An Extreme Value Approach to Estimating Interest-Rate Volatility: Pricing Implications for Interest-Rate Options

Turan G. Bali
Department of Economics and Finance, Zicklin School of Business, Baruch College, City University of New York, One Bernard Baruch Way, Box 10-225, New York, New York 10010 and Department of Finance, College of Administrative Sciences and Economics, Koç University, Fener Yolu Caddesi, Sarıyer 80910, Istanbul, Turkey, turan_bali@baruch.cuny.edu

This paper proposes an extreme value approach to estimating interest-rate volatility, and shows that during the extreme movements of the U.S. Treasury market the volatility of interest-rate changes is underestimated by the standard approach that uses the thin-tailed normal distribution. The empirical results indicate that (1) the volatility of maximal and minimal changes in interest rates declines as time-to-maturity rises, yielding a downward-sloping volatility curve for the extremes; (2) the minimal changes are more volatile than the maximal changes for all data sets and for all asymptotic distributions used; (3) the minimal changes in Treasury yields have fatter tails than the maximal changes; and (4) for both the maxima and minima, the extreme changes in short-term rates have thicker tails than the extreme changes in long-term rates. This paper extends the standard option-pricing models with lognormal forward rates to accommodate significant kurtosis observed in the interest-rate data. This paper introduces a closed-form option-pricing model based on the generalized extreme value distribution that successfully removes the well-known pricing bias of the lognormal distribution.

Key words: extreme value distributions; interest-rate options; term structure of interest rates; volatility; skewed fat-tailed distributions

History: Accepted by David A. Hsieh, finance; received May 7, 2005. This paper was with the author 4 months and 1 week for 2 revisions.

1. Introduction

The default-free short-term interest rate driving the changes in the entire term structure is fundamental to the classical approach to pricing fixed-income securities and derivatives. Both equilibrium and arbitrage-free term structure models use the short rate and its estimated volatility as inputs to determine the value of interest-rate-contingent claims. This makes the accurate estimation of the level and volatility of interest-rate changes crucial to pricing interest-rate derivatives. For example, to value interest rate-sensitive claims, one needs to model the instantaneous and time-series properties of interest-rate volatility. This follows because both the current level and the stochastic properties of volatility will affect the distribution of future interest-rate levels, which determine a derivative’s price. The interested reader may wish to consult the online supplement (provided in the e-companion)1 for a more detailed discussion on interest-rate models.

A major issue in estimating the term structure of interest-rate volatility is the modeling of the distribution of interest rates. In practice, the empirical distribution is difficult to handle, and a parametric model is often preferred. However, because there is no theoretical model for the exact distribution of interest-rate changes, an assumption has to be made. The methods based on a parametric distribution often assume that interest rates are normally (or lognormally) distributed. However, this paper provides strong evidence that the levels and changes in short-term and long-term interest rates are not normally distributed. The tails of the empirical distribution are found to be thicker than the tails of the normal distribution, which means that large interest-rate changes actually occur more frequently than predicted by the unconditional normal model. The method proposed here is based on the extreme value theory (EVT), and thus provides accurate predictions of catastrophic market risks during highly volatile periods.

The existing literature indicates that using a better representation of the yield and volatility curves can enhance the performance of arbitrage-free term structure models. Therefore, the choice of a volatility estimator for short-term and long-term interest rates is crucial to pricing interest-rate-dependent securities. Previous research on pricing interest-rate-contingent claims employs historical, moving average,

---

1 An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.
pricing models also implies that log-spot or forward motion assumption utilized by standard options-excess kurtosis of asset returns. The geometric Brownian motion cannot explicitly account for the skewness and the same maturity date. Moreover, this class of models have well-known drawbacks. First, the constancy required in the lognormal case implied volatility smiles or smirks are inconsistent with the thin-tailed normal distribution. The empirical results indicate that (1) the volatility of maximal and minimal changes in interest rates declines as time to maturity rises, yielding a downward-sloping volatility curve for the extremes; (2) the minimal changes are more realistic than the maximal changes for all data sets and for all asymptotic distributions used; (3) the minimal changes in Treasury yields have fatter tails than the maximal changes; (4) for both the maxima and minima, the extreme changes in short-term rates have thicker tails than the extreme changes in long-term rates. These results have important implications for pricing interest-rate derivatives.

Closed-form solutions for interest-rate derivatives—in particular, caps, floors, and bond options—have been obtained by a number of authors for Markovian term structure models with normally distributed interest rates or, alternatively, lognormally distributed bond prices (see, e.g., Jamshidian 1989, Heath et al. 1992, Miltersen et al. 1997). These models support the Black-Scholes-type (1973) formulas most frequently used by practitioners for pricing bond options and swaptions.

The Black-Scholes and lognormal distribution-based models have well-known drawbacks. First, the implied volatility smiles or smirks are inconsistent with the constancy required in the lognormal case for volatility across different strikes for options with the same maturity date. Moreover, this class of models cannot explicitly account for the skewness and excess kurtosis of asset returns. The geometric Brownian motion assumption utilized by standard options-pricing models also implies that log-spot or forward interest rates are identically and independently distributed (i.i.d.) normal variables, and thus exhibit no moment dependencies, such as asymmetric and conditional volatility with skewed and/or fat-tailed distributions. However, our empirical results from alternative specifications of the interest-rate process indicate the presence of significant volatility clustering and leptokurtosis in spot and forward interest-rate changes. To accommodate (especially) the tail thickness of the interest-rate distribution, we derive a closed-form option-pricing model with the generalized extreme value (GEV) distribution.

It is well known that the option prices are sensitive to the tail-shape change, which is a matter that is distinct from its sensitivity to the variance of the return distribution. However, the lack of closed-form solutions to the option-pricing model, the large number of parameters needed, or the lack of easy implementation of implied parameters have prevented many of the proposed models intended to deal with both the fat tails and the skew in asset prices from being of practical use in pricing and hedging.

This paper develops a closed-form option-pricing model based on the GEV distribution with only three parameters defining the location, scale, and tail shape. It provides a flexible framework that encompasses as special cases a number of classes of distributions that have been assumed to date in more restrictive settings. The GEV option-pricing model successfully removes the well-known pricing biases associated with the lognormal distribution by introducing flexibility in terms of changes in tail shape. Although the newly proposed model can be used to price options written on individual stocks, stock indices, currency, and futures contracts, in this study, we price interest-rate options traded on the Chicago Board Options Exchange (CBOE). The contracts that we study are European-style options on the 13-week T-bill discount rate that have the ticker symbol IRX. The IRX options are equivalent to European-type options written on zero-coupon bonds. Hence, we are dealing with the simplest contracts imaginable.

The empirical results indicate that the GEV option-pricing model provides more accurate predictions of interest-rate option prices than the lognormal and the fat-tailed generalized error distribution (GED). The relative performance of the EVT-based option-pricing model becomes much superior to the lognormal and GED distributions during extremely volatile periods.

This paper is organized as follows. Section 2 presents the extreme-value and skewed fat-tailed distributions. Section 3 describes the constant-maturities U.S. Treasury yields data. Section 4 presents the maximum-likelihood estimation results. Section 5 shows the EVT-based volatility formula and summarizes
results from the GEV and normal distributions. Section 6 provides an option-pricing model based on the GEV, GED, and lognormal distributions. Section 7 evaluates the option-pricing performance of alternative distribution functions. Section 8 concludes the paper.

2. Distributions

2.1. Extreme Value Distributions

Extreme interest-rate movements are measured by the daily changes in Treasury yields: \( r_t = R_t - R_{t-1} \), where \( R_t \) and \( r_t \) are the level and change in interest rates at time \( t \), respectively. Let us call \( f(r) \) the probability density function and \( F(r) \) the cumulative distribution function of \( r \). Let \( r_1, r_2, \ldots, r_n \) be a sequence of interest rate changes on days \( 1, 2, \ldots, n \). Extremes are defined as the maxima and minima of the \( n \) i.i.d. random variables \( r_1, r_2, \ldots, r_n \). Let \( M_n \) represent the highest daily interest-rate changes (maximum) and \( m_n \) denote the lowest daily interest-rate changes (minimum) over \( n \) trading days:

\[
M_n = \max(r_1, r_2, \ldots, r_n), \quad (1)
\]

\[
m_n = \min(r_1, r_2, \ldots, r_n)
\]

\[
= \max(-r_1, -r_2, \ldots, -r_n). \quad (2)
\]

As shown in Gumbel (1958), if the variables \( r_1, r_2, \ldots, r_n \) are statistically independent and drawn from the same distribution, then the exact distribution of the maximum and minimum can be written as a function of the parent distribution \( F(R) \) and the length of selection period \( n \): The exact distribution of \( M_n \) is \( H_{\max,n}(M_n) = [F(R)]^n \) and the exact distribution of \( m_n \) equals \( H_{\min,n}(m_n) = 1 - [1 - F(R)]^n \).

To find a limit distribution for \( M_n \), it is transformed such that the limit distribution of the new variable is a nondegenerate one. Following the Fisher-Tippett theorem (Fisher and Tippett 1928), the variate, \( M_n \), is reduced with a location parameter, \( \mu_n \), and a scale parameter, \( \sigma_n \), in such a way that \( x = (M_n - \mu_n)/\sigma_n \rightarrow H_{\max}(x) \). Assuming the existence of a sequence of such coefficients \( (\mu_n, \sigma_n > 0) \), three types of nondegenerated distributions are obtained for the standardized maximum:

- Frechet: \( H_{\max,\xi}(x) = \exp(-x^{-1/\xi}), \quad (3) \)
- Weibull: \( H_{\max,\xi}(x) = \exp(-(-x)^{-1/\xi}), \quad (4) \)
- Gumbel: \( H_{\max,\xi}(x) = \exp[-\exp(-x)]. \quad (5) \)

Jenkinson (1955) proposes a GEV distribution, which includes the three limit distributions in (3)–(5), distinguished by Gnedenko (1943):

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

\[
1 + \frac{\xi \left( M - \mu \right)}{\sigma} \geq 0, \quad \xi > 0, \quad \xi < 0, \quad \xi = 0 \quad \text{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; and the Gumbel distribution is thin tailed because its tail is rapidly decreasing. The shape parameter \( \xi \), called the tail index, 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.}

One potential problem with Equations (3)–(6) is that the daily interest-rate changes are not i.i.d. As explained by Leadbetter et al. (1983), Resnick (1987), and Castillo (1988), when the data are dependent, it is possible that extreme value distributions cannot be described as in (3)–(6). Not surprisingly, no general statement can be made without further assumptions on the exact nature of the dependence structure. However, the theory of extremes for the case of dependence, while not completely developed, has identified a number of empirically relevant cases for which inferences based on standard extreme value theory remain valid. For example, this is true if the daily interest-rate changes are stationary, and follow an MA(q), AR(p), or ARMA(p, q) model (see Leadbetter et al. 1983, Resnick 1987, and Castillo 1988 for detailed treatments). Because these conditions describe our data reasonably well, we prefer to stick to the extreme value distributions described in (3)–(6) as opposed to making assumptions on the dependence structure that would lead to other forms of extreme value distributions.

