ELSEVIER

Journal of Empirical Finance 10 (2003) 455–477

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

Disturbing extremal behavior of spot rate dynamics

Turan G. Bali[a,*], Salih N. Neftci[b,1]

[a] Department of Economics and Finance, Baruch College, Zicklin School of Business, City University of New York, 17 Lexington Avenue, Box 10-225, New York, NY 10010, USA

[b] Graduate School, City University of New York, Fifth Floor, 365 Fifth Avenue, New York, NY 10016-4309, USA

Abstract

This paper presents a study of extreme interest rate movements in the US federal funds market over almost a half century of daily observations from the mid 1950s through the end of 2000. We analyze the fluctuations of the maximal and minimal changes in short-term interest rates and test the significance of time-varying paths followed by the mean and volatility of extremes. We formally determine the relevance of introducing trend and serial correlation in the mean and of incorporating the level and GARCH effects in the volatility of extreme changes in the federal funds rate. The empirical findings indicate the existence of volatility clustering in the standard deviation of extremes and a significantly positive relationship between the level and the volatility of extremes. The results point to the presence of an autoregressive process in the means of both local maxima and local minima values. The paper proposes a conditional extreme value approach to calculating value at risk (VaR) by specifying the location and scale parameters of the generalized Pareto distribution (GPD) as a function of past information. Based on the estimated VaR thresholds, the statistical theory of extremes is found to provide more accurate estimates of the rate of occurrence and the size of extreme observations.

JEL classification: G12; C13; C22
Keywords: Extreme value theory; Volatility; Interest rates; Value at risk

1. Introduction

The key role that the short-term risk-free interest rate plays in the valuation of almost all securities has made it one of the most critical variables to model in financial economics. For

* Corresponding author. Tel.: +1-646-312-3506; fax: +1-646-312-3451.
E-mail addresses: Turan_Bali@baruch.cuny.edu (T.G. Bali), sneftci@gc.cuny.edu (S.N. Neftci).
[1] Tel.: +1-212-817-8261; fax: +1-212-817-1514.

0927-5398/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0927-5398(02)00070-1

the last two decades, an enormous amount of work has been directed towards modeling and estimation of the interest rate process.[2] However, most of the empirical studies concern average properties like the conditional mean, volatility, or correlations, and very little attention has been given to the extreme movements themselves.[3] Almost no empirical study has been presented to analyze the extreme changes in the mean and volatility of the short rate process. The present study, to our knowledge, is the first to examine the stochastic behavior of short-term interest rates in terms of extreme values. We estimate the asymptotic distribution of extreme daily changes in short rates and then test the presence of level and GARCH effects in the volatility of extremes. In addition, we test the presence of trend and autoregressive process in the means of the maximal and minimal values.

There may be two reasons why a decision maker would be interested in the "extremes" of a market variable (or risk factor). One is related to insolvency and to properties of the underlying utility functions, and the other is related to the functioning of financial markets. In the first case, the behavior towards risk taking may be sharply nonlinear. Aversion to market risk may be mild, but aversion to risks associated with insolvency-type extreme risks may be very high. When this is combined with practical constraints that require asymmetric treatment of upside potential and downside risk, it may put special emphasis on extreme movements.[4] The other possibility is more technical but, nevertheless, potentially as important. Funds managers report two types of realized returns. One is average quarterly realized returns, and the other reported by hedge funds often on demand is the *maximum drawdown*, which is in fact the *extreme* loss for the life of the hedge fund. Now, in a world where these are the only information concerning the true continuous time returns, extreme values will carry additional information about the risk taken by hedge fund managers. In fact, it is not difficult to see that the maximum drawdown may indeed be more informative of the two measures concerning the practices of the hedge fund.[5]

We assume that during the extreme price or interest rate movements in financial markets (like interest rate or stock market booms and crashes), what is of interest to market

[2] For modeling the conditional distribution of interest rates using parametric or nonparametric approaches, see Chan et al. (1992), Longstaff and Schwartz (1992), Ait-Sahalia (1996a,b), Brenner et al. (1996), Koedijk et al. (1997), Conley et al. (1997), Stanton (1997), Andersen and Lund (1997), and Bali (2000) among others.

