S(\lambda)^{-1}$$

$$S(\lambda) = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}$$

$$A = \Gamma(2/k)\Gamma(1/k)^{-.5}\Gamma(3/k)^{-.5}$$

$$\delta = 2\lambda A S(\lambda)^{-1},$$

$\mu, \sigma, \lambda, k$ , and *sign* are as defined previously.

The inverse hyperbolic sine (IHS) probability density function is

$$\begin{aligned} IHS(y; \mu, \sigma, k, \lambda) \\ = \frac{k}{\sqrt{2\pi(\theta^2 + (y - \mu + \delta\sigma)^2/\sigma^2)\sigma^2}} \\ \times \exp\left(-\frac{k^2}{2}\left(\ln\left((y - \mu + \delta\sigma)/\sigma + \sqrt{\theta^2 + (y - \mu + \delta\sigma)^2/\sigma^2}\right) - (\lambda + \ln\theta)\right)^2\right), \end{aligned} \quad (12)$$

where  $\theta = 1/\sigma_w$ ,  $\delta = \mu_w/\sigma_w$ ,  $\mu_w = 0.5(e^\lambda - e^{-\lambda})e^{0.5k^{-2}}$ ,  $\sigma_w = 0.5(e^{2\lambda+k^{-2}} + e^{-2\lambda+k^{-2}} + 2)^{0.5}(e^{k^{-2}} - 1)^{0.5}$ ,  $\mu_w$  and  $\sigma_w$  are the mean and standard deviation of  $w = \sinh(\lambda + z/k)$ ,  $\sinh$  is the hyperbolic sine function,  $z$  is a standardized normal variable, and  $\mu$  and  $\sigma$  are the mean and standard deviation of  $y$ .<sup>9</sup> Note that negative (positive) values of  $\lambda$  generate negative (positive) skewness and zero values generate no skewness. Smaller values of  $k$  result in more leptokurtic distributions.

The EGB2 probability density function is

$$EGB2(y; \mu, \sigma, p, q) = C \frac{e^{p(y-\mu+\delta\sigma)/(\theta\sigma)}}{(1 + e^{(y-\mu+\delta\sigma)/(\theta\sigma)})^{p+q}}, \quad (13)$$

where  $C = 1/(B(p, q)\theta\sigma)$ ,  $\delta = (\psi(p) - \psi(q))\theta$ ,  $\theta = 1/\sqrt{\psi'(p) + \psi'(q)}$ ,  $p$  and  $q$  are positive scaling constants,  $B(\cdot)$  is the beta function,  $\psi(z) = d \ln \Gamma(z)/dz$  and  $\psi'(z) = d\psi(z)/dz$  are the psi function and its first derivative, and  $\mu$  and  $\sigma$  are the mean and standard deviation of  $y$ . The EGB2 is symmetric for equal values of  $p$  and  $q$ , positively skewed for values of  $p > q$ , and negatively skewed for values of  $p < q$ . The EGB2 converges to the normal distribution for infinite values of  $p$  and  $q$ .

## VAR MODELS WITH ALTERNATIVE DISTRIBUTION FUNCTIONS

The discrete time version of the geometric Brownian motion governing financial price movements is:

$$R_t = \ln(P_{t+\Delta t}) - \ln(P_t) = \mu^* \Delta t + z\sigma^* \sqrt{\Delta t}, \quad (14)$$

where  $\mu^*$  and  $\sigma^*$  are the annualized mean and standard deviation of  $R_t$ ,  $\Delta t$  is the length of time between two successive prices, and  $\Delta W_t = z\sqrt{\Delta t}$  is a discrete approximation of

<sup>9</sup> The IHS parameterization used in Johnson (1949) and Johnson, Kotz, and Balakrishnan (1994) is  $w = \sinh((-\lambda+z)/k)$ . Our parameterization makes the interpretation of the scaling parameters  $\lambda$  and  $k$  easier and comparable across models.

the Wiener process. The latter term has mean zero, variance  $\Delta t$ , and follows the normal distribution. In the case of time-varying mean and variance, the above equation can be modified to reflect these dependencies as:

$$R_t = \mu_t^* \Delta t + z_t^* \sqrt{\Delta t}, \quad (15)$$

where  $\mu_t^*$  and  $\sigma_t^*$  are annualized measures of the conditional mean and conditional standard deviation of the log-return  $R_t$ . For simplicity, the above equation can be rewritten as

$$R_t = \mu_t + z_t \sigma_t, \quad (16)$$

where  $\mu_t = \mu_t^* \Delta t$  and  $\sigma_t = \sigma_t^* \sqrt{\Delta t}$  are, respectively, the conditional mean and conditional standard deviation of daily returns. Note that by construction the standardized return  $z_t = (R_t - \mu_t)/\sigma_t$  preserves the properties of zero mean and unit variance.

There are different ways to model the conditional first and second moments of the distribution of the returns. The most popular approach is based on the family of generalized autoregressive conditional heteroskedasticity (GARCH) models. In a later section of this article we use the absolute GARCH-in-mean to model the conditional first and second moments of returns.

There is substantial empirical evidence showing that the distributions of returns of financial assets are typically skewed to the left and peaked around the mode, and have fat tails. The fat tails suggest that extreme outcomes happen much more frequently than would be predicted by the normal distribution. Similarly, the distribution of standardized returns, although normalized, follows a similar pattern.

The conditional threshold for  $R_t$  at a given coverage probability (or confidence level)  $\rho$ , denoted by  $\vartheta_t$ , is obtained from the solution of the following cumulative distribution of returns,

$$\Pr(R_t \leq \vartheta_t | I_{t-1}) \equiv \int_{-\infty}^{\vartheta_t} f(R_t | I_{t-1}) dR_t = \rho, \quad (17)$$

where  $\Pr(\cdot)$  denotes the probability and  $f(R_t | I_{t-1})$  is the conditional probability density function for  $R_t$ , conditional on the information set  $I_{t-1}$ . The above probability function can be expressed in terms of the standardized returns as follows:

$$\begin{aligned} \Pr(R_t \leq \vartheta_t | I_{t-1}) &= \Pr\left(\frac{R_t - \mu_t}{\sigma_t} \leq \frac{\vartheta_t - \mu_t}{\sigma_t} \middle| I_{t-1}\right) \\ &= \Pr\left(z_t \leq a = \frac{\vartheta_t - \mu_t}{\sigma_t}\right) = \int_{-\infty}^a f(z_t) dz_t = F(a) = \rho, \end{aligned} \quad (18)$$

where  $F$  is a cumulative density function of  $f(z_t)$ . The density  $f(z_t)$  and the threshold  $a$  associated with the coverage probability  $\rho$  do not depend on the information set  $I_{t-1}$ .

The latter is a byproduct of the assumption that the series of standardized returns  $z_t$  are by construction identically and independently distributed. The latter assumption is consistent with empirical evidence related to the GARCH models for economic and financial price series. It follows easily from the above equation that conditional threshold equation is

$$\vartheta_t = \mu_t + \alpha \sigma_t, \quad (19)$$

where the unconditional threshold  $\alpha = F^{-1}(\rho)$  is obtained by inverting the cumulative probability density function  $F$ .

In the case of standard normal distribution and a coverage probability of 1 percent (i.e.,  $\rho = 1$  percent),  $\alpha = 2.326$ , and  $\vartheta_t = \mu_t - 2.326\sigma_t$ . In the case of the fat-tailed distributions, the threshold  $\alpha$  is a function of both the skewness and kurtosis parameters. In general fatter tails will result in larger absolute values for  $\alpha$ . In the case of Box-Cox-GEV distribution,

$$\alpha = -\frac{1}{\xi} \left[ 1 + \frac{1}{\phi} \ln \left( 1 - \frac{\rho N \phi}{n} \right)^{-\xi} \right], \quad (20)$$

where  $\mu$  (location),  $\sigma$  (scale),  $\xi$  (shape), and  $\phi$  are the Box-Cox-GEV distribution parameters and  $n$  and  $N$  are respectively the number of extremes and total data points. In the special case of the GPD,  $\phi = 0$ , and

$$\alpha = -\frac{1}{\xi} \left[ 1 - \left( \frac{\rho N}{n} \right)^{-\xi} \right]. \quad (21)$$

In the special case of GEV,  $\phi = 1$ , and

$$\alpha = -\frac{1}{\xi} \left[ \left( 1 + \ln \left( 1 - \frac{\rho N}{n} \right) \right)^{-\xi} \right]. \quad (22)$$

## DATA

The data set consists of daily (percentage) log-returns (log-price changes) for the S&P 500 composite index. The time period of investigation for the S&P 500 is January 4, 1950 to December 29, 2000 (12,832 observations). Panel A of Table 1 shows that the unconditional mean of daily log-returns on S&P 500 is 0.0341 percent with a standard deviation of 0.87 percent. The maximum and minimum values are 8.71 percent and  $-22.90$  percent. Panel A also reports the skewness and excess kurtosis statistics for testing the distributional assumption of normality. The skewness statistic for daily S&P 500 returns is negative and statistically significant at the 1 percent level. The excess kurtosis statistic is considerably high and significant at the 1 percent level, implying that the distribution of equity returns has much fatter tails than

**TABLE 1**  
Descriptive Statistics

**Panel A: Summary Statistics of Daily Returns**

|                     | S&P 500<br>(January 4, 1950–December 29, 2000) |
|---------------------|------------------------------------------------|
| No. of observations | 12,832                                         |
| Maximum             | 8.7089                                         |
| Minimum             | −22.899                                        |
| Mean                | 0.0341                                         |
| Std. dev.           | 0.8740                                         |
| Skewness            | −1.6216**                                      |
| Kurtosis            | 45.525**                                       |