To make sure that our results are not driven by a misspecified extreme value distribution, we follow Diebold et al. (1998) and conduct the following robustness check: We first fit a conditional mean-volatility model to the raw daily interest-rate changes, standardize the data by the estimated conditional mean and volatility, and then repeat our analysis based on the standardized residuals.

Two parametric approaches are commonly used to estimate the extreme value distributions: (1) the maximum-likelihood method, which yields parameter estimators that are unbiased, asymptotically normal, and of minimum variance; and (2) the regression method, which provides a graphical method for determining the type of asymptotic distribution. In this paper, the maximum-likelihood method is used to estimate interest-rate volatility in extreme values.

The GEV distribution in Equation (6) has a density function for the maxima,

\[
h_{\max}(\Phi; x) = \frac{1}{\sigma} \left[ 1 + \frac{\xi \left( M - \mu \right)}{\sigma} \right]^{(1+\xi)/\xi}
\]
\[ \delta = 2\lambda AS(\lambda)^{-1}, \quad S(\lambda) = \sqrt{1 + 3\lambda^2 - 4\lambda^2\lambda^2}, \]
\[ A = \Gamma\left(\frac{2}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right) \Gamma\left(\frac{3}{\nu}\right)^{-1/2}, \]
where \( \mu = E(r) \) and \( \sigma \) are the expected value and the standard deviation of \( r \), \( \lambda \) is a skewness parameter, “sign” is the sign function, and \( \Gamma(\cdot) \) is the gamma function. The scaling parameters \( \nu \) and \( \lambda \) obey the following constraints: \( \nu > 0 \) and \( -1 < \lambda < 1 \). The parameter \( \nu \) controls the height and tails of the density function, and the skewness parameter \( \lambda \) controls the rate of descent of the density around the mode of \( r \), where \( \text{mode}(r) = \mu - \delta \sigma \). In the case of positive skewness (\( \lambda > 0 \)), the density function is skewed to the right. This is because for values of \( r < \mu - \delta \sigma \), the interest-rate change \( r \) is weighted by a greater value than unity and for values of \( r > \mu - \delta \sigma \) by a value less than unity. The opposite is true for negative \( \lambda \). Note that \( \lambda \) and \( \delta \) have the same sign, thus, in the case of positive skewness (\( \lambda > 0 \)), the mode of \( r \) is less than the expected value of \( r \). The parameter \( \delta \) is Pearson’s skewness, \( [\mu - \text{mode}(r)]/\sigma = 3 \).

The SGED parameters are estimated by maximizing the log-likelihood function of \( r_i \) with respect to the parameters \( \mu, \sigma, \nu, \) and \( \lambda \):
\[ \log L = n \log C - n \log \sigma - \frac{1}{\theta^2 \sigma^2} \sum_{i=1}^{n} \left( \frac{|r_i - \mu + \delta \sigma|}{1 + \text{sign}(r_i - \mu + \delta \sigma)\lambda^\nu} \right)^\nu, \]
where “sign” is the sign of the residuals \( (r_i - \mu + \delta \sigma) \), and \( n \) is the sample size.

### 2.2. Skewed Fat-Tailed Distributions

#### 2.2.1. Skewed Generalized Error Distribution

Subbotin (1923) introduces the generalized error distribution (GED) as special cases of the Laplace, normal, and uniform distributions. The symmetric GED density is given by
\[ f(r; \mu, \sigma, \nu) = \frac{v \exp[-(1/2)|z_i|/\nu]}{\Pi^{(v+1)/2} \Gamma(1/\nu)}, \]
where \( r_i = R_i - R_{i-1} \) is the interest-rate change, \( R_i \) is the interest-rate level, \( \Gamma(a) = \int_0^\infty x^{a-1}e^{-x} dx \) is the gamma function, \( z_i = (r_i - \mu)/\sigma \) is the standardized interest-rate change, \( \Pi = [2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/\nu)]^{1/2} \), and \( \nu > 0 \) is the degrees of freedom or tail-thickness parameter.

The GED is used by Box and Tiao (1962) to model prior densities in Bayesian estimation, by Nelson (1991) to model the distribution of stock market returns, and by Hsieh (1989a, b) to model the distribution of exchange rates. Bali and Theodossiou (2007) introduce an asymmetric (or skewed) version of the GED and apply it to stock return data. The skewed generalized error distribution (SGED) adds an additional moment, skewness, to the GED formulation. The probability density function for the SGED is
\[ f(r; \mu, \sigma, \nu, \lambda) = \frac{C}{\sigma} \exp\left( -\frac{|r_i - \mu + \delta \sigma|}{[1 + \text{sign}(r_i - \mu + \delta \sigma)\lambda^\nu \theta^\nu \sigma^\nu]} \right), \]
where
\[ C = \frac{\nu}{2\theta} \Gamma\left(\frac{1}{\nu}\right)^{-1}, \quad \theta = \Gamma\left(\frac{1}{\nu}\right)^{-1/2} \Gamma\left(\frac{3}{\nu}\right)^{-1/2} S(\lambda)^{-1}, \]

\[ S(\lambda) = \sqrt{1 + 3\lambda^2 - 4\lambda^2\lambda^2}, \]
\[ A = \Gamma\left(\frac{2}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right) \Gamma\left(\frac{3}{\nu}\right)^{-1/2}, \]
where \( \mu = E(r) \) and \( \sigma \) are the expected value and the standard deviation of \( r \), \( \lambda \) is a skewness parameter, “sign” is the sign function, and \( \Gamma(\cdot) \) is the gamma function. The scaling parameters \( \nu \) and \( \lambda \) obey the following constraints: \( \nu > 0 \) and \( -1 < \lambda < 1 \). The parameter \( \nu \) controls the height and tails of the density function, and the skewness parameter \( \lambda \) controls the rate of descent of the density around the mode of \( r \), where \( \text{mode}(r) = \mu - \delta \sigma \). In the case of positive skewness (\( \lambda > 0 \)), the density function is skewed to the right. This is because for values of \( r < \mu - \delta \sigma \), the interest-rate change \( r \) is weighted by a greater value than unity and for values of \( r > \mu - \delta \sigma \) by a value less than unity. The opposite is true for negative \( \lambda \). Note that \( \lambda \) and \( \delta \) have the same sign, thus, in the case of positive skewness (\( \lambda > 0 \)), the mode of \( r \) is less than the expected value of \( r \). The parameter \( \delta \) is Pearson’s skewness, \( [\mu - \text{mode}(r)]/\sigma = 3 \).

The SGED parameters are estimated by maximizing the log-likelihood function of \( r_i \) with respect to the parameters \( \mu, \sigma, \nu, \) and \( \lambda \):
\[ \log L = n \log C - n \log \sigma - \frac{1}{\theta^2 \sigma^2} \sum_{i=1}^{n} \left( \frac{|r_i - \mu + \delta \sigma|}{1 + \text{sign}(r_i - \mu + \delta \sigma)\lambda^\nu} \right)^\nu, \]
where “sign” is the sign of the residuals \( (r_i - \mu + \delta \sigma) \), and \( n \) is the sample size.

#### 2.2.2. Skewed \( t \) Distribution

Bollerslev (1987) and Bollerslev and Wooldridge (1992) use the standard normal distribution for testing the standardized student \( t \) distribution. The symmetric standardized \( t \) density with \( 2 < \nu < \infty \) is given by
\[ f(r; \mu, \sigma, \nu) = \Gamma\left(\frac{\nu+1}{2}\right) \Gamma\left(\frac{\nu}{2}\right)^{-1} \left[ (\nu-2)\sigma^2 \right]^{-1/2} \left[ 1 + \frac{(r_i - \mu)^2}{(\nu-2)\sigma^2} \right]^{-(\nu+1)/2}. \]
It is well known that for \( 1/\nu \rightarrow 0 \), the \( t \) distribution approaches a normal distribution; but for \( 1/\nu > 0 \), the \( t \) distribution has fatter tails than the corresponding normal distribution.

Hansen (1994) introduces a generalization of the student \( t \) distribution where asymmetries may occur, while maintaining the assumption of a zero mean and

\[ \delta = 2\lambda AS(\lambda)^{-1}, \quad S(\lambda) = \sqrt{1 + 3\lambda^2 - 4\lambda^2\lambda^2}, \]
\[ A = \Gamma\left(\frac{2}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right) \Gamma\left(\frac{3}{\nu}\right)^{-1/2}, \]
where \( \mu = E(r) \) and \( \sigma \) are the expected value and the standard deviation of \( r \), \( \lambda \) is a skewness parameter, “sign” is the sign function, and \( \Gamma(\cdot) \) is the gamma function. The scaling parameters \( \nu \) and \( \lambda \) obey the following constraints: \( \nu > 0 \) and \( -1 < \lambda < 1 \). The parameter \( \nu \) controls the height and tails of the density function, and the skewness parameter \( \lambda \) controls the rate of descent of the density around the mode of \( r \), where \( \text{mode}(r) = \mu - \delta \sigma \). In the case of positive skewness (\( \lambda > 0 \)), the density function is skewed to the right. This is because for values of \( r < \mu - \delta \sigma \), the interest-rate change \( r \) is weighted by a greater value than unity and for values of \( r > \mu - \delta \sigma \) by a value less than unity. The opposite is true for negative \( \lambda \). Note that \( \lambda \) and \( \delta \) have the same sign, thus, in the case of positive skewness (\( \lambda > 0 \)), the mode of \( r \) is less than the expected value of \( r \). The parameter \( \delta \) is Pearson’s skewness, \( [\mu - \text{mode}(r)]/\sigma = 3 \).