[3] Although limited work has been devoted to extreme movements of stock prices, surprisingly little effort was spent on empirical studies analyzing extreme interest rate changes. Notable exceptions that use extreme value theory to explain the fluctuations in financial markets are Rothschild and Stiglitz (1970), Parkinson (1980), Neftci (1985), and Jansen and de Vries (1991). The most recent work in this area has been performed by Longin (1996, 2000), Kearns and Pagan (1997), McCulloch (1997), Danielsson and de Vries (1997a,b), Neftci (2000), McNeil and Frey (2000), Jansen et al. (2000), and Bali (in press).

[4] For example, Gollier et al. (1996) and Santomero and Babbel (1996) consider empirical models of banks and insurance companies that maximize a utility function subject to a VaR-type solvency constraint.

[5] As discussed by Gourieroux et al. (2000) and Jansen et al. (2000), there is growing interest on the economic foundations of extreme value theory and its applications to value at risk (VaR) and portfolio selection.

Optimal asset allocation in a traditional mean-variance portfolio theory is based on a trade-off between expected return and risk that is generally measured by the volatility of asset returns. Alternatively, as shown by Jansen et al. (2000), the risk can be based on a safety-first criterion (probability of failure). Jansen et al. (2000) use extreme value theory to study the optimal portfolio selection under limited downside risk as an alternative to standard mean-variance efficient frontiers. Similarly, Gourieroux et al. (2000) modify the mean-variance efficient portfolio by using VaR as a risk measure instead of the volatility. In their model, the VaR efficient portfolio selection depends on the loss probability in the context of capital adequacy requirement.

participants is not how much the interest rates change from 1 day to another, or what happens to this daily volatility. Instead, they are concerned with the mean and volatility of interest rate changes at the tails of the distribution. We make the point that it is quite possible for the standard deviation of a short rate process to "dampen" over time, while the volatility of extreme movements increases. Which "volatility" is more relevant to pricing interest rate sensitive claims and to estimating catastrophic market risks during extraordinary periods? This question deserves asking. In this study, we illustrate these two ways of looking at short rate volatility using extreme value theory (EVT).

The paper also deals with how extremes in the interest rate process can be used to analyze changes in volatility. It is our belief that volatility dynamics is directly related to the changing parameters of extremal distributions instead of some stochastic variation in the variance of the short rate process. That is to say, volatility dynamics can be captured to a large extent by a parsimonious extremal distribution theory, albeit with time-varying parameters. To this end, we basically examine the volatility of maximal (minimal) daily changes in the short rate and test the time variation followed by the parameters of the asymptotic extremal distributions. This is a difficult problem since if the parameters are time varying, then the convergence may not be to asymptotic extremal distributions. In fact, the issue of asymptotic convergence of extreme value distributions with time-varying parameters has not yet been investigated by the former researchers, and it is beyond the scope of this paper. The present study is a first attempt towards detecting any time variation in extreme value distributions. Essentially, our results are based on the fact that under the null hypothesis of no time variation the distributions are known. We use these distributions to build the maximum likelihood estimation model.

First, we consider the extremes of a stochastic process and then estimate trends on these to see if the trends are significantly positive or negative. We specify the *location* parameter (i.e., the mean of extreme interest rate changes) of the asymptotic distribution as a function of time and then test whether the mean of the extremes is constant or follows a time-dependent path. Second, we test if the mean of the extremes is a function of past information by specifying an autoregressive process for the location parameter. We estimate the parameters of the linear AR(1) specification of the mean and then determine if the last local extrema comprise adequate information, which can be used to explain the behavior of the current local extrema. Third, we estimate the *scale* parameter of the generalized Pareto distribution (GPD), which is a measure of the volatility of extremes, and then test if it depends on the level of extreme values. In other words, we test the presence of level effect in the volatility of extreme interest rate changes. Finally, the volatility parameter is parameterized as a function of the last period's unexpected news and the last period's volatility to determine the presence of GARCH effects in the standard deviation of extremes.