**Panel B: Summary Statistics of Extremes**

|        | <i>n</i> | Mean    | Std. Dev. | Maximum | Minimum |
|--------|----------|---------|-----------|---------|---------|
| Maxima | 642      | 1.9607  | 0.7039    | 8.7089  | 1.3491  |
| Minima | 642      | −1.9741 | 1.1656    | −1.3150 | −22.900 |

*Notes:* Panel A presents several descriptive statistics of daily percentage returns on the S&P 500 stock market index, computed using the formula,  $R_t = 100 \times (\ln P_t - \ln P_{t-1})$ , where  $P_t$  is the value of the stock market index at time  $t$ . Panel B presents statistics on the maximal and minimal returns obtained from the 5 percent of the right and left tails of the empirical distribution. \*\* denotes significance at the 1 percent level.

the normal distribution. The fat-tail property is more dominant than skewness in the sample.<sup>10</sup>

The maximal and minimal returns are obtained from the daily data described above. Following the extreme value theory, we define the extremes as excesses over high thresholds (see Embrechts, Kluppelberg, and Mikosch 1997, pp. 352–355). Specifically, the extreme changes are defined as the 5 percent of the right and left tails of the empirical distribution. Panel B of Table 1 shows the means, standard deviations, and maximum and minimum values of the extremes. In addition to the 5 percent tails, the extremes are obtained from the 2.5 percent and 10 percent tails of the empirical distribution. The qualitative results are found to be robust across different threshold levels. To save space we do not to present the empirical findings based on the 2.5 percent and 10 percent tails.

#### Empirical Results for the Extreme Value Distributions

Panel A of Table 2 presents the parameter estimates of the Box-Cox-GEV, generalized Pareto, and GEVs. The empirical results show unambiguously that the asymptotic

<sup>10</sup> In addition to the S&P 500 stock market index, at an earlier stage of the study, we use the Dow Jones Industrial Average (DJIA) that cover the period from May 26, 1896 to December 29, 2000 (28,758 observations). The risk measurement performance of alternative distribution functions turns out to be similar to those obtained from the S&P 500. To preserve space, we do not present the empirical results for the DJIA. They are available from the authors upon request.

**TABLE 2**  
Parameter Estimates

| Panel A: Regression Method Estimates of the Extreme Value Distributions |                        |                     |                    |              |        |        |
|-------------------------------------------------------------------------|------------------------|---------------------|--------------------|--------------|--------|--------|
| Maxima                                                                  | $\mu_{max}$            | $\sigma_{max}$      | $\xi_{max}$        | $\phi_{max}$ | Log-L  | LR     |
| Box-Cox-GEV                                                             | 0.01485<br>(1013.19)   | 0.00440<br>(188.51) | 0.2083<br>(70.621) | 0.51         | 892.52 | –      |
| GPD                                                                     | 0.01362<br>(582.82)    | 0.00503<br>(152.53) | 0.1701<br>(51.295) | 0.00         | 856.16 | 72.72  |
| GEV                                                                     | 0.01528<br>(855.58)    | 0.00316<br>(126.71) | 0.3291<br>(59.807) | 1.00         | 574.77 | 635.50 |
| Minima                                                                  | $\mu_{min}$            | $\sigma_{min}$      | $\xi_{min}$        | $\phi_{min}$ | Log-L  | LR     |
| Box-Cox-GEV                                                             | –0.01559<br>(–1296.42) | 0.00303<br>(180.99) | 0.4527<br>(118.10) | 0.94         | 660.89 | –      |
| GPD                                                                     | –0.01360<br>(–444.77)  | 0.00383<br>(90.081) | 0.3906<br>(65.745) | 0.00         | 510.03 | 301.72 |
| GEV                                                                     | –0.01479<br>(–1139.78) | 0.00283<br>(167.18) | 0.4578<br>(108.19) | 1.00         | 541.24 | 239.30 |

(continued)

**TABLE 2**  
(Continued)

| Panel B: Maximum Likelihood Estimates of the Normal and Flexible Distributions |                     |                       |                       |                       |                      |         |        |           |               |                |           |              |  |
|--------------------------------------------------------------------------------|---------------------|-----------------------|-----------------------|-----------------------|----------------------|---------|--------|-----------|---------------|----------------|-----------|--------------|--|
|                                                                                | $\mu$               | $\sigma$              | $\lambda$             | $k$                   | $n$                  | $S$     | $K$    | $Log-L$   | $LR_{Normal}$ | $LR_{Laplace}$ | $KS$      | $Lilliefors$ |  |
| SGT                                                                            | 0.03559<br>(4.82)** | 0.85012<br>(86.63)**  | -0.02455<br>(-2.23)*  | 1.60084<br>(26.42)**  | 5.21132<br>(12.79)** | -0.1336 | 11.045 | -15140.31 | 2677.86**     | 184.33**       | 0.00928   | 0.00919      |  |
| SGED                                                                           | 0.03383<br>(4.64)** | 0.84232<br>(124.57)** | -0.02775<br>(-3.23)** | 1.05063<br>(110.61)** | -                    | -0.1097 | 5.6125 | -15223.02 | 2512.46**     | 18.93**        | 0.01480** | 0.01471**    |  |
| IHS                                                                            | 0.03502<br>(4.67)** | 0.85512<br>(85.29)**  | -0.04021<br>(-2.27)*  | 1.30407<br>(49.92)**  | -                    | -0.1712 | 10.041 | -15146.84 | -             | -              | 0.00946   | 0.00937      |  |
| Normal                                                                         | 0.03408<br>(0.03)   | 0.87397<br>(0.87)**   | 0                     | 2                     | -                    | 0       | 3      | -16479.24 | -             | -              | 0.06393** | 0.06385**    |  |
|                                                                                | $\mu$               | $\sigma$              | $q$                   | $p$                   | $n$                  | $S$     | $K$    | $Log-L$   | $LR_{Normal}$ | $LR_{Laplace}$ | $KS$      | $Lilliefors$ |  |
| EGB2                                                                           | 0.03412<br>(4.81)** | 0.83361<br>(125.93)** | 0.25270<br>(9.80)**   | 0.23774<br>(10.03)**  | -                    | -0.1166 | 5.6238 | -15191.77 | 2574.95**     | -              | 0.01495** | 0.01487**    |  |

Notes: Panel A presents the parameter estimates of the extreme value distributions. Asymptotic  $t$ -statistics are given in parentheses.  $Log-L$  denotes the maximized log-likelihood values and  $LR$  the log-likelihood ratio test statistics between the Box-Cox-GEV with GPD and GEV distributions. Panel B presents the parameter estimates of the normal and four flexible distributions. Asymptotic  $t$ -statistics are given in parentheses.  $\mu$ ,  $\sigma$ ,  $S$ , and  $K$  are the estimated mean, standard deviation, skewness, and kurtosis statistics.  $LR_{Normal}$  is the LR statistic from testing the null hypothesis that the daily returns are distributed as normal against the alternative hypothesis that they are distributed as SGT, SGED, and EGB2.  $LR_{Laplace}$  is the LR statistic for testing the null hypothesis that the series are distributed as Laplace against the alternative hypothesis that they are distributed as SGT and SGED. Kolmogorov-Smirnov ( $KS$ ) and Lilliefors statistics are used for testing the null hypothesis that the data follow the corresponding distribution. \*, \*\* denote significance at the 5 percent and 1 percent level, respectively.

distribution of the maximal and minimal returns on S&P 500 belongs to the domain of attraction of the Box-Cox-GEV distribution. A likelihood ratio (LR) test between the GEV and Box-Cox-GEV distributions leads to a rejection of the GEV distribution.<sup>11</sup> The LR test between the GPD and Box-Cox-GEV distributions also indicates a rejection of the GPD distribution.

The parameters of the Box-Cox-GEV distribution are estimated using a one-dimensional grid search method based on the nonlinear least square estimation technique. Since one of our goals in this section is to determine whether the asymptotic distribution of extremes belongs to the domain of attraction of GEV or GPD, the value of  $\phi$  is expected to be between zero and one. In some cases, we had a difficulty in obtaining a “global maximum” of the log-likelihood function. Hence, the parameter  $\phi$  is estimated by scanning this range  $[0, 1]$  in increments of 0.1. When a minimum of the sum of squares is found, greater precision is desired, and the area to the right and left of the current optimum is searched in increments of 0.01. As shown in Panel A of Table 2,  $\phi_{\max}$  is estimated to be 0.51 for the maximal returns on S&P 500, and  $\phi_{\min}$  is found to be 0.94 for the minimal returns. The LR test results indicate that both the GPD with  $\phi = 0$  and the GEV with  $\phi = 1$  are strongly rejected in favor of the Box-Cox-GEV distribution with  $0 < \phi < 1$ .

The tail index  $\xi$  for the GPD and GEV distributions is found to be positive and statistically different from zero. This implies a rejection of the thin-tailed ( $\xi = 0$ ) Gumbel and exponential distributions with rapidly decreasing tails against the fat-tailed ( $\xi > 0$ ) Frechet and Pareto distributions with slowly decreasing tails, and *a fortiori* a rejection of the short-tailed ( $\xi < 0$ ) Weibull and uniform distributions. The asymptotic *t*-statistics of the estimated shape parameters ( $\xi$ ) clearly indicate the nonnormality of extremes. Another notable point in Panel A of Table 2 is that the estimated shape parameters for the minimal returns ( $\xi_{\min}$ ) turn out to be greater than those for the maximal returns ( $\xi_{\max}$ ). Specifically, the estimated  $\xi_{\max}$  values are in the range of 0.17 to 0.33 for the maximal returns, while the estimates of  $\xi_{\min}$  vary from 0.39 to 0.46 for the minimal returns. Since the higher  $\xi$  the fatter the distribution of extremes, the minimal returns have thicker tails than the maximal returns.