The SGED parameters are estimated by maximizing the log-likelihood function of \( r_i \) with respect to the parameters \( \mu, \sigma, \nu, \) and \( \lambda \):
\[ \log L = n \log C - n \log \sigma - \frac{1}{\theta^2 \sigma^2} \sum_{i=1}^{n} \left( \frac{|r_i - \mu + \delta \sigma|}{1 + \text{sign}(r_i - \mu + \delta \sigma)\lambda^\nu} \right)^\nu, \]
where “sign” is the sign of the residuals \( (r_i - \mu + \delta \sigma) \), and \( n \) is the sample size.

For \( \nu = 2 \), the GED yields the normal distribution, while for \( \nu = 1 \), it yields the Laplace or the double-exponential distribution. If \( \nu < 2 \), the density has thinner tails than the normal, whereas for \( \nu > 2 \), it has thinner tails.
unit variance. The skewed $t$ density that provides a flexible tool for modeling the empirical distribution of interest rates is given by

$$f(z_i; \mu, \sigma, v, \lambda) = \begin{cases} \frac{bc}{\Gamma((v+1)/2)} \left( \frac{b(z_i + a)}{v-2} \right)^2 \Gamma(v/2), & \text{if } z_i < -a/b, \\ \frac{bc}{\Gamma((v+1)/2)} \left( \frac{b(z_i + a)}{v-2} \right)^2 \Gamma(v/2), & \text{if } z_i \geq -a/b, \end{cases}$$

(13)

where $z_i = (r_i - \mu)/\sigma$ is the standardized interest-rate change, and the constants $a, b,$ and $c$ are given by

$$a = 4\lambda c \left( \frac{v-2}{v-1} \right), \quad b^2 = 1 + 3\lambda^2 - a^2,$$  

$$c = \frac{\Gamma((v+1)/2)}{\sqrt{\pi (v-2) \Gamma(v/2)}}.$$  

(14)

Hansen shows that this density is defined for $2 < v < \infty$ and $-1 < \lambda < 1$. This density has a single mode at $-a/b$, which is of the opposite sign with the parameter $\lambda$. Thus, if $\lambda > 0$, the mode of the density is to the left of zero and the variable is skewed to the right, and vice versa when $\lambda < 0$. Furthermore, if $\lambda = 0$, Hansen’s distribution reduces to the traditional standardized $t$ distribution. If $\lambda = 0$ and $v = \infty$, it reduces to a normal density.

The parameters of the skewed $t$ density are estimated by maximizing the log-likelihood function of $r_i$ with respect to the parameters $\mu, \sigma, v,$ and $\lambda$:

$$\text{LogL} = n \ln b + n \ln \Gamma\left(\frac{v+1}{2}\right) - \frac{n}{2} \ln \pi$$  

$$- n \ln \Gamma(v/2) - n \ln \sigma$$  

$$- \frac{v+1}{2} \sum_{i=1}^{n} \ln \left(1 + \frac{d_i^2}{v-2}\right),$$  

(15)

where $d_i = (b z_i + a)/(1 - As)$ and $s$ is a sign dummy taking the value of one if $b z_i + a < 0$, and $s = -1$ otherwise.\(^4\)

3. Data

The data set is obtained from the Federal Reserve H.15 database, and consists of daily observations for constant-maturity Treasury yields on the 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, and 30-year U.S. government bonds. The time period of investigation for the annualized yields on Treasury securities extends from January 4, 1982 to December 31, 2001 for a total of 4,998 daily observations.\(^5\) Yields on Treasury securities at constant maturity are interpolated by the U.S. Treasury from the daily yield curve. This curve, which relates the yield on a security to its time to maturity, is based on the closing market bid yields on actively traded Treasury securities in the over-the-counter market. These market yields are calculated from composites of quotations obtained by the Federal Reserve Bank of New York. The constant maturity yield values are read from the yield curve at fixed maturities of 3 and 6 months and 1, 2, 3, 5, 7, 10, and 30 years.

Table 1 provides descriptive statistics of the level and first differences of the data series. The unconditional means (M) of the interest-rate levels reveal a positive relationship between time to maturity and sample mean rate, implying an upward-sloping yield curve. The unconditional average levels of the 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, and 30-year rates are about 6.22%, 6.46%, 6.69%, 7.14%, 7.34%, 7.62%, 7.83%, 7.91%, and 8.12%, respectively. The relationship between time to maturity and the unconditional standard deviation (SD) of interest-rate changes is negative, yielding a downward-sloping volatility curve, as expected, given that long-term interest rates are averages of current and future expected short-term rates. The unconditional daily standard deviations of the 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, and 30-year Treasury yields are about 0.0899%, 0.0858%, 0.0804%, 0.0772%, 0.0773%, 0.0760%, 0.0755%, 0.0726%, and 0.0655%, respectively. The unconditional means of interest-rate changes are in the range of $-0.0017\%$ to $-0.0020\%$.

The skewness (S) and excess kurtosis (K) statistics are reported for testing the distributional assumption of normality. The statistics in Table 1 indicate that the daily changes in short-term rates (3-month, 6-month) are positively skewed, while the daily changes in medium- and long-term rates (1, 2, 3, 5, 7, 10, and 30-year) are negatively skewed. The skewness statistics are not large, but they are statistically significant. The excess kurtosis values for interest-rate levels are found to be between 3.10 and 4.11, and they are statistically different from zero, implying that the distributions of interest-rate levels have fat tails compared to the normal distribution. As expected, the excess kurtosis statistics for interest-rate changes are much greater than those for interest-rate levels. Specifically, K ranges from 8.63 (for seven-year Treasury) to 44.57

\(^4\) For the efficient or simulated method of moment estimation of interest-rate models, see Gallant and Tauchen (1997, 1998).

\(^5\) In our data set, the short-term rates cover a longer sample period than the long-term rates. Although some of the interest-rate series start from the early 1950s, we use the longest common sample from January 4, 1982 to December 31, 2001.
The daily data extending from January 4, 1982 to December 31, 2001 yield 1,000 weekly, 238 monthly, and 80 quarterly extremes that are modeled with the GEV distribution.

4. Estimation Results

Table 2 presents the maximum-likelihood estimates of the GEV distribution based on the one-week (five trading days), one-month (21 trading days), and three-month (63 trading days) extremes. The location, scale, and shape parameters of the GEV are found to be statistically significant at the 1% level. The empirical results are clear cut and allow one to determine unambiguously the type of extreme value distribution: For both the largest falls and rises of 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, and 30-year Treasury rates, the asymptotic distribution belongs to the domain of attraction of the Frechet distribution. Although not presented in the paper, a likelihood-ratio test between the Frechet case and the Gumbel case leads to a firm rejection of the Gumbel distribution (and a fortiori a rejection of the Weibull distribution).6

Table 1

<table>
<thead>
<tr>
<th>Treasury yields</th>
<th>Interest-rate levels</th>
<th>M</th>
<th>SD</th>
<th>S</th>
<th>K</th>
<th>Max.</th>
<th>Min.</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-month</td>
<td></td>
<td>6.222</td>
<td>2.2483</td>
<td>0.8708</td>
<td>4.0650</td>
<td>15.49</td>
<td>1.66</td>
<td>-1.8776</td>
</tr>
<tr>
<td>6-month</td>
<td></td>
<td>6.4648</td>
<td>2.3632</td>
<td>0.9132</td>
<td>4.1088</td>
<td>15.67</td>
<td>1.73</td>
<td>-2.2054</td>
</tr>
<tr>
<td>1-year</td>
<td></td>
<td>6.6882</td>
<td>2.4233</td>
<td>0.9115</td>
<td>3.9107</td>
<td>15.23</td>
<td>1.93</td>
<td>-2.2485</td>
</tr>
<tr>
<td>2-year</td>
<td></td>
<td>7.1446</td>
<td>2.4603</td>
<td>0.9362</td>
<td>3.6736</td>
<td>15.18</td>
<td>2.32</td>
<td>-2.2909</td>
</tr>
<tr>
<td>3-year</td>
<td></td>
<td>7.3398</td>
<td>2.4394</td>
<td>0.9592</td>
<td>3.5663</td>
<td>15.16</td>
<td>2.70</td>
<td>-2.4148</td>
</tr>
<tr>
<td>5-year</td>
<td></td>
<td>7.6160</td>
<td>2.3940</td>
<td>0.9890</td>
<td>3.4284</td>
<td>15.02</td>
<td>3.47</td>
<td>-2.5214</td>
</tr>
<tr>
<td>7-year</td>
<td></td>
<td>7.8264</td>
<td>2.3605</td>
<td>0.9744</td>
<td>3.3057</td>
<td>14.98</td>
<td>3.95</td>
<td>-2.5749</td>
</tr>
<tr>
<td>10-year</td>
<td></td>
<td>7.9117</td>
<td>2.3382</td>
<td>0.9296</td>
<td>3.1941</td>
<td>14.95</td>
<td>4.16</td>
<td>-2.5385</td>
</tr>
<tr>
<td>30-year</td>
<td></td>
<td>8.1153</td>
<td>2.1846</td>
<td>0.9058</td>
<td>3.1060</td>
<td>14.80</td>
<td>4.70</td>
<td>-2.4849</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interest-rate changes</th>
<th>M</th>
<th>SD</th>
<th>S</th>
<th>K</th>
<th>Max.</th>
<th>Min.</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-month</td>
<td>-0.0020</td>
<td>0.0899</td>
<td>0.4939</td>
<td>44.574</td>
<td>1.69</td>
<td>-1.04</td>
<td>-29.731</td>
</tr>
<tr>
<td>6-month</td>
<td>-0.0023</td>
<td>0.0856</td>
<td>0.4300</td>
<td>35.150</td>
<td>1.33</td>
<td>-0.94</td>
<td>-21.792</td>
</tr>
<tr>
<td>1-year</td>
<td>-0.0023</td>
<td>0.0804</td>
<td>-0.4496</td>
<td>21.907</td>
<td>1.04</td>
<td>-0.83</td>
<td>-32.222</td>
</tr>
<tr>
<td>2-year</td>
<td>-0.0022</td>
<td>0.0772</td>
<td>-0.4914</td>
<td>14.026</td>
<td>0.80</td>
<td>-0.84</td>
<td>-32.129</td>
</tr>
<tr>
<td>3-year</td>
<td>-0.0021</td>
<td>0.0773</td>
<td>-0.3341</td>
<td>10.812</td>
<td>0.63</td>
<td>-0.79</td>
<td>-32.140</td>
</tr>
<tr>
<td>5-year</td>
<td>-0.0020</td>
<td>0.0760</td>
<td>-0.3191</td>
<td>9.4942</td>
<td>0.55</td>
<td>-0.77</td>
<td>-32.365</td>
</tr>
<tr>
<td>7-year</td>
<td>-0.0019</td>
<td>0.0755</td>
<td>-0.2580</td>
<td>8.6277</td>
<td>0.44</td>
<td>-0.77</td>
<td>-32.524</td>
</tr>
<tr>
<td>10-year</td>
<td>-0.0018</td>
<td>0.0726</td>
<td>-0.2468</td>
<td>8.6650</td>
<td>0.44</td>
<td>-0.75</td>
<td>-32.290</td>
</tr>
<tr>
<td>30-year</td>
<td>-0.0017</td>
<td>0.0655</td>
<td>-0.3321</td>
<td>9.4883</td>
<td>0.42</td>
<td>-0.76</td>
<td>-32.268</td>
</tr>
</tbody>
</table>