In recent years, a central issue in risk management has been to determine capital requirement for financial institutions to meet catastrophic market risk.[6] The standard

[6] This increased focus on risk management has led to the development of various methods and tools to measure the risks financial institutions face. A primary tool for financial risk assessment is the value at risk (VaR), which is defined as the maximum loss expected on a portfolio of assets over a certain holding period at a given confidence level (probability).

volatility models can be successful in estimating the maximum likely loss of an institution *under normal market conditions*, which corresponds to the normal functioning of financial markets during ordinary periods. However, as pointed out by Longin (2000) and Bali (in press), the volatility measures based on the distribution of all returns cannot produce accurate estimates of market risks during highly volatile periods. We, therefore, think that an alternative approach that well approximates the tails of the empirical distribution may be more appropriate in estimating the conditional volatility of risk factors and value at risk (VaR) during extraordinary periods. In order to improve the existing unconditional EVT-based methods, we propose a *conditional* extreme value approach to estimating VaR by specifying the mean and volatility parameters of the extreme value distribution as a function of past information.

The paper is organized as follows. Section 2 presents a framework for the asymptotic distribution of extremes. Section 3 provides the maximum likelihood estimation methodology for the extremes. Section 4 describes the data. Section 5 presents the empirical results. Section 6 discusses risk management performance of the extreme value approach. Section 7 concludes.

2. The framework

The market practice for pricing a very large majority of financial derivatives requires that the spot rate be modeled as a continuous time mean-reverting process. Modeling of the diffusion component of such processes plays a crucial role in the pricing effort. All continuous time interest models used in practice do this by imposing quite restrictive conditions on the stability and smoothness properties of the diffusion components. In particular, all these approaches imply that the behavior of extremes of the continuous time spot rates is stable over time. The notion of extremes is clearly at the center of pricing risk, and the implication that spot rate has a stable extremal (asymptotic) distribution is essential for current pricing models to be correct.

Yet, in this paper, we show that when we place ourselves within the extremal theory of continuous time stochastic processes, the extremes of the interest rate may have a much more complicated dynamics than the one assumed by the current models. In fact, we see that although the extremal theory fits well, the parameters of the extremal distribution are not constant and depend on the sample path properties of the spot rate.

The extremal theory for continuous time stochastic processes is similar to the extremal theory for i.i.d. sequences. In this latter case, we let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. nondegenerate random variables with cumulative distribution function (cdf) $F(x)$ and investigate the fluctuations of the sample maxima (minima) of the sequence $\{X_1, X_2, \ldots, X_n\}$, where

$$M_1 = X_1,\ M_2 = \max(X_1, X_2), \ldots,\ M_n = \max(X_1, \ldots, X_n), \qquad n \geq 2. \tag{1}$$

Corresponding results for minima can be obtained from those for maxima by using the identity:

$$\min(X_1, \ldots, X_n) = -\max(-X_1, \ldots, -X_n). \tag{2}$$

The exact cdf of the maximum M_n is easy to write:

$$P(M_n \leq x) = P(X_1 \leq x, \ldots, X_n \leq x) = F^n(x), \qquad x \in \Re, \qquad n \in N. \tag{3}$$