### Empirical Results for the Flexible Distributions

Panel B of Table 2 presents the estimated parameters of the SGT, SGED, inverse hyperbolic sign (IHS), EGB2, and normal distributions for the S&P 500 log-returns. The estimates are obtained using the maximum likelihood method and the iterative algorithm described in Theodossiou (1998).

The first two columns of the table present the estimates for the mean and standard deviation of log-returns on S&P 500. As expected, these estimates are quite similar across distributions and do not differ much from the simple arithmetic means and standard deviations of log-returns presented in Panel A of Table 1.

The third column presents the skewness parameter  $\lambda$ , which is negative and statistically significant, indicating negatively skewed distributions. The fourth and fifth

<sup>11</sup> The LR statistic is calculated as  $LR = -2 [\text{Log-L}^* - \text{Log-L}]$ , where  $\text{Log-L}^*$  is the value of the log likelihood under the null hypothesis, and  $\text{Log-L}$  is the log likelihood under the alternative. This statistic is distributed as  $\chi^2$  with one degree of freedom.

column present the estimates for the kurtosis parameters  $k$  and  $n$  (SGT only). In the case of SGT, the values of  $k = 1.6$  and  $n = 5.21$  are quite different from those of the normal distribution of  $k = 2$  and  $n = \infty$ . These parameter estimates indicate that the S&P 500 log-returns are characterized by excess kurtosis. Note that in the SGT model, the parameter  $k$  controls mainly the peakness of the distribution around the mode while the parameter  $n$  controls mainly the tails of the distribution, i.e., adjusting the tails to the extreme values. The parameter  $n$  has the degrees of freedom interpretation as in the Student- $t$  distribution.

Columns 6 and 7 present the standardized measures for skewness  $S = E(y-\mu)^3/\sigma^3$  and kurtosis  $K = E(y-\mu)^4/\sigma^4$  based on the parameter estimates for  $k$ ,  $n$ , and  $\lambda$ .<sup>12</sup> The standardized skewness,  $S$ , is negative indicating that the distribution of log-returns is skewed to the left. Note that the moments of the SGT exist up to the value of  $n$ . Since  $n$  is estimated to be above five, the SGT has a defined kurtosis statistic. As shown in Panel B, the standardized kurtosis for S&P 500 is 11.045.

The above results provide strong support to the hypothesis that stock returns are not normal. The normality hypothesis is also rejected by the LR statistics for testing the null hypothesis of normality against that of SGT. Note that the LR statistics, presented in column 9, are quite large and statistically significant at the 1 percent level. To test the overall fit of the SGT, we also use the Kolmogorov–Smirnov (KS) and Lilliefors (1967) test statistics.<sup>13</sup> The KS and Lilliefors test statistics, presented in the last column of Panel B, are small and statistically insignificant at both the 1 percent and 5 percent levels, providing support for the null hypothesis that the data are SGT and IHS distributed.

The SGED estimate for the kurtosis parameter  $k$  is close to one ( $k = 1.05$ ) and it is considerably lower than that of SGT. This is because the SGED's kurtosis is controlled by parameter  $k$  only, thus, to account for the excess kurtosis in the data the parameter  $k$  has to be smaller than that of SGT.<sup>14</sup> The skewness parameter  $\lambda$  is negative and statistically significant for the S&P 500 data. The standardized skewness and kurtosis parameters are smaller than those of SGT. SGED's nesting property allows us to test the null hypothesis that the data follow the SGED against the alternative hypothesis that they follow the SGT. The log-LR for testing the latter hypothesis is 165.42 for the S&P 500. This ratio, which follows the chi-square distribution with one degree of freedom, is large and statistically significant at the 1 percent level, thus suggesting that the SGT provides a better fit than the SGED. Moreover, the KS and Lilliefors statistics reject the null hypothesis that the S&P 500 log-returns follow the SGED. The superiority of the SGT over the SGED can be attributed to the fact that it provides a better fit to the tails of the distribution because of the parameter  $n$ .

<sup>12</sup> The formulas for computing  $S$  and  $K$  are presented in Hansen, McDonald, and Theodossiou (2005). These formulas are functions of the parameter estimates for skewness and kurtosis.

<sup>13</sup> The Lilliefors (1967) test statistic is used to test the null hypothesis that data come from a normally distributed (or SGT distributed) population (for example), when the null hypothesis does not specify which normal (or which SGT) distribution, i.e., does not specify the parameters.

<sup>14</sup> The SGED is derived asymptotically from SGT by setting  $n = \infty$ .

The IHS and EGB2 distributions are not linked directly to each other or the SGT, but the EGB2 is linked with the normal distribution. Specifically, as the parameters  $p$  and  $q$  approach infinity the EGB2 converges to the normal distribution. The possible values for kurtosis are limited to the range [3, 9] for EGB2; see Hansen, McDonald, and Theodossiou (2005). Like in the case of SGT and SGED, the results for IHS and EGB2 indicate that the S&P 500 log-returns exhibit skewness and significant excess kurtosis. The LR statistic for testing the null hypothesis of normal distribution against the alternative hypothesis of EGB2 rejects normality. The LR ratio for IHS does not exist because the normal distribution is not nested within IHS. The KS and Lilliefors statistics indicate that the hypothesis that the data follow IHS cannot be rejected but the EGB2 hypothesis is rejected. The latter may be attributed to the inability of EGB2 to model kurtosis values outside the range of 3 to 9.

### **VAR CALCULATIONS WITH ALTERNATIVE DISTRIBUTION FUNCTIONS**

Table 3 compares the in-sample performance of VaR measures based on the unconditional distribution functions for the S&P 500 daily log-returns. The top numbers at each row are the VaR thresholds estimated by the extreme value, normal, and flexible distributions. The results indicate that the extreme tails yield threshold points,  $\vartheta_{Box-Cox-GEV}$ , that are up to 30 percent higher than the normal thresholds,  $\vartheta_{Normal}$ . The multiplication factors ( $\vartheta_{Box-Cox-GEV} / \vartheta_{Normal}$ ) for the extreme tails (0.5 percent and 1 percent) are in the range of 1.13 to 1.23 for the maxima and 1.09 to 1.22 for the minima. Although not presented in this article, the multiplication factors for the extreme tails of the DJIA are in the range of 1.22 to 1.40 for the minimal returns and 1.12 to 1.28 for the maximal returns. The VaR measures for the flexible distributions are similar to those of the extreme value distributions. Specifically, the average ratio of VaR thresholds for the maximal (minimal) returns on S&P500 are  $\vartheta_{Box-Cox-GEV} / \vartheta_{SGT} = 1.0058$  (0.9803),  $\vartheta_{Box-Cox-GEV} / \vartheta_{SGED} = 0.9923$  (0.9609),  $\vartheta_{Box-Cox-GEV} / \vartheta_{EGB2} = 1.0091$  (0.9753), and  $\vartheta_{Box-Cox-GEV} / \vartheta_{IHS} = 1.0012$  (0.9679). These findings imply that the tails of the empirical distribution approximated by the flexible distributions are similar to those of the extreme value distributions.

Table 3 presents the estimated counts or the actual number of observations falling in various tails of the extreme value, normal, and flexible distributions. The results show that the normal VaR thresholds at various tails are quite inadequate. Given that the S&P500 data include 12,832 returns from December 4, 1950 to December 29, 2000, one would expect 64, 128, 192, 257, 321, and 642 returns to fall respectively into the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent positive and negative tails. The number of returns for the normal VaR thresholds falling into the positive (negative) 0.5 percent tail are 135 (130), 1 percent tail are 197 (178), 1.5 percent tail are 251 (223), 2 percent tail are 290 (275), 2.5 percent tail are 337 (316), and 5 percent tail are 549 (516). Based on these results, the normal distribution underestimates the actual VaR thresholds at the 0.5 percent, 1 percent, 1.5 percent, and 2 percent tails and overestimates the VaR thresholds for the 5 percent tail.