Notes: The daily data are obtained from the Federal Reserve H.15 database and consist of constant maturity Treasury yields on 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, and 30-year U.S. government bonds. The time period of investigation for the annualized yields on Treasury securities extends from January 4, 1982 to December 31, 2001, giving a total of 4,998 daily observations. Interest rates are expressed in annualized percentage terms. The means (M), standard deviations (SD), skewness (S), kurtosis (K), maximum (Max.), minimum (Min.), the Augmented Dickey-Fuller (ADF) unit root statistics with a 10% (5%) [1%] critical value of $-2.5674 \ (-2.8626) [-3.4348]$ are reported. `*, **, ***` denote the 10%, 5%, 1% level of significance, respectively.

In all cases, the tail index $\xi$ is estimated to be positive and statistically different from zero. A notable point in Table 2 is that the estimated shape parameters for the short-term rates are greater than those for the long-term rates. Another notable point is that, for all data sets, the tail index $\xi$ for the minimal changes in Treasury yields is estimated to be higher than that for the maximal changes. More specifically, the estimated tail index for the maximal changes is found to be in the range of 0.05 to 0.26 for the one-week, 0.06 to 0.40 for the one-month, and 0.08 to 0.43 for the three-month extremes. For the minimal changes in Treasury rates, the estimate of $\xi$ varies from 0.15 to 0.31 for the one-week, 0.26 to 0.43 for the one-month, and 0.35 to 0.46 for the three-month extremes. Because the higher $\xi$, the fatter the distribution of extremes, the minimal changes in Treasury yields have fatter tails than the maximal changes. In addition, for both the maxima and minima, the extreme changes in short-term rates have thicker tails than the extreme changes in long-term rates.

The maximum-likelihood estimates of the location parameters ($\mu_{\text{max}}, \mu_{\text{min}}$) that determine the mean of the extremes along with $\sigma$ and $\xi$ are statistically significant at the 1% level. The scale parameters ($\sigma_{\text{max}}, \sigma_{\text{min}}$) that determine the volatility of extremes along with $\xi$ are estimated to be positive and statistically significant at the 1% level. As expected, the estimated $\mu$ and $\sigma$ values are greater for the (for three-month Treasury). A notable point is that the tails of the empirical distribution for the short-term rates are much thicker than those for the long-term rates. The augmented Dickey-Fuller (ADF) statistics indicate acceptance of the null hypothesis of a unit root for the levels of interest rates (except for the seven-year Treasury), and rejection of the null for interest-rate changes at the 1% level.

In this paper, the extremes are obtained from the highest (maximum) and lowest (minimum) daily interest-rate changes over five trading days (one week), 21 trading days (one month), and 63 trading days (three months). The daily data extending from January 4, 1982 to December 31, 2001 yield 1,000 weekly, 238 monthly, and 80 quarterly extremes that are modeled with the GEV distribution.

6 The likelihood ratio (LR) test statistic is calculated as $LR = -2(\text{LogL}^* - \text{LogL})$, where $\text{LogL}^*$ is the value of the log-likelihood under the null hypothesis, and $\text{LogL}$ is the log-likelihood under the alternative. This statistic is distributed as Chi-square with degrees of freedom equal to the number of restrictions under the null hypothesis (in this case, it equals one). To save space, we do not present the maximum-likelihood estimates of the Gumbel distribution and corresponding LR statistics that are well above the critical value at the 1% level. They are available upon request from the authors.
three-month extremes than for the one-month and one-week extremes.

5. Estimating Volatility Based on the GEV Distribution

One of our objectives is to estimate the volatility of short-term and long-term interest rates during the extreme movements of the Treasury securities market. The volatility of extremes is determined based on the maximum-likelihood parameter estimates of the GEV distribution. Using the first and second moments of the maxima and minima for the GEV, we derive the volatility of extremes. As shown in Equation (16), the standard deviation (or volatility) of extremes is determined by the scale and shape parameters of the extreme value distribution:

$$
\text{Volatility}_{\text{GEV}} = \frac{\sigma \left[ \Gamma(1 - 2\xi) - \left[ \Gamma(1 - \xi) \right]^2 \right]^{0.5}}{\xi},
$$

where $\Gamma(.)$ is the gamma function. Equation (16) shows that the volatility of extremes is a positive

7 The supplement presents the derivation of Equation (16) based on the first and second moments of the GEV distribution.
function of the scale and shape parameters, which measure dispersion and tail thickness of extremes, respectively.

The supplement presents the annualized volatility of one-week, one-month, and three-month extremes obtained from Equation (16) based on the estimated scale and shape parameters of the GEV distribution. In addition, the original daily data (January 4, 1982 until December 31, 2001) for constant-maturity Treasury yields are used to estimate the volatility of interest-rate changes based on the normal distribution. The supplement shows that for both the maximal and the minimal changes in interest rates, the volatilities of extremes decrease with the increase in time to maturity. This implies the presence of downward-sloping term structure of interest-rate volatility during the extreme movements of the U.S. Treasury market. For the minimal changes in Treasury yields, as time to maturity rises from 3 months to 30 years, VolatilityGEV decreases from 1.92% to 1.22% for the one-week, from 3.86% to 1.39% for the one-month, and from 6.94% to 2.77% for the three-month extremes. For the maximal changes, the corresponding estimates of VolatilityGEV decrease from 1.85% to 1.18% for the one-week, from 3.13% to 1.09% for the one-month, and from 4.14% to 1.17% for the three-month extremes. The results indicate that the minimal changes in Treasury rates are more volatile than the maximal changes. This may be because of the higher absolute values of the minima compared to the maxima over the estimation period. The results suggest that the so-called level effect in interest-rate volatility does exist in the volatility of extreme values as well. As expected, the normal distribution underestimates the volatility of Treasury yields during the extreme movements of the market because the volatility curves for both the maxima and minima are placed above the volatility of original series.

It is possible that the mean and volatility of extremes change over time. By modeling time variation in GEV parameters, we can generate a conditional extreme value volatility estimator and capture the time-series variation in the entire term structure during the extreme movements of the market. We specify the location, scale, and shape parameters of the GEV distribution as a function of the past extreme changes in interest rates and find significant serial correlation and clustering in the GEV parameters. The results presented in the supplement show that the time-varying location, scale, and shape parameters can be used to generate a conditional extreme value volatility estimator.9

6. Pricing Implications for Interest Rate Options

6.1. Option Pricing Model Based on the Lognormal Distribution

Miltersen et al. (1997) derive a unified model that gives closed-form solutions for caps and floors written on interest rates as well as puts and calls written on zero-coupon bonds. Miltersen et al. model the evolution of the discretely compounded forward rate over a fixed period of length \( \Phi \). Letting \( P(t, T) \) denote the price of a zero-coupon bond at time \( t \) for a bond maturing at time \( T \), the simple forward rate at time \( t \) for the future time interval \( [T, T + \Phi] \) is given as \( P(t, T + \Phi) = P(t, T)/[1 + \Phi \cdot f(t, T, \Phi)] \). For any maturity \( T \), the forward rate \( f(t, T, \Phi) \) for the fixed period of length \( \Phi \) is assumed to evolve according to the lognormal diffusion under the empirical measure, i.e.,

\[
\frac{df(t, T, \Phi)}{f(t, T, \Phi)} = \mu(t, T, \Phi)dt + \sigma(t, T, \Phi)dW_t, \tag{17}
\]

where \( W_t \) is a one-dimensional standard Wiener process defined on a filtered probability space.