According to this, extremes happen near the upper end of the support of the distribution, hence, intuitively, the asymptotic behavior of M_n must be related to the cdf $F(x)$ in its right tail. In this case, this tail has finite support. We let

$$x_F = \sup\{x \in \Re:\ F(x) < 1\} \tag{4}$$

denote the right endpoint of $F(x)$. We immediately obtain, for all $x < x_F$,

$$P(M_n \leq x) = F^n(x) \rightarrow 0, \qquad n \rightarrow \infty, \tag{5}$$

and, in the case $x_F < \infty$, we have, for $x \geq x_F$,

$$P(M_n \leq x) = F^n(x) = 1. \tag{6}$$

Thus, $M_n \xrightarrow{P} x_F$ in probability as $n \rightarrow \infty$, where $x_F \leq \infty$. To obtain an asymptotic distribution theory, we need to look at the convergence in distribution of the *centered* and *normalized* maxima.[7] Here, the well-known Fisher-Tippett (1928) theorem has the following content: if there exist normalizing constants $\sigma_n > 0$ and centering constants $\mu_n \in \Re$ such that

$$\sigma_n^{-1}(M_n - \mu_n) \xrightarrow{d} G, \quad n \rightarrow \infty, \tag{7}$$

for some nondegenerate distribution, G, then G belongs to the type of one of the three so-called standard extreme value distributions: Frechet, Gumbel, and Weibull (see Embrechts et al., 1997).

All this is the theory concerning the extremes of an i.i.d. sequence. Yet, the extremal theory for stochastic processes eventually boils down to the same sets of results, albeit under quite different conditions. In our case, given that the valuation of interest sensitive securities is done in a setting where the spot rate follows a mean-reverting diffusion, we need to discuss how this extension is done for the extremes of a process r_t that follows the stochastic differential equation:

$$\mathrm{d}r_t = a(r_t)\mathrm{d}t + b(r_t)\mathrm{d}W_t \qquad t \in [0, \infty] \tag{8}$$

where W_t is a standard Wiener process, and the drift and diffusion parameters, $a(.)$ and $b(.)$, satisfy the usual integrability conditions. To obtain an asymptotic distribution for the

[7] The true small-sample distribution of an *extreme value* is not known. What the extreme value theory does is substitute for it an asymptotic distribution that will be a good approximation if the sample of extremes is large. This weak convergence notion is in fact a refinement of the Central Limit Theorems, and the refinement occurs at the tails. The basic reference for how quickly the sample distribution would converge to the asymptotic distributions is Leadbetter et al. (1983) and the references therein.

extremes of this continuous time process, we select a small finite interval Δ>0 and define the equally distant time periods:

$$t_0 = 0, t_1 = \Delta, \ldots, t_i = i\Delta, \ldots \tag{9}$$

Now, let

$$x_{t_i} = \max\{r_t \colon (i-1)\Delta \le t \le i\Delta\} \tag{10}$$

then, under the assumption that the *jumps in the density of* x_t *die-out sufficiently quickly* relative to the tail, we get a nondegenerate convergence as in the i.i.d. case (see Leadbetter et al., 1983, p. 4):

$$P(\sigma_n^{-1}(M_n - \mu_n)) \rightarrow G \tag{11}$$

where the maximal changes, M_n, over n trading days is now defined as:

$$M_n = \max\{x_1, x_2, \ldots, x_n\} \tag{12}$$

and where G is a nondegenerate extremal distribution. On the condition that the process is strongly mixing G will be one of the three standard types of extreme value distributions.

M_n in Eq. (12) is the maximal (or the highest) daily changes in short-term interest rates reached over n trading days (or over time periods of different lengths: 1 month, 1 quarter, 1 year, etc.). This is a standard way of obtaining the extremes for the generalized extreme value (GEV) distribution of Jenkinson (1955) that nests the Frechet, Gumbel, and Weibull distributions (see Embrechts et al., 1997, pp. 120–121). The GEV distribution describes the limit distributions of normalized maxima and minima. To test the significance of time-varying paths followed by the mean and volatility of extremes, we chose n in our empirical work as 1 week, 2 weeks, 1 month, and 1 quarter. The results turn out to be robust across different time periods. Since our emphasis is on conditional value at risk and the excesses over high thresholds, we do not present the parameter estimates from the GEV distribution in order to save space. They are available from the authors upon request.