The last row of Table 3 gives the average mean absolute percentage errors (MA percentEs) based on the actual and expected counts. Specifically, the Average MA percentE is calculated as

**TABLE 3**  
In-Sample Performance of Unconditional Distribution Functions for the S&P 500

| Maxima       | Expected | Box-Cox-GEV            | GPD                    | GEV                    | Normal                    | SGT                    | SGED                    | EGB2                    | IHS                    |
|--------------|----------|------------------------|------------------------|------------------------|---------------------------|------------------------|-------------------------|-------------------------|------------------------|
| 0.5% VaR     | 64       | 2.7659%<br>64 (0.00)   | 2.7790%<br>63 (0.02)   | 2.768%<br>69 (0.36)    | 2.2853%<br>135 (59.57)**  | 2.7554%<br>64 (0.00)   | 2.6715%<br>74 (1.44)    | 2.6347%<br>77 (2.43)    | 2.8001%<br>62 (0.07)   |
| 1% VaR       | 128      | 2.2934%<br>125 (0.09)  | 2.2926%<br>125 (0.09)  | 2.2537%<br>134 (0.25)  | 2.0673%<br>197 (31.91)**  | 2.2865%<br>125 (0.09)  | 2.2891%<br>125 (0.09)   | 2.2526%<br>134 (0.25)   | 2.3090%<br>122 (0.32)  |
| 1.5% VaR     | 192      | 2.0411%<br>195 (0.03)  | 2.0336%<br>197 (0.11)  | 2.0246%<br>199 (0.22)  | 1.9307%<br>251 (16.49)**  | 2.0306%<br>198 (0.16)  | 2.0642%<br>186 (0.22)   | 2.0290%<br>198 (0.16)   | 2.0417%<br>193 (0.00)  |
| 2% VaR       | 257      | 1.8703%<br>262 (0.11)  | 1.8603%<br>264 (0.21)  | 1.8724%<br>259 (0.02)  | 1.8290%<br>290 (4.25)*    | 1.8564%<br>266 (0.34)  | 1.9039%<br>249 (0.23)   | 1.8703%<br>262 (0.11)   | 1.8608%<br>264 (0.21)  |
| 2.5% VaR     | 321      | 1.7413%<br>319 (0.01)  | 1.7317%<br>329 (0.21)  | 1.7572%<br>308 (0.53)  | 1.7471%<br>337 (0.83)     | 1.7251%<br>330 (0.27)  | 1.7792%<br>295 (2.19)   | 1.7472%<br>315 (0.11)   | 1.7250%<br>330 (0.27)  |
| 5% VaR       | 642      | 1.3429%<br>648 (0.07)  | 1.3618%<br>630 (0.22)  | 1.3607%<br>630 (0.22)  | 1.4717%<br>549 (14.76)**  | 1.3353%<br>651 (0.14)  | 1.3895%<br>594 (3.81)   | 1.3649%<br>626 (0.40)   | 1.3268%<br>663 (0.74)  |
| Minima       | Expected | Box-Cox-GEV            | GPD                    | GEV                    | Normal                    | SGT                    | SGED                    | EGB2                    | IHS                    |
| 0.5% VaR     | 64       | -2.7486%<br>59 (0.43)  | -2.7898%<br>56 (1.09)  | -2.7301%<br>62 (0.07)  | -2.2172%<br>130 (52.26)** | -2.7941%<br>56 (1.09)  | -2.7177%<br>65 (0.01)   | -2.6889%<br>68 (0.23)   | -2.8743%<br>51 (2.92)  |
| 1% VaR       | 128      | -2.2154%<br>137 (0.58) | -2.2181%<br>135 (0.35) | -2.2011%<br>137 (0.58) | -1.9991%<br>178 (17.34)** | -2.3000%<br>113 (1.92) | -2.3138%<br>112 (2.19)  | -2.2827%<br>118 (0.86)  | -2.3472%<br>108 (3.43) |
| 1.5% VaR     | 192      | -1.9640%<br>198 (0.16) | -1.9489%<br>204 (0.69) | -1.9561%<br>202 (0.47) | -1.8626%<br>223 (4.68)*   | -2.0302%<br>179 (0.98) | -2.0762%<br>168 (3.30)  | -2.0452%<br>174 (1.86)  | -2.0601%<br>171 (2.53) |
| 2% VaR       | 257      | -1.8052%<br>272 (0.92) | -1.7821%<br>280 (2.11) | -1.8027%<br>274 (1.17) | -1.7609%<br>275 (1.31)    | -1.8464%<br>247 (0.37) | -1.9069%<br>219 (5.92)* | -1.8765%<br>226 (3.89)* | -1.8656%<br>235 (1.92) |
| 2.5% VaR     | 321      | -1.6899%<br>331 (0.33) | -1.6650%<br>339 (1.04) | -1.6919%<br>331 (0.33) | -1.6790%<br>316 (0.07)    | -1.7079%<br>321 (0.00) | -1.7751%<br>281 (5.28)* | -1.7457%<br>294 (2.36)  | -1.7199%<br>311 (0.31) |
| 5% VaR       | 642      | -1.2977%<br>662 (0.68) | -1.3601%<br>597 (3.34) | -1.3339%<br>619 (0.85) | -1.4036%<br>516 (27.66)** | -1.2970%<br>662 (0.68) | -1.3635%<br>592 (4.14)* | -1.3393%<br>616 (1.09)  | -1.2928%<br>664 (0.81) |
| Average MA%E | -        | 3.12%                  | 4.95%                  | 4.29%                  | 34.53%                    | 4.26%                  | 8.45%                   | 6.87%                   | 6.59%                  |

Notes: This table presents statistics on the in-sample performance of VaR thresholds based on the unconditional distribution functions for the S&P500 daily returns spanning the period January 4, 1950 to December 29, 2000 (12,832 observations). The column "Expected" gives the number of observations expected to fall in the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent left (minima) and right (maxima) tails of the return distribution. The top numbers at each row give the estimated VaR thresholds for each distribution. The numbers below the thresholds give the actual counts along with their respective LR statistics (in parentheses). Average MA percentE is the average mean absolute percentage error based on the actual and expected counts. \*, \*\* denote significance at the 5 percent and 1 percent levels, respectively.

$$\text{Average MA}\%E = \frac{1}{6} \sum_{i=1}^6 \left( \frac{\tau_i - \rho_i N}{\rho_i N} \right), \quad (23)$$

where  $\tau_i$  and  $\rho_i N$  are the actual and expected number of observations that fall below the threshold for the coverage probabilities  $\rho_i = 0.5$  percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent;  $i = 1, 2, \dots, 6$ . Equation (23) measures the relative performance of alternative distribution functions.

The normal VaR estimates have an MA percentE of 34.53 percent for the S&P 500. The Box-Cox-GEV distribution has the best overall performance, although the GPD, GEV, SGT, EGB2, and IHS appear to be very similar in performance to the Box-Cox-GEV. All five distributions estimate the actual VaR thresholds very well. Their MA percentEs are 3.12 percent for the Box-Cox-GEV distribution, 4.95 percent for the GPD, 4.29 percent for the GEV, 4.26 percent for the SGT, 6.59 percent for the IHS, and 6.87 percent for the EGB2. The SGED performs better than the normal distribution, however, its MA percentE's is slightly greater than those of the other flexible distributions. The results indicate that the extreme value and flexible distributions are superior to the normal distribution in calculating VaR.

Given the obvious importance of VaR estimates to financial institutions and their regulators, evaluating the accuracy of the distribution functions underlying them is a necessary exercise. The evaluation of VaR estimates is based on hypothesis testing using the binomial distribution. This latter test is currently embodied in the market risk amendment (MRA).<sup>15</sup> Under the hypothesis-testing method, the null hypothesis is that the VaR estimates exhibit the property characteristics of accurate VaR measures. If the null hypothesis is rejected, the VaR estimates do not exhibit the specified property, and the underlying distribution function can be said to be "inaccurate." Otherwise, the VaR model is "acceptably accurate."

Under the MRA, banks will report their VaR estimates to their regulators, who observe when actual portfolio losses exceed these estimates. As discussed by Kupiec (1995), assuming that the VaR measures are accurate, such exceptions can be modeled as independent draws from a binomial distribution with a probability of occurrence equal to (say) 1 percent. The LR statistic for testing the null hypothesis that the actual and the expected number of observations falling below each threshold are statistically the same is computed using the formula

$$LR = 2 [\tau \ln(\tau/(\rho N)) + (N - \tau) \ln((N - \tau)/(N - \rho N))], \quad (24)$$

where  $N$  is the number of sample observations,  $\rho$  is the coverage probability,  $\rho N$  is the expected number of observations to fall below the threshold  $a$ , and  $\tau$  is the actual number of sample observations that falls below the threshold  $a$ . The LR statistic has an asymptotic chi-square distribution with one degree of freedom,  $\chi^2(1)$ .

<sup>15</sup> In August 1996, the U.S. bank regulatory agencies adopted the MRA to the 1998 Basle Capital Accord. The MRA, which became effective in January 1998, requires that commercial banks with significant trading activities set aside capital to cover the market risk exposure in their trading accounts. The market risk capital requirements are to be based on the VaR estimates generated by the banks' own risk management models.

Table 3 presents in parentheses the LR statistics from testing the null hypothesis that the reported VaR estimates are acceptably accurate. According to the LR statistics for the binomial method, the standard approach that assumes normality of asset returns is strongly rejected for the 0.5 percent, 1 percent, 1.5 percent, and 5 percent VaRs for both the maximal and minimal returns. The extreme value (Box-Cox-GEV, GPD, GEV) and skewed fat-tailed (SGT, IHS) distributions produce acceptably accurate VaR estimates for all VaR tails and for both the maximal and minimal returns on S&P 500. In general, the SGED and EBG2 distributions also produce accurate VaR measures with few exceptions for the minimal returns: the 2 percent VaR estimates of EGB2, and the 2 percent, 2.5 percent, and 5 percent VaR estimates of SGED are found to be inaccurate at the 5 percent level of significance.

As discussed by Christoffersen (1998), VaR estimates can be viewed as interval forecasts of the lower and upper tails of the return distribution. Interval forecasts can be evaluated conditionally and unconditionally, that is, with or without reference to the information available at each point in time. The  $LR_{uc}$  given in Equation (24) is an unconditional test statistic because it simply counts exceedences (or violations) over the entire period. However, in the presence of serial correlation, volatility clustering or volatility persistence, the conditional accuracy of VaR estimates becomes an important issue. The VaR models that ignore mean-volatility dynamics may have correct unconditional coverage, but at any given time, they will have incorrect conditional coverage. In such cases, the  $LR_{uc}$  test is of limited use since it will classify inaccurate VaR estimates as acceptably accurate. Moreover, as indicated by Kupiec (1995) and Christoffersen (1998), the unconditional coverage tests have low power against alternative hypotheses if the sample size is small. This problem does not exist here since our daily data sets cover a long period of time (28,758 observations for DJIA and 12,832 observations for S&P 500).