In the Miltersen et al. framework, \( \text{Call}(t, T, X) \) denotes the price at time \( t \) of a European interest-rate call option that matures at time \( T \) and has an exercise rate of \( X \). The payoff of a T-bill interest-rate option is based on the discount rate. Letting \( P(t, T) \) denote the time \( t \) price of the T-bill with \( (T - t) \) maturity (paying $1 at maturity), the annualized discount rate is given as \( (100/(T - t))(1 - P(t, T)) \). Defining the exercise value of the interest-rate call options as (discount rate – strike rate), the value of the call option at maturity is

\[
\text{Call}(t, T, X) = \max \left[ \frac{100}{\Phi} (1 - P(T, T + \Phi)) - X \right] = \frac{100}{\Phi} \max[\bar{X} - P(T, T + \Phi)], \tag{18}
\]

8 Chan et al. (1992), Andersen and Lund (1997), and Bali (2000) show that the volatility of interest-rate changes is high when the level of the interest rate is high. Therefore, higher absolute values of the minima cause higher volatilities.
where $\bar{X} = 1 - (\Phi X/100)$. The value of the interest-rate call option at maturity is simply a constant multiplied by the payoff of a European put option that provides the holder the right to sell a zero-coupon bond with a maturity of $\Phi$ years at time $T$ for a price of $\bar{X}$. Therefore, Miltersen et al. (1997) directly apply the closed-form solution for the price of a European put option with a zero-coupon bond as the underlying asset. Hence, they show that the no-arbitrage price of the European interest-rate call option is given as

$$\text{Call}(t, T, X) = \frac{100}{\Phi} \left[ (1 - \bar{X})P(t, T + \Phi)N(-d_2) - (1 - \bar{X})P(t, T + \Phi)N(-d_1) \right],$$

where $N(.)$ denotes the standard normal cumulative distribution function, and

$$d_{1,2} = \frac{1}{\sigma(t, T, \Phi)} \left[ \ln \left( \frac{(1 - \bar{X})P(t, T + \Phi)}{\bar{X}P(t, T - P(t, T + \Phi))} \right) \pm \frac{\gamma^2(t, T, \Phi)}{2} \right],$$

$$\gamma^2(t, T, \Phi) = \int_t^T \sigma^2(s, T, \Phi) \, ds.$$ 

In a similar manner, they derive a closed-form expression for the put interest-rate option price:

$$\text{Put}(t, T, X) = \frac{100}{\Phi} \left[ (1 - \bar{X})P(t, T + \Phi)N(d_1) - \bar{X}(P(t, T - P(t, T + \Phi))N(d_2) \right].$$

For constant volatility specification, $\gamma(t, T, \Phi) = \gamma_0$, the option-pricing model of Miltersen et al. reduces to the well-known Black (1976) model. There is a subtle difference between the common usage of the two models: A Black (1976) volatility is attached to each contract while the volatility parameter $\gamma_0$ in the Miltersen et al. model is backed out from a cross-section of option contracts. Bali (2002) considers two additional specifications of the instantaneous volatility of the forward rate: (1) affine volatility: $\gamma(t, T, \Phi) = \gamma_0 + \gamma_1(T - t)$, and (2) exponential volatility: $\gamma(t, T, \Phi) = \gamma_0 \exp(-\gamma_1(T - t))$. All three specifications (constant, affine, and exponential) have one common feature, namely, that they are time invariant, i.e., the volatility only depends on $(T - t)$, and not on the specific calendar date, $t$. Bali (2002) shows that the option-pricing performance of the three volatility specifications are mixed; no one is preferred above the others in all respects. In other words, for the IRX interest-rate options, the additional parameter, $\gamma_1$, in the affine and exponential volatility function is not justified.

### 6.2. Option-Pricing Model Based on the GEV Distribution

The Black-Scholes (1973) and lognormal distribution-based models cannot explicitly account for the negative skewness and the excess kurtosis of asset returns. A large literature has developed that aims to extract the risk-neutral density function from traded option prices so that the skewness and fat-tail properties of the distribution are better captured than is the case in lognormal models.

We will now introduce an option-pricing model based on the GEV distribution that overcomes the problems associated with the standard option-pricing models. Although the newly proposed model can be used to price European call/put options written on individual stocks, stock indices, currency, interest rates, and futures contracts, our focus in this paper will be on pricing interest-rate options. To give the generalized version of the option-pricing model, we will now use different notations than those in §6.1.

According to the asset-pricing framework of Harrison and Pliska (1981), the risk-neutral probability density function (RND) exists under an assumption of no arbitrage. By definition of a no-arbitrage equilibrium, the current price of an asset is the present discounted value of its expected future payoffs given a risk-free interest rate where the expectation is evaluated by the RND function. Let $S_t$ denote the underlying asset price at time $t$ (in our case, the interest rate index called IRX). The European call option $C_t$ is written on this asset with strike $X$ and maturity $T$. Following the Harrison and Pliska (1981) result on the arbitrage-free European call option price, there exists an RND function, $g(S_t)$, such that the equilibrium call option price can be written as

$$C_t = e^{-r(T-t)}E^Q_t(\max(S_T) - X, 0) = e^{-r(T-t)}\int_X^\infty (S_T - X)g(S_T) \, dS_T,$$

where $r$ is the risk-free interest rate and $E^Q(.)$ is the risk-neutral expectation operator conditional on all information available at time $t$, and $g(S_T)$ is the risk-neutral density function of the underlying asset at maturity. Similarly, the arbitrage-free option-pricing formula for a put option is given by

$$P_t = e^{-r(T-t)}E^Q_t(X - \max(S_T), 0) = e^{-r(T-t)}\int_0^X (X - S_t)g(S_t) \, dS_T.$$ 

In an arbitrage-free framework, the following martingale condition must also be satisfied:

$$S_t = e^{-r(T-t)}E^Q_t(S_T).$$

We now assume that the distribution of asset return for a holding period equal to time to maturity of the
option is represented by the GEV distribution. We derive closed-form solutions for the call- and put-pricing equations by analytically solving the integrals in (22) and (23). For the purpose of obtaining an analytical closed-form solution, it was found to be necessary to define returns as simple returns. Further, following the convention in extreme value theory where Frechet-type distributions with \( \xi > 0 \) for the tail index parameter are associated with losses, we model asset returns in terms of losses:

\[
L_T = -R_T = -\frac{S_T - S_t}{S_t} = 1 - \frac{S_T}{S_t},
\]

(25)

where \( R_T \) in our empirical analysis is the return on the 13-week interest-rate index called IRX.

By assuming that \( L_T \) follows the standardized GEV distribution, the density function for the negative returns is given by

\[
h(L_T) = \frac{1}{\sigma} \left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-1 - 1/\xi} \cdot \exp \left(-\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-1/\xi}\right).
\]

(26)

The RND function \( g(S_T) \) in (22) for the underlying asset price \( S_T \) is given by the general formula

\[
g(S_T) = h(L_T) \left| \frac{dL_T}{dS_T} \right| = h(L_T) \frac{1}{S_T}.
\]

(27)

By substituting (26) into (27), we obtain the RND function of the underlying asset price in terms of the standardized GEV density function

\[
g(S_T) = \frac{1}{S_T \sigma} \left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-1 - 1/\xi} \cdot \exp \left(-\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-1/\xi}\right),
\]

(28)

with

\[1 + \frac{\xi}{\sigma} (L_T - \mu) = 1 + \frac{\xi}{\sigma} \left(1 - \frac{S_T}{S_t} - \mu\right) > 0.
\]

Because the statistical results indicate strong rejection of the Weibull and Gumbel distributions in favor of the Frechet distribution with \( \xi > 0 \), we derive the closed-form option-pricing solutions for \( \xi > 0 \). When \( \xi > 0 \), the price RND function in (28) is truncated on the right, and hence, the upper limit of the integral for the call option price in (22) becomes \( S_t \left(1 - \mu + \frac{\sigma}{\xi} \right) \).

Substituting \( g(S_T) \) in (28) into the call price equation in (22) yields

\[
C_t = e^{-r(T-t)} \int_X S_t \left(1 - \mu + \frac{\sigma}{\xi} \right) \left(S_t - X\right) \frac{1}{S_T \sigma} \left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-1 - 1/\xi} \cdot \exp \left(-\left(1 + \frac{\xi(L_T - \mu)}{\sigma}\right)^{-1/\xi}\right) dS_T.
\]

(29)

Using the following change of variable,

\[
q = 1 + \frac{\xi}{\sigma} (L_T - \mu) = 1 + \frac{\xi}{\sigma} \left(1 - \frac{S_T}{S_t} - \mu\right),
\]

(30)

the underlying asset price \( S_T \) and \( dS_T \) can be written in terms of \( q \) as follows:

\[
S_T = S_t \left(1 - \mu + \frac{\sigma}{\xi} (q - 1)\right) \quad \text{and} \quad dS_T = -S_t \frac{\sigma}{\xi} dq.
\]

(31)

Also, the GEV density function in (28) becomes

\[
g(q) = \frac{1}{S_t \sigma} \left(q^{-1 - \frac{1}{\xi}}\right) \exp(-q^{-1/\xi}).
\]

(32)

Under the change of variable, the lower limit of the integral for the call option price equation in (29) becomes

\[
\Pi = 1 + \frac{\xi}{\sigma} \left(1 - \frac{X}{S_t} - \mu\right),
\]

(33)

while the upper limit of the integral in (29) becomes zero. Substituting for \( S_T \) and \( dS_T \) as defined in (31) into (28) and using the new limits of integral, we have

\[
C_t e^{r(T-t)} = \int_0^\infty \left(S_t \left(1 - \mu + \frac{\sigma}{\xi} (q - 1)\right) - X\right) \frac{1}{S_t \sigma} \left(q^{-1 - \frac{1}{\xi}}\right) \exp(-q^{-1/\xi}) dq
\]

\[
\cdot \exp(-q^{-1/\xi}) \left(-S_t \frac{\sigma}{\xi} dq\right).
\]

(34)

Simplifying and rearranging (34), we have

\[
C_t e^{r(T-t)} = -\frac{1}{\xi} \int_0^\Pi \left(S_t \left(1 - \mu + \frac{\sigma}{\xi} (q - 1)\right) - X\right) \frac{S_t \sigma}{\xi} dq \exp(-q^{-1/\xi}) dq
\]

\[
\cdot \left(\frac{1}{S_t \sigma} \psi_1 - \left(S_t \left(1 - \mu + \frac{\sigma}{\xi}\right) - X\right) \psi_2\right).
\]