Bali (in press) indicates that the generalized Pareto distribution (GPD) of Pickands (1975) provides more accurate estimates of the actual VaR thresholds than the generalized extreme value (GEV) distribution of Jenkinson (1955). Therefore, we use the generalized Pareto distribution (GPD) and define the extremes as *excesses over high thresholds* (see Embrechts et al., 1997, pp. 352–355). The key idea of this alternative approach to estimate the asymptotic distribution of extremes can be explained as follows. First, we choose a high threshold u so that all r_i>u>0 are defined to be in the positive tail of the distribution, where r_1,r_2,. . .,r_n are a sequence of interest rate changes on days 1, 2,. . ., n. Then we denote the number of exceedences of u by

$$k_u = \mathrm{card}\{i \colon\ i = 1, \ldots, n,\ r_i > n\}, \tag{13}$$

and the corresponding excesses by M_1,M_2,. . .,M_{k_u}. The excess distribution function of r is given by:

$$F_u(y) = P(r - u \le y \mid r > u) = P(M \le y \mid r > u), \quad y \ge 0 \tag{14}$$

Using the threshold u, we now define the probabilities associated with r:

$$P(r \leq u) = F(u) \tag{15}$$

$$P(r \leq u + y) = F(u + y) \tag{16}$$

where $y>0$ is an exceedence of the threshold u. Finally, let $F_u(y)$ be given by

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)} \tag{17}$$

We thus obtain the $F_u(y)$, the conditional distribution of *how* extreme r_i is, given that it already qualifies as an extreme. Pickands (1975) shows that $F_u(y)$ will be very close to the generalized Pareto distribution $H_\xi(x)$ given in Eq. (18) if u is a high threshold:

$$H_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp(-x) & \xi = 0 \end{cases} \tag{18}$$

where

$$x = \frac{M_u - \mu_u}{\sigma_u} \geq 0 \qquad \text{if } \xi \geq 0 \tag{19}$$

$$0 < x = \frac{M_u - \mu_u}{\sigma_u} \leq -\frac{1}{\xi} \qquad \text{if } \xi < 0 \tag{20}$$

Notice that the generalized Pareto distribution presented in Eq. (18) encompasses the standard *Pareto* distribution, the *uniform* distribution on $[-1,0]$, and the standard *exponential* distribution with $\xi = 0$.[8] In Eq. (18), the shape parameter, ξ, determines the tail behavior of the distributions. For $\xi>0$, the Pareto distribution has a polynomially decreasing tail. For $\xi = 0$, the tail of the exponential distribution decreases exponentially. For $\xi < 0$, the distribution is short tailed.

[8] The extreme value distributions nested within the generalized Pareto distribution for the standardized maximum are as follows:

$$\begin{aligned} &\text{Pareto}: && H_1(x) = 1 - x^{-1/\xi} && \text{for } x \geq 1, \\ &\text{Uniform}: && H_2(x) = 1 - (-x)^{1/\xi} && \text{for } x \in [-1, 0] \\ &\text{Exponential}: && H_3(x) = 1 - \exp(-x) && \text{for } x \geq 0. \end{aligned}$$

3. Maximum likelihood methodology for the extremes

Our setup corresponds to the standard parametric case of statistical inference, and hence, maximum likelihood methodology can be utilized in a straightforward fashion. Suppose that $H_\Omega(x)$ has density function $h_\Omega(x)$, where $\Omega=(\xi, \mu, \sigma)\in\Re\times\Re\times\Re_+$ consists of a *shape* parameter ξ, *location* parameter μ, and *scale* parameter σ. Then the likelihood function based on the data $\boldsymbol{X}=(M_1, M_2,\ldots, M_k)$ is given by

$$L(\Omega;\boldsymbol{X}) = \prod_{i=1}^{k} h_\Omega(M_i). \tag{21}$$

Denote the log-likelihood function by $l(\Omega;\boldsymbol{X})=\ln L(\Omega;\boldsymbol{X})$. The maximum likelihood estimator (MLE) for Ω then equals

$$\hat{\Omega}_k = \arg\max_{\Omega\in\Theta} l(\Omega;\boldsymbol{X}), \tag{22}$$

where $\hat{\Omega}_k=\hat{\Omega}_k(M_1, M_2,\ldots, M_k)$ maximizes $l(\Omega;\boldsymbol{X})$ over an appropriate parameter space Θ.