The conditional coverage test developed by Christoffersen (1998) determines whether the VaR estimates exhibit both correct unconditional coverage and serial independence. In other words, if a VaR model produces acceptably accurate conditional thresholds then not only the exceedences implied by the model occur  $x$  percent (say 1 percent) of the time, but they are also independent and identically distributed over time. Given a set of VaR estimates, the indicator variable is constructed as

$$E_t = \begin{cases} 1 & \text{if exceedence occurs} \\ 0 & \text{if no exceedence occurs} \end{cases} \quad (25)$$

and should follow an i.i.d. Bernoulli sequence with the targeted exceedence rate (say 1 percent). The  $LR_{cc}$  test is a joint test of two properties: correct unconditional coverage and serial independence,

$$LR_{cc} = LR_{uc} + LR_{ind}, \quad (26)$$

which is asymptotically distributed as the chi-square with two degrees of freedom,  $\chi^2(2)$ . The  $LR_{ind}$  is the LR test statistic for the null hypothesis of serial independence

against the alternative of first-order Markov dependence.<sup>16</sup> We calculated the  $LR_{ind}$  and  $LR_{cc}$  statistics for all the distribution functions considered in the article.<sup>17</sup> We also computed the first-order serial correlation coefficient  $corr(E_t, E_{t-1})$ , a diagnostic suggested by Christoffersen and Diebold (2000), and tested its statistical significance. The  $LR_{cc}$  and  $corr(E_t, E_{t-1})$  statistics indicate strong rejection of the null hypothesis, implying that given an exceedence (or violation) on one day there is a high probability of a violation the next day.

We have so far compared the in-sample performance of alternative distribution functions based on the unconditional coverage. However, the real test is the out-of-sample performance of conditional distributions. The risk manager, by definition, obtains VaR estimates in real time and hence must use parameters obtained from an already observed sample to evaluate the risks associated with current and future random movements in market variables. Hence, a true test for the extreme value and skewed fat-tailed distributions is their performance outside the sample, which is not used in the estimation of the underlying parameters. To measure the out-of-sample performance, we proceed as follows.

A 10-year rolling sample, starting January 4, 1950, is used to estimate the parameters of the distributions and a 1-year holdout sample (year subsequent to the estimation) to evaluate their performance. Specifically, the first rolling (estimation) sample includes the returns for the years 1950 to 1959 (2,514 returns) and the first holdout sample includes the returns for the year 1960. The estimated VaR thresholds (based on rolling sample) are then used to compute the number of returns in the 1960 holdout sample falling in the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent left (minima) and right (maxima) tails of each estimated probability distribution. Next, the estimation sample is rolled forward by removing the returns for the year 1950 and adding the returns for the year 1960. Consequently, the new holdout sample includes the returns for the year 1961. The new estimated VaR thresholds are then used to compute the number of returns in the new holdout sample falling in the aforementioned left and right tails. The procedure continues until the sample is exhausted. Given that the returns on S&P500 span the period January 4, 1950 to January 29, 2000, the 10-year rolling estimation procedure yields a total holdout sample of 41 years, or 10,302 observations.

Given that there are 10,302 daily returns in the total holdout sample, we would expect 51, 103, 154, 206, 257, and 515 observations to fall respectively into the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent left and right tails. These numbers along with the number and percentage of returns falling into the aforementioned left and right tails of the total holdout sample are then used to compute the LR statistics for evaluating the out-of-sample performance of unconditional distribution of returns. The results are presented in Table 4.

<sup>16</sup> The null hypothesis that exceedances are i.i.d. is not a prediction of the model. The first-order serial correlation observed in the large falls and rises of stock market returns is a stylized fact and it is our motivation to introduce a conditional VaR model that takes into account serial correlation in the conditional mean and volatility of stock returns.

<sup>17</sup> We decide not to present the  $LR_{ind}$  and  $LR_{cc}$  statistics since they are found to be significant at the 5 percent or 1 percent level without any exception. They are available upon request.

**TABLE 4**  
Out-of-Sample Performance of Unconditional Distribution Functions for the S&P 500

| Maxima       | Expected | Box-Cox-GEV              | GED                      | GEV                       | Normal                     | SGT                       | SGED                      | EGB2                      | IHS                       |
|--------------|----------|--------------------------|--------------------------|---------------------------|----------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| 0.5% VaR     | 51       | 2.6689%<br>71 (6.62)*    | 2.6701%<br>70 (5.99)*    | 2.6633%<br>71 (6.62)*     | 2.2097%<br>148 (120.04)**  | 2.6173%<br>80 (13.45)**   | 2.5427%<br>85 (18.17)**   | 2.5087%<br>90 (23.49)**   | 2.6359%<br>78 (11.74)**   |
| 1% VaR       | 103      | 2.2025%<br>136 (9.59)**  | 2.2110%<br>132 (7.56)**  | 2.1918%<br>138 (10.84)**  | 1.9985%<br>206 (80.26)**   | 2.1791%<br>149 (18.07)**  | 2.1884%<br>145 (15.21)**  | 2.1508%<br>153 (21.16)**  | 2.1932%<br>146 (15.90)**  |
| 1.5% VaR     | 154      | 1.9982%<br>192 (8.57)**  | 1.9897%<br>193 (9.01)**  | 1.9644%<br>201 (12.97)**  | 1.8662%<br>254 (54.17)**   | 1.9403%<br>221 (25.42)**  | 1.9791%<br>207 (16.19)**  | 1.9414%<br>214 (20.57)**  | 1.9501%<br>219 (23.99)**  |
| 2% VaR       | 206      | 1.8493%<br>247 (7.82)**  | 1.8398%<br>249 (8.58)**  | 1.8167%<br>260 (13.33)**  | 1.7677%<br>300 (38.09)**   | 1.7778%<br>280 (24.15)**  | 1.8295%<br>258 (12.23)**  | 1.7927%<br>267 (16.66)**  | 1.7843%<br>280 (24.15)**  |
| 2.5% VaR     | 257      | 1.7402%<br>311 (10.68)** | 1.7317%<br>313 (11.47)** | 1.7192%<br>316 (12.70)**  | 1.6884%<br>338 (23.25)**   | 1.6553%<br>340 (24.37)**  | 1.7129%<br>316 (12.52)**  | 1.6773%<br>330 (18.99)**  | 1.6593%<br>338 (23.25)**  |
| 5% VaR       | 515      | 1.3561%<br>560 (4.01)*   | 1.3504%<br>563 (4.56)*   | 1.3390%<br>568 (5.54)*    | 1.4216%<br>515 (0.00)      | 1.2906%<br>623 (22.01)**  | 1.3462%<br>565 (4.78)*    | 1.3184%<br>593 (11.60)**  | 1.2886%<br>626 (23.22)**  |
| Minima       | Expected | Box-Cox-GEV              | GED                      | GEV                       | Normal                     | SGT                       | SGED                      | EGB2                      | IHS                       |
| 0.5% VaR     | 51       | -2.7035%<br>57 (0.57)    | -2.7099%<br>57 (0.57)    | -2.6881%<br>62 (2.02)     | -2.1514%<br>154 (133.04)** | -2.6441%<br>67 (4.23)*    | -2.5714%<br>79 (12.58)**  | -2.5410%<br>78 (11.74)**  | -2.6872%<br>62 (1.98)     |
| 1% VaR       | 103      | -2.2301%<br>118 (2.10)   | -2.2387%<br>117 (1.84)   | -2.2177%<br>125 (4.44)*   | -1.9402%<br>221 (102.39)** | -2.1845%<br>133 (7.98)**  | -2.2002%<br>134 (8.05)**  | -2.1659%<br>135 (9.04)**  | -2.2164%<br>125 (4.37)*   |
| 1.5% VaR     | 154      | -2.0012%<br>188 (6.89)** | -2.0131%<br>186 (6.11)*  | -1.9824%<br>198 (11.41)** | -1.8080%<br>274 (75.96)**  | -1.9338%<br>204 (14.46)** | -1.9810%<br>198 (11.27)** | -1.9463%<br>201 (12.82)** | -1.9576%<br>196 (10.29)** |
| 2% VaR       | 206      | -1.8422%<br>239 (5.12)*  | -1.8395%<br>240 (5.43)*  | -1.8124%<br>250 (8.97)**  | -1.7095%<br>324 (58.42)**  | -1.7632%<br>266 (16.13)** | -1.8244%<br>249 (8.44)**  | -1.7904%<br>255 (10.89)** | -1.7813%<br>257 (11.77)** |
| 2.5% VaR     | 257      | -1.7333%<br>288 (3.56)   | -1.7300%<br>290 (4.03)*  | -1.7002%<br>311 (10.68)** | -1.6301%<br>364 (39.74)**  | -1.6347%<br>337 (22.69)** | -1.7022%<br>311 (10.51)** | -1.6693%<br>323 (15.60)** | -1.6484%<br>329 (18.49)** |
| 5% VaR       | 515      | -1.3341%<br>572 (6.40)*  | -1.3222%<br>575 (7.08)** | -1.3124%<br>580 (8.29)**  | -1.3634%<br>551 (2.46)     | -1.2527%<br>639 (28.83)** | -1.3186%<br>580 (8.07)**  | -1.2927%<br>603 (14.72)** | -1.2552%<br>637 (27.93)** |
| Average MA%E | -        | 19.43%                   | 19.15%                   | 24.14%                    | 77.72%                     | 34.29%                    | 30.65%                    | 34.85%                    | 31.05%                    |

Notes: This table presents statistics on the out-of-sample performance of VaR thresholds based on unconditional distribution functions for the S&P 500. The statistics are based on the total out-of-sample period January 4, 1960 to December 29, 2000 (10,302 observations). The column "Expected" gives the number of observations expected to fall in the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent left (minima) and right (maxima) tails of the return distribution. The top numbers at each row give the *average* estimated VaR thresholds for each distribution. The numbers below the thresholds give the *average* actual counts along with their respective LR statistics (in parentheses). Average MA percentE is the average mean absolute percentage error based on the actual and expected counts. \*, \*\* denote significance at the 5 percent and 1 percent levels, respectively.