(35)

The integrals \( \psi_1 \) and \( \psi_2 \) in (35) can be evaluated in terms of the incomplete gamma function (see the online supplement), and their solutions are

\[
\psi_1 = \int_0^\Pi q^{-1/\xi} \exp(-q^{-1/\xi}) dq = -\xi \Gamma(1 - \xi, \Pi^{-1/\xi}),
\]

(36)

\[
\psi_2 = \int_0^\Pi (q^{-1 - \frac{1}{\xi}}) \exp(-q^{-1/\xi}) dq = [\xi \exp(-\Pi^{-1/\xi})]_0^\Pi
\]

\[
= \xi (-\exp(-\Pi^{-1/\xi})).
\]

(37)
Substituting $\psi_1$ and $\psi_2$ into (35), we obtain a closed-form solution for the GEV call option price:

$$C_t = e^{-r(T-t)} \left[ -\frac{S_t \sigma}{\xi} \Gamma(1-\xi, \Pi^{-1/\xi}) - \left( S_t \left( 1 - \mu \frac{\sigma}{\xi} \right) - X \right) \cdot \left( -\exp(-\Pi^{-1/\xi}) \right) \right]. \quad (38)$$

Rearranging Equation (38) yields

$$C_t = e^{-r(T-t)} \left[ S_t \left( \left( 1 - \mu \frac{\sigma}{\xi} \right) e^{-\Pi^{-1/\xi}} - \frac{\sigma}{\xi} \Gamma(1-\xi, \Pi^{-1/\xi}) \right) - X e^{-\Pi^{-1/\xi}} \right]. \quad (39)$$

The derivation of a closed-form solution for a put option price under GEV returns can be found in the appendix, and yields the following equation:

$$P_t = e^{-r(T-t)} \left[ X \left( e^{-\omega^{-1/\xi}} - e^{-\Pi^{-1/\xi}} \right) - S_t \left( \left( 1 - \mu \frac{\sigma}{\xi} \right) e^{-\Pi^{-1/\xi}} - e^{-\omega^{-1/\xi}} \right) - \frac{\sigma}{\xi} \Gamma(1-\xi, \omega^{-1/\xi}, \Pi^{-1/\xi}) \right], \quad (40)$$

where $\omega = 1 + (\xi/\sigma)(1 - \mu) > 0$. Note that $\omega$ is a constant, given a set of parameters $\mu$, $\sigma$, and $\xi$.

### 6.3. Option-Pricing Model Based on the Fat-Tailed Distributions

The forward rates are generally assumed to be lognormally distributed, which yields a closed-form solution for interest-rate caps and floors (see Heath et al. 1992, Milletersen et al. 1997, Brace et al. 1997). To test the empirical validity of the lognormality assumption, we estimate the empirical distribution of log-forward rate changes using the normal, GED, symmetric $t$, SGED, and skewed $t$ distributions. We use the daily three-month and six-month T-bill rates to generate the forward interest-rate series. The three-month T-bill rates extend from January 4, 1954 to December 31, 2002, yielding a total of 12,237 daily observations. The daily six-month T-bill data cover the period from December 9, 1958 to December 31, 2002, giving a total of 10,997 observations. Our empirical analyses are based on the forward interest rates obtained from the common sample period December 9, 1958 until December 31, 2002.

Table 3 presents the maximum-likelihood estimates for the daily data from December 9, 1958 to December 31, 2002. A notable point in Table 3 is that the tail-thickness parameter ($\nu$) of the GED, SGED, symmetric $t$, and skewed $t$ distributions is highly significant and indicates that the tails of the distribution of log-forward rate changes are substantially thicker than the tails of the normal distribution. Another important point in Table 3 is that the skewness parameters ($\lambda$) of the SGED and skewed $t$ distributions are estimated to be negative, but they are not statistically different from zero. Comparing the maximized log-likelihood values of the SGED versus GED and the skewed $t$ versus symmetric $t$ indicates that the null hypothesis of $\lambda = 0$ cannot be rejected based on the likelihood-ratio test. The maximized log-likelihood values also show that the normal distribution is rejected in favor of the GED and symmetric $t$ distributions. These results provide strong evidence that the empirical distribution of forward rates is far from lognormal, and the fat-tail property is more dominant than skewness in the sample.

Overall, the parameter estimates for the extreme value and skewed fat-tailed distributions indicate the presence of significant kurtosis in spot and forward rate changes. To accommodate the tail-thickness of the interest-rate distribution, we now derive an option-pricing model with GED and symmetric $t$ distributions.

We consider a European call option written on a forward interest rate with a strike rate of $X$ and with a current level of $F_0$. The value of this option at expiration $T$ is equal to $C_T = \max(F_T - X, 0)$, where $F_T$ is the forward rate at the expiration of the option. The expected value of the option at time $T$ is

$$E[\max(F_T - X, 0)] = \int_{X}^{\infty} (F_T - X) g(F) dF, \quad (41)$$

where $g(F)$ is the probability density function of the forward interest rate $F$. Because of the lognormality assumption, the forward rate at time $T$ can be expressed as

$$F_T = F_0 e^{\xi}, \quad (42)$$

$$F_T = F_0 e^{\xi}, \quad (42)$$

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\lambda$</th>
<th>$\nu$</th>
<th>LogL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.000099</td>
<td>0.01609</td>
<td>0.0</td>
<td>0.0</td>
<td>30,396.71</td>
</tr>
<tr>
<td>GED</td>
<td>-0.000081</td>
<td>0.01553</td>
<td>0.0</td>
<td>0.9023</td>
<td>31,845.37</td>
</tr>
<tr>
<td>Symmetric $t$</td>
<td>0.000113</td>
<td>0.01932</td>
<td>0.0</td>
<td>2.6390</td>
<td>38,176.83</td>
</tr>
<tr>
<td>SGED</td>
<td>-0.000052</td>
<td>0.01545</td>
<td>-0.0023</td>
<td>0.9223</td>
<td>31,846.01</td>
</tr>
<tr>
<td>Skewed $t$</td>
<td>0.000736</td>
<td>0.01931</td>
<td>-0.0176</td>
<td>2.6406</td>
<td>38,177.12</td>
</tr>
</tbody>
</table>

Notes. This table presents the parameter estimates of the normal, GED, symmetric $t$, SGED, and skewed $t$ distributions. The results are based on the daily forward interest rates spanning the period from December 9, 1958 to December 31, 2002. Asymptotic t-statistics are given in parentheses. LogL is the maximized log-likelihood value.
where \( \delta = \ln(F_t/F_0) = \mu T + z\sigma \sqrt{T} \) is the change in log-forward rates from time 0 to \( T \), \( \mu T \) and \( \sigma^2 T \) are the mean and variance of \( \delta \), and \( z \) is a standardized (i.e., zero mean and unit variance) GED or symmetric \( t \) random variable with tail-thickness parameter \( v \). First, the expected value of the call option at time \( T \) can be expressed in terms of the distribution of GED or symmetric \( t \) random variable \( z \). Then, the option price at time \( t = 0 \) can be calculated as the present value of \( E(C_T) \):

\[
C_0 = E_0 e^{(r-T)T} \int_0^\infty e^{z\sigma\sqrt{T}} g(z; \mu, \sigma, v) \, dz - e^{-rT} X[1 - G(q; \mu, \sigma, v)],
\]

(43)

where \( r \) is the risk-free interest rate or equilibrium (required) rate of return for buyers and issuers of call options, \( q = -[\ln(F_0/X) + \mu T]/\sigma \sqrt{T} \), \( g(z; \mu, \sigma, v) \) is the standardized GED or symmetric \( t \) density function with tail-thickness parameter \( v \), and \( G(q; \mu, \sigma, v) = \int_q^{\infty} g(z; \mu, \sigma, v) \, dz \) is the cumulative standardized GED or symmetric \( t \). In Equation (43), the parameters \( \mu, \sigma, \) and \( v \) can be estimated by maximizing the GED or symmetric \( t \) sample log-likelihood function using the data that best reflect the distribution of changes in log-forward rates during the life of the option.

What about the conditional distribution of forward interest rates? Is it lognormal, or is it skewed and fat-tailed? We consider the following AR(1)-GARCH specification with alternative distribution functions to investigate the presence of substantial volatility persistence in the forward-rate process. The maximized log-likelihood values and the standard errors of the tail-thickness parameters \( (v) \) of the conditional GED, SGED, symmetric \( t \), and skewed \( t \) distributions indicate strong rejection of the normality assumption. The skewness parameters \( (\lambda) \) of the SGED and skewed \( t \) distributions are estimated to be negative, but they are not statistically different from zero. These results provide strong evidence that the conditional distribution of forward rates is far from lognormal. It is symmetric and fat tailed.

When pricing IRX options, we use the GED distribution along with the AR(1)-GARCH(1, 1) model given in Equations (44)–(45). As shown in Equation (43), there is no closed-form solution to the option price with the GED-GARCH model. However, once we estimate the conditional mean, volatility, and tail-thickness parameters, we can compute the standardized GED density denoted by \( g(z; \mu, \sigma, v) \) and the cumulative standardized GED function denoted by \( G(q; \mu, \sigma, v) \). In other words, the integral in Equation (43) can be solved numerically.

7. Option-Pricing Performance of Alternative Distribution Functions

7.1. Data and Preliminary Analysis

The interest-rate options data are obtained from the Option Metrics database, which contains a complete record of all bid/ask quotes and trades for all options at the Chicago Board Options Exchange (CBOE). The contracts we consider are European-style options on the 13-week T-bill discount rate, which have the ticker symbol IRX. It can be shown that the IRX options are equivalent to European-type options written on zero-coupon bonds. Hence, we are dealing with the simplest contracts imaginable. The underlying asset of the IRX contract is 10 times the annualized discount rate of the most recently auctioned 13-week T-bill, i.e., the discount yield quoted in terms of the yield indicator. IRX is also known as the short-term interest-rate index.