The generalized Pareto distribution presented in Eq. (18) has the density function

$$h(x) = \begin{cases} \left(\frac{1}{\sigma}\right)(1+\xi x)^{-(1+\xi)/\xi}, & \xi\neq 0 \\ \left(\frac{1}{\sigma}\right)\exp(-x), & \xi = 0 \end{cases} \tag{23}$$

which yields the following log-likelihood function:

$$\ln L(\Omega;\boldsymbol{X}) = -k\ln\sigma - k\left(\frac{1+\xi}{\xi}\right)\sum_{i=1}^{k}\ln\left[1+\xi\left(\frac{M_i-\mu}{\sigma}\right)\right], \tag{24}$$

with a nonzero shape parameter ξ. Differentiating the log-likelihood function given in Eq. (24) with respect to Ω 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.[9]

Our main interest is in the volatility of interest rate changes in terms of extreme values. Note that the μ and σ parameters account for the first- and second-order moments of the short rate implied by the extremes only. In fact, if the short-term interest rate is normally distributed, the parameters of the extreme value distribution will coincide with the standard estimates of μ and σ.

We test the significance of time trend and the adequacy of last local extrema in explaining the stochastic behavior of the current local extrema. To test the significance of

[9] Jenkinson (1969) and Prescott and Walden (1980) suggest variants of the Newton–Raphson scheme to solve a set of nonlinear equations given by the first-order conditions.

time dependency and AR(1) specification in the asymptotic distribution of extremes, we first assume that both location and scale parameters of the generalized Pareto distribution are constant. Then, under the alternative hypothesis H_a, we parameterize the location parameter as a linear function of time,

$$\mu_{t_i} = \mu + \alpha t_i, \tag{25}$$

and test the statistical significance of α.

Similarly, to determine whether incorporating AR(1) process in the mean of the extremes improves the predictability of current local maxima or minima, we specify the location parameter as a linear function of the last local extrema,

$$\mu_{t_i} = \mu + \beta M_{t_{i-1}}, \tag{26}$$

and test whether the coefficient, β, on the last local extrema, $M_{t_{i-1}}$ is statistically significant.

More importantly, we test the presence of level effect in the volatility of extreme changes in short rates by defining the scale parameter as a function of the level of the last local maxima,[10]

$$\sigma_{t_i} = \sigma M_{t_{i-1}}^{\gamma}, \tag{27}$$

where γ determines the sensitivity of current volatility to the level of extremes. We test whether γ equals zero to see if the constant volatility assumption is appropriate to use in explaining the stochastic behavior of extremes. When $\gamma = 0$ in Eq. (27), the level effect disappears, and we are back to the original framework with constant volatility.

The literature on time-varying volatility models of the short-term interest rate provides enough evidence that the conditional variance of interest rate changes is very sensitive to the last period's unexpected news. This paper tests the presence of GARCH effect in the volatility of the extremes by parameterizing the standard deviation of the current maxima as a function of the last period's unexpected news as well as the last period's standard deviation:

$$\sigma_{t_i} = \sigma + \lambda_1 |\varepsilon_{t_{i-1}}| + \lambda_2 \sigma_{t_{i-1}}, \tag{28}$$

where $\varepsilon_{t_i} = M_{t_i} - \mu_{t_i}$ measures the deviation of the predicted maxima, μ_{t_i}, from the actual, M_{t_i}, and can thus be viewed as an unexpected information shock to the interest rate market during its largest falls and rises. We test the hypothesis $\lambda_1 = \lambda_2 = 0$ to determine whether the standard deviation of the extremes is constant or accommodates volatility clustering. Eqs. (25)–(28) present various alternative hypotheses H_a that we consider. The important point