Note that the top numbers at each row of Table 4 are the average estimated VaR thresholds. The next level of numbers at each row give the average estimated counts and the corresponding LR statistics (in parentheses). Interestingly, all distributions exhibit an inferior out-of-sample performance. With a few exceptions, all distributions are rejected for most of the VaR tails. The exceptions are, for the minimal returns, the 0.5 percent, 1.5 percent, and 2.5 percent VaR estimates of Box-Cox-GEV, the 0.5 percent and 1 percent VaR estimates of GPD, the 0.5 percent VaR estimate of GEV, the 5 percent VaR estimate of the normal, and the 0.5 percent VaR estimate of IHS distributions are found to accurate at the 5 percent level of significance. Moreover, the out-of-sample VaR estimates have considerably high MA percentEs. Specifically, the MA percentEs are in the range of 19 percent to 24 percent for the extreme value distributions, 31 percent to 35 percent for the flexible distributions, and about 78 percent for the normal distribution.

### A CONDITIONAL VaR APPROACH WITH FLEXIBLE DISTRIBUTIONS

The results presented in the previous section imply that the actual thresholds are time varying. As a consequence, the out-of-sample performance of the unconditional VaR models is poor. This section extends the analysis using an AR(1) ABS-GARCH(1,1) to model the dynamic behavior of returns for the purpose of computing time-varying VaR thresholds. The model is as follows:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + z_t \sigma_t = \mu_t + z_t \sigma_t \quad (27)$$

$$\sigma_t = \beta_0 + \beta_1 |z_{t-1}| \sigma_{t-1} + \beta_2 \sigma_{t-1}, \quad (28)$$

where  $R_t$  is the return for period  $t$ ,  $\mu_t = \alpha_0 + \alpha_1 R_{t-1}$  and  $\sigma_t$  are the conditional mean and conditional standard deviation of returns based on the information set  $I_{t-1}$ . The conditional threshold for the returns is computed by

$$\vartheta_t = \mu_t + \alpha \sigma_t, \quad (29)$$

where  $\alpha = F^{-1}(\rho)$  is a standardized threshold corresponding to a coverage probability  $\rho$ ; see the section “VaR Models With Alternative Distribution Functions” for additional discussion on the derivation of  $\vartheta_t$  and computation of  $\alpha$ .

Table 5 presents the in-sample performance of alternative distribution functions for estimating time-varying VaR thresholds (i.e., Equation (29)) based on the distribution of standardized returns. These VaR measures, which directly relate to the conditional distribution of returns, indicate that the two tails of the distribution are asymmetric. This is because the estimated thresholds for the extreme negative returns are greater than those for the extreme positive returns. According to the conditional coverage test results, all the extreme value and flexible distributions (except for the SGED) perform very well in estimating VaR, thus they are expected to provide good predictions of catastrophic market risks for all quantiles considered in the article. The SGT, EGB2, and IHS distributions produce very similar VaR thresholds for the standardized returns and generally perform better than the SGED distribution. The normal distribution is strongly rejected for the 0.5 percent, 1 percent, 1.5 percent, and 2 percent left tails (minimal returns) and for the 0.5 percent and 5 percent right tails (maximal returns). The

**TABLE 5**  
In-Sample Performance of Conditional Distribution Functions for Time-Varying VaR Thresholds

| Maxima       | Expected | Box-Cox-GEV            | GED                    | GEV                    | Normal                    | SGT                    | SGED                      | EGB2                    | IHS                    |
|--------------|----------|------------------------|------------------------|------------------------|---------------------------|------------------------|---------------------------|-------------------------|------------------------|
| 0.5% VaR     | 64       | 2.7181%<br>64 (0.00)   | 2.7273%<br>65 (0.01)   | 2.6745%<br>67 (0.12)   | 2.5603%<br>89 (8.63)**    | 2.8787%<br>51 (2.91)   | 2.8207%<br>55 (1.38)      | 2.8136%<br>55 (1.38)    | 2.8479%<br>54 (1.71)   |
| 1% VaR       | 128      | 2.3210%<br>136 (0.46)  | 2.3170%<br>138 (0.72)  | 2.2994%<br>144 (1.86)  | 2.3093%<br>144 (1.87)     | 2.4522%<br>111 (2.47)  | 2.4644%<br>110 (2.77)     | 2.4328%<br>116 (1.23)   | 2.4398%<br>113 (1.92)  |
| 1.5% VaR     | 192      | 2.1135%<br>200 (0.29)  | 2.1042%<br>203 (0.57)  | 2.1088%<br>201 (0.37)  | 2.1521%<br>197 (0.11)     | 2.2132%<br>176 (1.47)  | 2.2497%<br>161 (5.52)*    | 2.2095%<br>177 (1.29)   | 2.2073%<br>177 (1.29)  |
| 2% VaR       | 257      | 1.9753%<br>270 (0.70)  | 1.9642%<br>278 (1.77)  | 1.9828%<br>265 (0.28)  | 2.0350%<br>247 (0.37)     | 2.0473%<br>243 (0.75)  | 2.0940%<br>220 (5.59)*    | 2.0506%<br>243 (0.75)   | 2.0444%<br>244 (0.64)  |
| 2.5% VaR     | 321      | 1.8723%<br>342 (1.41)  | 1.8616%<br>349 (2.47)  | 1.8883%<br>331 (0.33)  | 1.9407%<br>296 (2.01)     | 1.9203%<br>303 (1.03)  | 1.9713%<br>278 (6.12)*    | 1.9269%<br>301 (1.27)   | 1.9188%<br>303 (1.03)  |
| 5% VaR       | 642      | 1.5661%<br>630 (0.22)  | 1.5729%<br>618 (0.92)  | 1.5727%<br>618 (0.92)  | 1.6237%<br>541 (17.46)**  | 1.5308%<br>630 (0.22)  | 1.5769%<br>579 (6.62)*    | 1.5394%<br>622 (0.63)   | 1.5310%<br>630 (0.22)  |
| Minima       | Expected | Box-Cox-GEV            | GED                    | GEV                    | Normal                    | SGT                    | SGED                      | EGB2                    | IHS                    |
| 0.5% VaR     | 64       | -3.0700%<br>59 (0.28)  | -3.1222%<br>55 (1.38)  | -3.0523%<br>63 (0.02)  | -2.6229%<br>123 (42.71)** | -3.0939%<br>63 (0.02)  | -3.0595%<br>64 (0.00)     | -3.0997%<br>60 (0.28)   | -3.1312%<br>58 (0.61)  |
| 1% VaR       | 128      | -2.5393%<br>137 (0.01) | -2.5538%<br>126 (0.04) | -2.5234%<br>131 (0.06) | -2.3718%<br>180 (18.70)** | -2.6309%<br>120 (0.55) | -2.6723%<br>113 (1.92)    | -2.6709%<br>114 (1.67)  | -2.6695%<br>114 (1.67) |
| 1.5% VaR     | 192      | -2.2848%<br>198 (0.69) | -2.2786%<br>207 (1.09) | -2.2743%<br>209 (1.40) | -2.2147%<br>234 (8.52)**  | -2.3723%<br>180 (0.84) | -2.4391%<br>163 (4.82)*   | -2.4198%<br>167 (3.58)  | -2.4072%<br>169 (3.02) |
| 2% VaR       | 257      | -2.1230%<br>272 (0.02) | -2.1052%<br>266 (0.34) | -2.1175%<br>262 (0.11) | -2.0975%<br>291 (4.51)*   | -2.1929%<br>243 (0.75) | -2.2702%<br>210 (9.20)**  | -2.2414%<br>223 (4.70)* | -2.2240%<br>229 (3.14) |
| 2.5% VaR     | 321      | -2.0057%<br>331 (0.55) | -1.9820%<br>348 (2.30) | -2.0044%<br>336 (0.73) | -2.0033%<br>352 (3.03)    | -2.0556%<br>324 (0.03) | -2.1370%<br>266 (10.17)** | -2.1027%<br>286 (4.01)* | -2.0833%<br>303 (1.03) |
| 5% VaR       | 642      | -1.6490%<br>662 (0.15) | -1.6538%<br>627 (0.35) | -1.6579%<br>623 (0.57) | -1.6863%<br>595 (3.63)    | -1.6354%<br>654 (0.25) | -1.7095%<br>576 (7.28)**  | -1.6695%<br>616 (1.08)  | -1.6511%<br>638 (0.02) |
| Average MA%E | -        | 3.63%                  | 6.12%                  | 4.51%                  | 22.21%                    | 6.43%                  | 12.87%                    | 8.70%                   | 8.09%                  |

Notes: This table presents statistics on the in-sample performance of conditional distribution functions for time-varying VaR thresholds estimated by Equation (31). The results are based on the standardized returns on S&P 500 spanning the period January 4, 1950 to December 29, 2000 (12,832 observations). The column "Expected" gives the number of observations expected to fall in the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent left (minima) and right (maxima) tails of the return distribution. The top numbers at each row give the estimated VaR thresholds for each distribution. The numbers below the thresholds give the actual counts along with their respective LR statistics (in parentheses). Average MA percentE is the average mean absolute percentage error based on the actual and expected counts. \*, \*\* denote significance at the 5 percent and 1 percent levels, respectively.

in-sample VaR estimates of the conditional distributions yield considerably low MA percentEs except for the conditional normal density. Specifically, the MA percentEs are in the range of 3.63 percent to 6.12 percent for the extreme value distributions, 6.43 percent to 12.87 percent for the flexible distributions, and about 22.21 percent for the normal distribution.