The contracts expire on the Saturday immediately following the third Friday of the expiration month. The expiration months are the three nearest months and, in addition, two months from the March quarterly cycle (March, June, September, and December). The IRX options are European style, i.e., they can be exercised on the last business day (Friday) before the

---

10 As shown in the online supplement, we consider one-factor models with nonlinear, quadratic, and linear drift, along with the nonlinear, CEV, and linear diffusion functions. We also consider discrete-time GED-GARCH models with nonlinear, quadratic, and linear drift specifications. The results from daily and weekly three-month Treasury yields indicate that the conditional mean of interest rate changes is linear (or constant), whereas the conditional volatility follows a nonlinear and persistent process.

11 At an earlier stage of the study, we used two alternative specifications for the conditional variance process given in Equation (45). In the first specification, the conditional variance is defined as a function of the long period’s forward rate level, like a CEV process: \( \sigma^2 = \sigma^2_f f^{2\phi} \). In the other specification, it is parameterized as a function of the absolute value of the last period’s log-forward rate level: \( \sigma^2 = \sigma^2 |\ln f_{t-1}|^\phi \). In both cases, we find that there is no level effect in the conditional volatility of log-forward rate changes, i.e., \( \phi \) is not statistically different from zero. The estimation results with these two variance specifications are available upon request from the authors.
are estimated from the following econometric specification:

\begin{align*}
\text{LogL} &= \frac{1}{2} \sum_{t=1}^{T} \left( \ln f_t - \ln f_{t-1} - a_0 + a_1 \ln f_{t-1} + \epsilon_t \right), \\
E(\epsilon_t | \Omega_{t-1}) &= 0, \\
E(\epsilon_t^2 | \Omega_{t-1}) &= \sigma_t^2 = \beta_0 + \beta_1 \epsilon_{t-1} + \beta_2 \sigma_{t-1}^2.
\end{align*}

 expiration date. Both call and put options are traded. The IRX options have cash settlement and a size multiplier of $100, i.e., the price is 100 times the quoted premium at CBOE and the payoff at maturity is 100 times the difference between the discount rate (in terms of the yield indicator) and the strike price.

The interest-rate options data cover the period from January 4, 1996 to June 30, 2005, yielding a total of 2,378 business days. This period includes some extreme events such as the Asian crises (July–October 1997), the LTCM crisis (September 1998), the bursting of the high-technology dot-com bubble of 2000–2002, and the 9/11 attacks. Each day, we use the mid quotes (the average of the best bid and best offer) as the call/put option price. We focus on options with maturities between seven and 360 days for which the Black (1976) volatility is between 5% and 100%. We exclude options with the bid price of zero. This sampling procedure results in 54,766 observations for call prices and 71,545 observations for put prices. A total of 126,311 option prices from January 4, 1996 to June 30, 2005 yields an average of 53 observations per day. Even though call and put options are quoted equally frequently, 43% of the options in our sample are call options. This is mainly because of our filtration. If we loosen our filtration procedure, we obtain a similar number of call and put options.

Table 5 shows the summary statistics of the mid quotes for the IRX put and call prices. The contracts are grouped according to their time to maturity; short maturity with (7, 90) calendar days, medium maturity with (90, 180) calendar days, and long maturity with (180, 360) calendar days. Furthermore, a breakdown according to moneyness is presented, where moneyness for a call option is defined as the underlying discount rate divided by the exercise rate, and for a put option is defined as the inverse of this ratio. Options with moneyness in the interval (0, 0.9), [0.9, 1.1), and [1.1, \infty) are denoted ITM (in-the-money), ATM (at-the-money), and OTM (out-of-the-money), respectively.

Table 5 indicates that the average price of call and put options increases with the time to maturity and with the moneyness of the options, ceteris paribus.

### 7.2. Pricing Performance

For each day in our sample, we find the volatility parameter (\gamma_0) of the lognormal-based option-pricing model and the three parameters of the GEV distribution (\mu, \sigma, \xi) that minimize the sum of squared pricing errors:

\begin{equation}
\text{SSE}(t) = \min_{N_t} \sum_{i=1}^{N_t} (\widehat{\text{Price}}_i - \text{Price}_i)^2,
\end{equation}

where \(N_t\) is the number of observations for day \(t\), \(\widehat{\text{Price}}_i\) is the model price for option \(i\) interpreted as a function of the lognormal or GEV parameters, and \(\text{Price}_i\) is the actual price of the option. For the GED-GARCH model, we first estimate the conditional mean, conditional volatility, and tail-thickness parameters of the GED density and then numerically evaluate the integral in Equation (43) to find the model’s estimates for the call and put option prices. The pricing performance of the lognormal, GEV, and GED-GARCH models is evaluated based on the root mean square error (RMSE),
The data cover the period from January 4, 1996 to June 30, 2005.

Moneyness for out, at, in the money (OTM, ATM, ITM) options:
exercise rate for call options, and inversely for put options. Moneyness for dollars:
which represents the average pricing error in terms of

\[ \text{RMSE}(t) = \frac{1}{N} \sqrt{\text{SSE}(t)}. \]  

(47)

The first panel of Table 6 reports the average pricing errors for IRX call options. The GEV distribution outperforms the lognormal and GED-GARCH models at all time horizons. All models consistently display an improvement in performance as time to maturity decreases. Specifically, the GEV distribution removes the pricing bias that the lognormal distribution exhibits for options far from maturity, with the lognormal having an average error of $25.61 for long-term call options with 180 to 360 days to maturity, while the GEV distribution has an average error of $5.27, representing an 80% reduction. Although the GED-GARCH model performs much better than the lognormal distribution, reducing the average pricing error from $25.61 to $12.99, it cannot perform as well as the GEV distribution. For short-term call options with seven to 90 days to maturity, the average pricing error with the lognormal distribution is $7.68, whereas with the GEV distribution it is $2.33. Even though the lognormal distribution improves considerably for options close to maturity, the GEV distribution still represents a 70% reduction in pricing error.

For short-term call options, the GED-GARCH model with an average pricing error of $5.47 outperforms the lognormal distribution, but it still cannot produce more accurate estimates than the GEV distribution. We should note that the aforementioned average pricing errors are relatively low compared to an average contract price of $482 and average bid-ask spread of $42.

The second panel of Table 6 shows that even though the average pricing errors are greater for puts than for calls, a similar result is obtained for put options for all times to maturity and for all distributions. For long-term put options, the average pricing error is $26.51 for the lognormal distribution, while it reduces to $14.23 for the GED-GARCH model, and to $5.76 for the GEV distribution. For short-term put options, the average pricing error is much lower; $7.97 for the lognormal, $5.52 for the GED-GARCH model, and $2.48 for the GEV distribution. Overall, the pricing performance improves for all distributions as time to maturity declines, with the GEV distribution removing the pricing bias that the lognormal distribution exhibits for options far from maturity. Overall, the results indicate that incorporating nonlinearity in the conditional volatility of interest rates and accounting for the nonnormality of interest rates—as in the GED-GARCH model—yield more precise estimates of option prices. However, the pricing performance of the GEV distribution is still superior to the two alternative models considered in the paper.

To see whether the relative performance of the GEV distribution is even higher during the extreme movements of the market, we compute the average pricing errors for each year from 1996 to 2005. We first estimate the annualized volatility of returns on the IRX index to determine which years are relatively more volatile or stable. We use three measures of annual volatility: realized, GARCH, and implied. The annual realized volatility is computed as the square root of the sum of squared daily returns within each year: $\sigma_{\text{realized}} = \left( \sum_{t=1}^{\text{days}} R_t^2 \right)^{1/2}$, where $D_t$ is the number of trading days in year $t$ and $R_t$ is the IRX index return on day $d$. The annualized GARCH volatility is estimated as the square root of the sum of conditional variance of daily returns within each year. We estimate the daily conditional variance by applying the SGED AR(1) GJR-GARCH model in Equations (16)–(17) to daily returns on the IRX index. The annual GARCH volatility is defined as $\sigma_{\text{GARCH}} = \left( \sum_{t=1}^{\text{years}} \sigma_t^2 \right)^{1/2}$, where $D_t$ is the number of trading days in year $t$ and $\sigma_t^2$ is the conditional variance of daily IRX index returns on day $d$. The annual implied volatility is calculated for each year as the average implied volatilities of all at-the-money call and put options (with moneyness of 0.9 to 1.1). We consider options only with maturities between seven and 360 days for which the implied volatility is between 5% and 100%.
Table 6  Option-Pricing Performance of Alternative Distribution Functions

<table>
<thead>
<tr>
<th>Date</th>
<th>Lognormal</th>
<th>GED-GARCH</th>
<th>GEV</th>
<th>Lognormal</th>
<th>GED-GARCH</th>
<th>GEV</th>
<th>Lognormal</th>
<th>GED-GARCH</th>
<th>GEV</th>
<th>Lognormal</th>
<th>GED-GARCH</th>
<th>GEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>19.65</td>
<td>7.96</td>
<td>3.22</td>
<td>15.88</td>
<td>6.36</td>
<td>3.92</td>
<td>3.13</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>18.25</td>
<td>8.15</td>
<td>3.05</td>
<td>17.65</td>
<td>8.07</td>
<td>2.06</td>
<td>3.69</td>
<td>2.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>23.29</td>
<td>11.96</td>
<td>6.36</td>
<td>19.43</td>
<td>10.29</td>
<td>4.41</td>
<td>11.67</td>
<td>7.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>19.06</td>
<td>7.95</td>
<td>3.85</td>
<td>17.46</td>
<td>12.28</td>
<td>6.25</td>
<td>4.56</td>
<td>3.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>20.42</td>
<td>11.09</td>
<td>3.80</td>
<td>17.22</td>
<td>7.27</td>
<td>2.64</td>
<td>4.69</td>
<td>2.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>40.29</td>
<td>21.75</td>
<td>6.92</td>
<td>21.13</td>
<td>11.09</td>
<td>6.35</td>
<td>11.02</td>
<td>8.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>32.91</td>
<td>17.27</td>
<td>5.30</td>
<td>22.84</td>
<td>10.06</td>
<td>5.02</td>
<td>11.79</td>
<td>8.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>35.00</td>
<td>19.66</td>
<td>6.80</td>
<td>20.53</td>
<td>11.02</td>
<td>5.70</td>
<td>9.86</td>
<td>6.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>29.36</td>
<td>16.58</td>
<td>7.98</td>
<td>19.64</td>
<td>9.59</td>
<td>6.03</td>
<td>10.16</td>
<td>6.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>17.86</td>
<td>7.57</td>
<td>3.42</td>
<td>15.04</td>
<td>7.02</td>
<td>3.88</td>
<td>5.48</td>
<td>4.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>25.61</td>
<td>12.99</td>
<td>5.27</td>
<td>18.68</td>
<td>9.30</td>
<td>4.42</td>
<td>7.68</td>
<td>5.47</td>
<td>2.33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7  Annual Volatility of IRX Index Returns