[10] We measure the sensitivity of the volatility of extremes to the *absolute* level of the last local *minima* by defining the scale parameter of the GPD as $\sigma_{t_i} = \sigma |M_{t_{i-1}}|^{\gamma}$, where $|M_{t_{i-1}}|$ is the absolute value of the last local minima. minima.

is that under the null hypotheses, the underlying distributions become GPD asymptotically.[11]

4. Data

The data set used in this study is obtained from the Federal Reserve H.15 database and consists of the US federal funds interest rates from July 1, 1954 to December 29, 2000, giving a total of 11,698 daily observations.[12] Table 1 displays the means, standard deviations, maximum and minimum values of the levels, and first-differences of the federal funds rate. The unconditional average level of the fed funds rate is about 6.10% with a standard deviation of 3.36%, whereas the daily change averaged 0.0004% with a standard deviation of 0.46%. The maximum and minimum values of the level are 22.36% and 0.13%, while these statistics for the interest rate changes equal 7.79% and − 7.89%.

In this paper, local maxima and local minima values are obtained from the original daily data described above. Following the extreme value theory for the generalized Pareto distribution, we define the extremes as excesses over high thresholds (see Embrechts et al., 1997, pp. 352–355). Specifically, the extreme changes are defined as those more than two standard deviations away from the sample mean of daily interest rate changes, which correspond to 2.5% of the right and left tails of the distribution. Table 2 shows the number of extremes $k=292$ ($k=0.025\times 11{,}697$), means, standard deviations, and the maximum

[11] Gradshteyn and Ryzhik (1994, p. 341) show that

$$\int_0^{+\infty} x^r(1+qx^k)^{-m}\mathrm{d}x = k^{-1}\times q^{-(r+1)/k}\times\Gamma[(r+1)/k]\times\Gamma[m-(r+1)/k]\times\Gamma(m)^{-1}.$$

Letting $x=M-\mu$, $q=\varepsilon/\sigma$, $k=1$, $m=(1+\varepsilon)/\varepsilon$, and $r=1$ gives the expected value of the maxima minus μ:

$$E(M-\mu)=E(x)=\left(\frac{1}{\sigma}\right)\int_0^{+\infty} x[1+(\varepsilon/\sigma)x]^{-(1+\varepsilon)/\varepsilon}\mathrm{d}x=\left(\frac{\sigma}{\varepsilon^2}\right)\frac{\Gamma(2)\times\Gamma[(1-\varepsilon)/\varepsilon]}{\Gamma[(1+\varepsilon)/\varepsilon]}$$

where $\left(\frac{1}{\sigma}\right)\left[1+\varepsilon\left(\frac{M-\mu}{\sigma}\right)\right]^{-(1+\varepsilon)/\varepsilon}$ is the probability density function (pdf) of the generalized Pareto distribution. The above integral yields the mean of the extremes: $E(M)=\mu+\left(\frac{\sigma}{\varepsilon^2}\right)\frac{\Gamma[(1-\varepsilon)/\varepsilon]}{\Gamma[(1+\varepsilon)/\varepsilon]}$ that depends on the location, scale, and shape parameters of the GPD. Since the term $\left(\frac{\sigma}{\varepsilon^2}\right)\frac{\Gamma[(1-\varepsilon)/\varepsilon]}{\Gamma[(1+\varepsilon)/\varepsilon]}$ in $E(M)$ is trivial compared to μ in terms of their magnitudes, we assume that the location parameter, μ, captures most of the variation in the mean of the extremes. Therefore, the time-varying behavior of the conditional mean of the extremes is modeled by μ_{t_i}. Letting $x=M-\mu$, $q=\varepsilon/\sigma$, $k=1$, $m=(1+\varepsilon)/\varepsilon$, and $r=2$ gives

$$E(M-\mu)^2=E(x^2)=\left(\frac{1}{\sigma}\right)\int_0^{+\infty} x^2[1+(\varepsilon/\sigma)x]^{-(1+\varepsilon)/\varepsilon}\mathrm{d}x=\left(\frac{\sigma^2}{\varepsilon^3}\right)\frac{\Gamma(3)\times\Gamma[(1-2\varepsilon)/\varepsilon]}{\Gamma[(1+\varepsilon)/\varepsilon]}.$$