To further evaluate the out-of-sample performance of the conditional probability distributions, we use the 10-year rolling-window procedure described previously. Table 6 presents the out-of-sample risk measurement performance of conditional distribution functions for the standardized returns. The results for the time-varying VaR thresholds indicate the superior performance of the extreme value and flexible distributions. Interestingly, the results for the flexible distributions, although qualitatively similar, are much better than those of Table 5. According to the LR statistics from the conditional coverage tests, the extreme value and all flexible distributions produce acceptably accurate out-of-sample VaR estimates for all tails considered in this article. The normal distribution is rejected for most of the VaR tails. The SGT, IHS, and EGB2 distributions yield lower MA percentEs than the Box-Cox-GEV and GPD distributions. Overall, the out-of-sample performance results imply that the flexible distributions (except the SGED) perform better than some of the extreme value distributions.

## CONCLUSION

Risk measurement has been an important area to financial managers, actuaries, and insurance practitioners for a long period of time. VaR type measures are employed in many different types of financial and insurance applications including the determination of capital requirements, capital reserves, the setting of insurance deductibles, the setting of reinsurance cedance levels, as well as the estimation of expected claims and expected losses.

This article develops VaR measures that account for the dynamic behavior of the conditional mean and conditional volatility of portfolio returns over time. The performance of three extreme value distributions, i.e., GPD, GEV and Box-Cox-GEV and four skewed fat-tailed distributions, i.e., skewed-generalized error distribution (SGED), SGT, EGB2, and IHS, in estimating VaR thresholds is evaluated using various test statistics.

The results provide strong evidence that the SGT, EGB2, and IHS distributions perform as well as the more specialized extreme value distributions in modeling the tail behavior of portfolio returns. All three distributions provide produce similar VaR thresholds and perform better than the SGED and the normal distributions in approximating the extreme tails of the return distribution. The conditional coverage and the out-of-sample performance tests show that the actual VaR thresholds are time varying to a degree not captured by unconditional VaR measures.

The fact that the SGT, EGB2, IHS, and Box-Cox-GEV distributions performed well in modeling the extreme tails of the conditional distribution of financial asset returns may have significant implications for the modeling of insurance loss processes and in particular the payout tail in long tail lines. The latter is important in pricing, reserving and reinsurance decision making, solvency testing, and dynamic financial analysis (e.g., Cummings, McDonald and Merrill, 2004). To our knowledge the ability of the

**TABLE 6**  
Out-of-Sample Performance of Conditional Distribution Functions for Time-Varying VaR Thresholds

| Maxima       | Expected | Box-Cox-GEV            | GED                    | GEV                    | Normal                    | SGT                    | SGED                   | EGB2                   | IHS                    |
|--------------|----------|------------------------|------------------------|------------------------|---------------------------|------------------------|------------------------|------------------------|------------------------|
| 0.5% VaR     | 51       | 2.7966%<br>53 (0.04)   | 2.8013%<br>52 (0.00)   | 2.8310%<br>49 (0.13)   | 2.5242%<br>76 (10.12)**   | 2.8194%<br>50 (0.05)   | 2.7747%<br>56 (0.37)   | 2.7681%<br>49 (0.13)   | 2.8011%<br>48 (0.26)   |
| 1% VaR       | 103      | 2.4199%<br>98 (0.25)   | 2.4105%<br>99 (0.16)   | 2.3903%<br>101 (0.04)  | 2.2762%<br>133 (7.98)**   | 2.4110%<br>100 (0.10)  | 2.4281%<br>96 (0.52)   | 2.4009%<br>102 (0.01)  | 2.4060%<br>101 (0.05)  |
| 1.5% VaR     | 154      | 2.1993%<br>156 (0.01)  | 2.2011%<br>155 (0.00)  | 2.1877%<br>157 (0.04)  | 2.1209%<br>177 (3.10)     | 2.1807%<br>159 (0.12)  | 2.2188%<br>153 (0.02)  | 2.1843%<br>157 (0.03)  | 2.1798%<br>157 (0.03)  |
| 2% VaR       | 206      | 2.0478%<br>199 (0.25)  | 2.0500%<br>198 (0.32)  | 2.0144%<br>206 (0.00)  | 2.0053%<br>211 (0.11)     | 2.0202%<br>207 (0.00)  | 2.0670%<br>192 (1.04)  | 2.0296%<br>204 (0.03)  | 2.0209%<br>209 (0.03)  |
| 2.5% VaR     | 257      | 1.9288%<br>250 (0.23)  | 1.9401%<br>246 (0.54)  | 1.9062%<br>256 (0.01)  | 1.9121%<br>256 (0.02)     | 1.8969%<br>262 (0.06)  | 1.9471%<br>238 (1.62)  | 1.9089%<br>255 (0.03)  | 1.8983%<br>258 (0.00)  |
| 5% VaR       | 515      | 1.5401%<br>525 (0.20)  | 1.5500%<br>521 (0.07)  | 1.5148%<br>542 (1.46)  | 1.5990%<br>472 (4.04)*    | 1.5162%<br>546 (1.82)  | 1.5610%<br>504 (0.29)  | 1.5281%<br>538 (0.98)  | 1.5175%<br>547 (1.94)  |
| Minima       | Expected | Box-Cox-GEV            | GED                    | GEV                    | Normal                    | SGT                    | SGED                   | EGB2                   | IHS                    |
| 0.5% VaR     | 51       | -2.9865%<br>56 (0.38)  | -3.0001%<br>55 (0.24)  | -3.0285%<br>50 (0.05)  | -2.5961%<br>101 (37.12)** | -3.0122%<br>53 (0.04)  | -2.9677%<br>59 (1.02)  | -2.9937%<br>52 (0.00)  | -3.0363%<br>48 (0.26)  |
| 1% VaR       | 103      | -2.5801%<br>104 (0.01) | -2.5892%<br>103 (0.00) | -2.5888%<br>103 (0.00) | -2.3481%<br>145 (15.21)** | -2.5746%<br>104 (0.01) | -2.5995%<br>104 (0.01) | -2.5944%<br>103 (0.00) | -2.6009%<br>103 (0.00) |
| 1.5% VaR     | 154      | -2.3545%<br>144 (0.75) | -2.3551%<br>144 (0.75) | -2.3421%<br>146 (0.49) | -2.1929%<br>195 (9.81)**  | -2.3282%<br>147 (0.40) | -2.3774%<br>141 (1.28) | -2.3592%<br>147 (0.40) | -2.3525%<br>146 (0.51) |
| 2% VaR       | 206      | -2.1843%<br>191 (1.15) | -2.2066%<br>186 (2.06) | -2.1792%<br>192 (0.99) | -2.0772%<br>246 (7.32)*   | -2.1567%<br>202 (0.09) | -2.2164%<br>180 (3.59) | -2.1913%<br>191 (1.20) | -2.1784%<br>194 (0.77) |
| 2.5% VaR     | 257      | -2.0541%<br>252 (0.12) | -2.0792%<br>243 (0.86) | -2.0600%<br>250 (0.23) | -1.9841%<br>291 (4.17)    | -2.0251%<br>267 (0.32) | -2.0893%<br>237 (1.79) | -2.0604%<br>248 (0.40) | -2.0445%<br>253 (0.10) |
| 5% VaR       | 515      | -1.6523%<br>517 (0.01) | -1.6715%<br>507 (0.13) | -1.6302%<br>532 (0.58) | -1.6709%<br>499 (0.59)    | -1.6201%<br>546 (1.82) | -1.6804%<br>488 (1.62) | -1.6483%<br>526 (0.21) | -1.6309%<br>543 (1.47) |
| Average MA%E | —        | 3.75%                  | 3.91%                  | 2.78%                  | 25.45%                    | 3.16%                  | 7.03%                  | 2.71%                  | 3.48%                  |

Notes: This table presents statistics on the out-of-sample performance of conditional distribution functions for time-varying VaR thresholds estimated by Equation (31). The results are based on the standardized returns on the S&P 500 and the total out-of-sample period January 4, 1960 to December 29, 2000 (10,302 observations). The column "Expected" gives the number of observations that fall in the 0.5 percent, 1 percent, 1.5 percent, 2 percent, 2.5 percent, and 5 percent left (minima) and right (maxima) tails of the return distribution. The top numbers at each row give the *average* estimated VaR thresholds for each distribution. The numbers below the thresholds give the *average* actual counts along with their respective LR statistics (in parentheses). Average MA percentE is the average mean absolute percentage error based on the actual and expected counts. \*, \*\* denote significance at the 5 percent and 1 percent levels, respectively.

aforementioned distributions to estimate the severity distributions of insurance claims has not yet been investigated.