<table>
<thead>
<tr>
<th>Date</th>
<th>Realized (%)</th>
<th>GARCH (%)</th>
<th>Implied (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>10.29</td>
<td>13.38</td>
<td>19.16</td>
</tr>
<tr>
<td>1997</td>
<td>12.49</td>
<td>14.64</td>
<td>20.02</td>
</tr>
<tr>
<td>1998</td>
<td>23.17</td>
<td>24.68</td>
<td>20.45</td>
</tr>
<tr>
<td>1999</td>
<td>13.98</td>
<td>15.48</td>
<td>22.38</td>
</tr>
<tr>
<td>2000</td>
<td>20.03</td>
<td>22.01</td>
<td>19.78</td>
</tr>
<tr>
<td>2001</td>
<td>32.45</td>
<td>32.47</td>
<td>48.41</td>
</tr>
<tr>
<td>2002</td>
<td>23.64</td>
<td>24.25</td>
<td>70.60</td>
</tr>
<tr>
<td>2003</td>
<td>27.79</td>
<td>29.28</td>
<td>72.95</td>
</tr>
<tr>
<td>2004</td>
<td>27.39</td>
<td>28.81</td>
<td>81.39</td>
</tr>
<tr>
<td>2005</td>
<td>15.92</td>
<td>17.61</td>
<td>39.33</td>
</tr>
</tbody>
</table>

Notes: The table presents the annual volatility of IRX index returns for the sample period of 1996 to 2005. The annual realized volatility is computed as the square root of the sum of squared daily returns within each year. The annualized GARCH volatility is estimated as the square root of the sum of conditional variance of daily returns within each year. The annual implied volatility is calculated for each year as the average implied volatilities of at-the-money call and put options.

As shown in Table 7, the U.S. Treasury market was extremely volatile during the years 2001, 2002, 2003, and 2004. The year 1998 can also be viewed as volatile, especially compared to 1996, 1997, and 1999. As expected, the implied volatility measures are higher than the realized and GARCH volatility estimates, but they are all highly correlated and point out the same years as relatively more volatile or stable. We can now look at the average pricing errors in Table 6 more closely. The first panel of Table 6 shows that for long-term call options, RMSE for the lognormal distribution is $40.29, $32.91, $35, and $29.36 for 2001, 2002, 2003, and 2004, respectively. During these extremely large falls and rises of the market, the GEV distribution significantly reduces the pricing bias that the lognormal distribution exhibits for options far from maturity. More specifically, RMSE for the GEV distribution is $6.92, $5.30, $6.80, and $7.98 for 2001, 2002, 2003, and 2004, respectively. As shown in Table 6, similar results are obtained for other times to maturity and for put options. Overall, the average pricing errors are higher for all distributions during extremely volatile periods. However, the relative performance of the GEV option-pricing model becomes much superior to the lognormal distribution during the extreme movements of the market.

8. Conclusion

This paper proposes a new approach to estimating interest-rate volatility during highly volatile periods. A major issue in estimating interest-rate volatility is the modeling of the distribution of interest rates. Most of the existing models, which use a method based on a parametric distribution, assume that the interest rates are normally distributed. However, this study
clearly shows that the tails of the empirical distribution are much thicker than the tails of the normal distribution. The method proposed here estimates the term structure of interest-rate volatility based on the extreme changes in 3-month, 6-month, 1, 2, 3, 5, 7, 10, and 30-year Treasury yields. The empirical results indicate that the volatilities of the maximal and minimal changes in interest rates decline as time to maturity rises, yielding a downward-sloping volatility curve for the extremes. For all data sets and for all asymptotic distributions used, the volatility curve for the minima plots above the volatility curve for the maxima, implying that the minimal changes are more volatile than the maximal changes for all times to maturity.

The commonly used cap-floor pricing models are based on the assumption that forward rates follow the geometric Brownian motion. A key assumption of the geometric Brownian motion is that the ratio of two consecutive forward rates does not depend on past forward rates. Furthermore, the natural logarithm of two consecutive forward rates is far from normal, and the fat-tail property is more dominant than skewness. Therefore, the distribution of spot and forward interest rate process. Therefore, the conditional variance of the geometric Brownian motion is that the ratio of two consecutive forward rates does not depend on past forward rates. Furthermore, the natural logarithm of two consecutive forward rates is assumed to be normally distributed with the same annualized mean and variance over time. However, this paper provides strong evidence for the presence of serial correlation and extreme conditionally heteroscedastic volatility effects in the spot and forward interest-rate process. Therefore, the conditional variance of the forward-rate process, which is a key determinant of cap-floor prices, cannot be assumed as i.i.d. and constant over time. The nature of interest-rate option-pricing models also necessitates the use of probability distributions, which provide a good fit to the empirical distribution of spot and forward interest rates. Our results point out that the distribution of log-forward rates is far from normal, and the fat-tail property is more dominant than skewness. Therefore, we extend the standard option-pricing models with lognormal forward rates to accommodate significant leptokurtosis observed in the interest rate data. We introduce a closed-form option-pricing model based on the GEV distribution, and show that the newly proposed model provides more accurate predictions of actual option prices than the lognormal and generalized error distributions.

9. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org/.

Acknowledgments

The author thanks David Hsieh (the department editor), the associate editor, and two anonymous referees for their extremely helpful comments and suggestions. The author also benefited from discussions with Torben Andersen, Salih Neftci, Kenneth Singleton, David Weinbaum, and Liuren Wu. The suggestions of Linda Allen, John Merrick, Lin Peng, and Robert Schwartz have been very valuable. Financial support from the Eugene Lang Research Foundation of the Baruch College and PSC-CUNY Research Foundation of CUNY is gratefully acknowledged. All errors remain the author’s responsibility.

Appendix. Closed-Form Solution for the European Put Option Price with GEV

The derivation of the closed-form solution for the European put option price equation is similar to the derivation for the call option price equation. When applying the change of variable defined in (31) to the put option price equation, after having substituted for the price RND function \( g(S_T) \) in (28), the upper limit of integral \( X \) in the put option equation becomes \( \Pi \) as defined in (33), while the lower limit of the integral in the put option equation becomes \( \omega = 1 + (\xi/\sigma)(1 - \mu) > 0 \). Under these new limits of the integral we have

\[
P_e^{c(T-t)} = \frac{1}{\xi} \left[ \int_0^{q(x^{1/\xi})} g(q^{1-1/\xi}) \exp(-q^{-1/\xi}) dq \right] 
- \left( X - S_t \left( 1 - \mu - \frac{\sigma}{\xi} \right) \right) 
\cdot \left( \int_0^{q(x^{1/\xi})} \exp(-q^{-1/\xi}) dq \right) 
= \frac{1}{\xi} \left[ \int_0^{q(x^{1/\xi})} \exp(-q^{-1/\xi}) dq \right] 
\cdot \left( X - S_t \left( 1 - \mu - \frac{\sigma}{\xi} \right) \right) 
\phi_1 \right]. \quad (48)
\]

Evaluating the first integral in (48) yields

\[
\phi_1 = \int_0^{q(x^{1/\xi})} \exp(-q^{-1/\xi}) dq = \frac{\xi}{\sigma} (e^{-\omega^{-1/\xi}} - e^{-x^{-1/\xi}}). \quad (49)
\]

To solve the second integral in (48), consider the change of variable \( x = q^{-1/\xi} \) and \( q = x^{-\xi} \), \( dq = -\xi x^{-\xi-1} dx \), which yields

\[
\phi_2 = \int_{\omega^{-1/\xi}}^{1} x^{-\xi} e^{-x^{-\xi}} d(x^{-\xi}) = -\xi \int_{\omega^{-1/\xi}}^{1} x^{-\xi} e^{-x^{-\xi}} dx 
= -\xi \int_{\omega^{-1/\xi}}^{1} x^{(1-\xi)-1} e^{-x} dx. \quad (50)
\]

We can solve this integral by using the definition of the generalized gamma function (see the online supplement):

\[
\Gamma(a, y_0, y_1) = \Gamma(a, y_0) - \Gamma(a, y_1) = \int_{y_0}^{y_1} x^{a-1} e^{-x} dx, \quad (51)
\]

which yields

\[
\phi_2 = \int_0^{q(x^{1/\xi})} \exp(q^{-1/\xi}) dq = -\xi \Gamma(1 - \xi, \omega^{-1/\xi}, \Pi^{-1/\xi}). \quad (52)
\]

Substituting \( \phi_1 \) and \( \phi_2 \) into (48), we obtain a closed-form solution for the GEV put option price:

\[
P_e = e^{-r(T-t)} \left[ X (e^{-\omega^{-1/\xi}} - e^{-\Pi^{-1/\xi}}) 
- S_t \left( 1 - \mu - \frac{\sigma}{\xi} \right) (e^{-\Pi^{-1/\xi}} - e^{-\omega^{-1/\xi}}) 
- \frac{\sigma}{\xi} \Gamma(1 - \xi, \omega^{-1/\xi}, \Pi^{-1/\xi}) \right]. \quad (53)
\]
References