#### APPENDIX: MAXIMUM LIKELIHOOD ESTIMATION OF THE BOX-COX-GEV DISTRIBUTION

We examine the parameter estimates of  $F_{\max, \xi}(x)$  to evaluate which of the two commonly used extreme value distributions is closer to the truth. The results indicate that Box-Cox-GEV distribution provides a more accurate characterization of the actual data than the GPD and GEV distributions because the truth is somewhere in between. Two parametric approaches are commonly used to estimate the extreme value distributions: the maximum likelihood method that yields parameter estimators that are unbiased, asymptotically normal, and of minimum variance; and the regression method that provides a graphical method for determining the type of asymptotic distribution. In this article the maximum likelihood method is used to determine the correct specification of the limit distribution for the maxima and minima.

Suppose that the Box-Cox-GEV distribution for the maxima,  $F_{\max}(\Theta; x)$ , has a density function  $f_{\max}(\Theta; x)$ , where  $\Theta = (\xi, \mu, \sigma) \in \Re \times \Re \times \Re_+$  consists of a shape parameter  $\xi$ , a location parameter  $\mu$  and a scale parameter  $\sigma$ . Then the likelihood function based on the data with a time-series of length  $k$  for the maximum variable  $M_k = (M_1, M_2, \dots, M_k)$  is given by

$$L(\Theta; M_k) = \prod_{i=1}^k f_{\max}(\Theta; M_i). \quad (\text{A1})$$

Denote the log-likelihood function by  $l(\Theta; M_k) = \ln L(\Theta; M_k)$ . The maximum likelihood estimator (MLE) for  $\Theta$  then equals

$$\hat{\Theta}_k = \arg \max_{\Theta \in \Psi} l(\Theta; M_k), \quad (\text{A2})$$

where  $\hat{\Theta}_k = \hat{\Theta}_k(M_1, M_2, \dots, M_k)$  maximizes  $l(\Theta; M_k)$  over an appropriate parameter space  $\Psi$ .

The Box-Cox-GEV distribution presented in Equation (16) has a density function,

$$f_{\max}(\Theta; x) = \left(\frac{1}{\sigma}\right) \left[1 + \xi \left(\frac{M - \mu}{\sigma}\right)\right]^{-\left(\frac{1+\xi}{\xi}\right)} \left\{ \exp \left[ - \left(1 + \xi \left(\frac{M - \mu}{\sigma}\right)\right) \right]^{-1/\xi} \right\}^{\phi}, \quad (\text{A3})$$

which yields the following log-likelihood function:

$$\begin{aligned} \ln L_f = & -k \ln \sigma - k \left(\frac{1+\xi}{\xi}\right) \sum_{i=1}^k \ln \left(1 + \xi \left(\frac{M_i - \mu}{\sigma}\right)\right) \\ & - k \phi \sum_{i=1}^k \left(1 + \xi \left(\frac{M_i - \mu}{\sigma}\right)\right)^{-1/\xi}. \end{aligned} \quad (\text{A4})$$

Differentiating the log-likelihood function in Equation (A4) with respect to  $\mu, \sigma, \xi$ , and  $\phi$  yields the first-order conditions of the maximization problem. Clearly, no explicit

solution exists to these nonlinear equations, and thus numerical procedures or search algorithms are called for. In some cases, we had a difficulty in obtaining a "global maximum" of the log-likelihood function. Hence, the parameters are estimated by scanning the possible range in increments of 0.10. When a minimum of the sum of squares is found, greater precision is desired, and the area to the right and left of the current optimum is searched in increments of 0.01.

## REFERENCES

- Aiuppa, T. A., 1988, Evaluation of Pearson Curves as an Approximation of the Maximum Probably Annual Aggregate Loss, *Journal of Risk and Insurance*, 5: 425-441.
- Bali, T. G., 2003a, An Extreme Value Approach to Estimating Volatility and Value at Risk, *Journal of Business*, 76: 83-108.
- Bali, T. G., 2003b, The Generalized Extreme Value Distribution, *Economics Letters*, 79: 423-427.
- Box, G., and D. Cox, 1964, An Analysis of Transformations, *Journal of the Royal Statistical Society*, Series B, 211-264.
- Box, G., and G. C. Tiao, 1962, A Further Look at Robustness via Bayes Theorem, *Biometrika*, 49: 419-432.
- Cho, D., 1988, The Impact of Risk Management Decisions on Firm Value: Gordon's Growth Model Approach, *Journal of Risk and Insurance*, 55: 118-131.
- Christoffersen, P. F., 1998, Evaluating Interval Forecasts, *International Economic Review*, 39: 841-862.
- Christoffersen, P. F., and F. X. Diebold, 2000, How Relevant Is Volatility Forecasting for Financial Risk Management, *Review of Economics and Statistics*, 82: 12-22.
- Cummins, D. J., 1976, Risk Management and the Theory of the Firm, *Journal of Risk and Insurance*, 43: 587-609.
- Cummins, D. J., C. M. Lewis, and R. D. Phillips, 1999, Pricing Excess of Loss Reinsurance Contracts Against Catastrophic Loss, in: Kenneth Froot, ed., *The Financing of Catastrophe Risk* (Chicago: University of Chicago Press).
- Cummins, D. J., G. Dionne, J. B. McDonald, and B. M. Pritchett, 1990, Applications of the GB2 Family of Distributions in Modeling Insurance Loss Processes, *Insurance: Mathematics and Economics*, 9: 257-272.
- Cummins, D. J., J. B. McDonald, and C. Merrill, 2004, Risky Loss Distributions and Modeling the Loss Reserve Pay-Out Tail, Working Paper, University of Pennsylvania.
- Dowd, K., and D. Blake, 2006, After VaR: The Theory, Estimation, and Insurance Applications of Quantile-Based Risk Measures. *Journal of Risk and Insurance*, 73(2): 193-229.
- Embrechts, P., C. Kluppelberg, and T. Mikosch, 1997, *Modeling Extremal Events* (Berlin: Springer).
- Gnedenko, B. V., 1943, Sur la distribution limite du terme maximum d'une serie aleatoire, *Annals of Mathematics*, 44: 423-453.
- Hansen, B. E., 1994, Autoregressive Conditional Density Estimation, *International Economic Review*, 35: 705-730.

- Hansen, C. B., J. B. McDonald, and P. Theodossiou, 2005, Some Flexible Parametric Models and Leptokurtic Data, Working Paper, Brigham Young University.
- Hogg, R., and S. Klugman, 1984, *Loss Distributions* (New York: John Wiley & Sons).
- Jenkinson, A. F., 1955, The Frequency Distribution of the Annual Maximum (or Minimum) Values of Meteorological Elements, *Quarterly Journal of the Royal Meteorology Society*, 87: 145-158.
- Johnson, N. L., 1949, Systems of Frequency Curves Generated by Methods of Translation, *Biometrika*, 36: 149-176.
- Johnson, N. L., S. Kotz, and N. Balakrishnan, 1994, *Continuous Univariate Distributions*, Volume 1, 2nd edition (New York: John Wiley & Sons).
- Klugman, S. A., H. H. Panjer, and G. E. Willmot, 1998, *Loss Models: From Data to Decisions* (New York: John Wiley & Sons).
- Kupiec, P. H., 1995, Techniques for Verifying the Accuracy of Risk Measurement Models, *Journal of Derivatives*, 3: 73-84.
- Lilliefors, H., 1967, On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown, *Journal of the American Statistical Association*, 62: 399-402.
- Longin, F. M., 2000, From Value at Risk to Stress Testing: The Extreme Value Approach, *Journal of Banking and Finance*, 24: 1097-1130.
- McDonald, J. B., and W. K. Newey, 1988, Partially Adaptive Estimation of Regression Models via the Generalized t Distribution, *Econometric Theory*, 4: 428-457.
- McDonald, J. B., and Y. J. Xu, 1995, A Generalization of the Beta Distribution with Applications, *Journal of Econometrics*, 66: 133-152. Errata 69(1995), 427-428.
- McNeil, A. J., and R. Frey, 2000, Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach, *Journal of Empirical Finance*, 7: 271-300.
- Nelson, D., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59: 347-370.
- Paulson, A. S., and N. J. Faris, 1985, A Practical Approach to Measuring the Distribution of Total Annual Claims, in: D. J. Cummins, ed., *Strategic Planning and Modeling in Property-Liability Insurance* (Norwell, MA: Kluwer Academic Publishers).
- Pickands, J., 1975, Statistical Inference Using Extreme Order Statistics, *Annals of Statistics*, 3: 119-131.
- Ramlau-Hansen, H., 1988, A Solvency Study in Non-Life Insurance. Part 1. Analysis of Fire, Windstorm, and Glass Claims, *Scandinavian Actuarial Journal*, 80: 3-34.
- Santomero, A. M., and D. F. Babbel, 1997, Financial Risk Management by Insurers: An Analysis of the Process, *Journal of Risk and Insurance*, 64: 231-270.
- Sprecher, C. R., and M. A. Pertl, 1983, Large Losses, Risk Management and Stock Prices, *Journal of Risk and Insurance*, 50: 107-117.
- Subbotin, M. T. H., 1923, On the Law of Frequency of Error, *Matematicheskii Sbornik*, 31: 296-301.
- Theodossiou, P., 1998, Financial Data and the Skewed Generalized t Distribution, *Management Science*, 44: 1650-1661.
- Theodossiou, P., 2004, Skewness and Kurtosis in Financial Data and the Pricing of Options, Working Paper, Rutgers University.

Copyright of Journal of Risk & Insurance is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.