Then, the variance of M is a function of σ and ε: $E[M-E(M)]^2=\dfrac{\sigma^2\left[2\varepsilon\Gamma\left(\frac{1-2\varepsilon}{\varepsilon}\right)-\Gamma\left(\frac{1-\varepsilon}{\varepsilon}\right)^2\right]}{\varepsilon^4\Gamma\left(\frac{1+\varepsilon}{\varepsilon}\right)^2}$.

We should note that the time-varying behavior of the standard deviation of extremes is captured by σ_{t_i} because the shape parameter, ε, is assumed to be constant in the paper.

[12] Federal funds are typically overnight loans between banks of their deposits at the Federal Reserve. The federal funds market is very sensitive to the credit needs of the banks, so the interest rate on these loans, called the federal funds rate, is a closely watched barometer of the tightness of credit market conditions in the banking system and the stance of monetary policy; when it is high, it indicates that the banks are strapped for funds, whereas when it is low, the banks' credit needs are low.

Table 1
Descriptive statistics (July 1, 1954 to December 29, 2000)

Federal funds rate	n	Mean	Standard deviation	Maximum	Minimum
r_t	11,698	0.060954	0.0336	0.2236	0.0013
$r_t - r_{t-1}$	11,697	0.000004	0.0046	0.0779	− 0.0789

This table shows the means, standard deviations, maximum, and minimum values of the level, r_t, and daily changes, $r_t - r_{t-1}$, in the federal funds rate. The daily data are obtained from the Federal Reserve H.15 database over the period from July 1, 1954 to December 29, 2000, giving a total of 11,698 daily observations. The sample periods and the number of observations (n) are specified above.

and minimum values of the extremes. In addition to the 2.5% tails, the extremes are obtained from the 5% and 10% tails of the empirical distribution. Since the qualitative results are found to be robust across different threshold levels, we choose not to present the maximum likelihood estimates based on the 5% and 10% tails.[13]

5. Empirical results

Table 3 presents the estimated parameters of the GPD, asymptotic t-statistics, and the maximized log-likelihood values of the Constant Mean–Constant Volatility model. According to the asymptotic t-statistics, all of the estimated GPD parameters are statistically significant at the 1% or 5% level. The estimated shape parameter for the local maxima ($\xi_{\max} = 0.2722$) is greater than that for the local minima ($\xi_{\min} = 0.2146$). Since the higher ξ, the fatter the distribution of extremes, the distribution of maximal changes has thicker tails than that of minimal changes. The volatility measured by the estimated scale parameter in the constant volatility specification turns out to be around 0.004, which is very close to the unconditional standard deviation ($\sigma = 0.005$) of daily changes in the federal funds rate shown in Table 1.

When we allow the volatility to depend on the level of extreme values, as in the case of the Constant Mean–Level Volatility model with nonzero sensitivity parameter ($\gamma \neq 0$), the estimated shape parameters, ξ, drop slightly for both the local maxima and local minima. For the Constant Mean–Level Volatility specification, the sensitivity parameters turn out to be less than one and highly significant. The current volatility of local maxima is found to be more sensitive to the level of extremes (with $\gamma_{\max} = 0.6886$) than that of the local minima (with $\gamma_{\min} = 0.4291$). Positive estimated values of γ imply that the volatility of extremes is high when the level of local maxima or the absolute value of local minima is high.

When the current volatility, σ_{t_i}, is defined as a function of the last period's unexpected shocks, $\varepsilon_{t_{i-1}}$, and the last period's volatility, $\sigma_{t_{i-1}}$, as in the case of the Constant Mean–GARCH Volatility model, the shape parameter is estimated to be somewhat higher ($\xi_{\max} = 0.2827$ and $\xi_{\min} = 0.2459$) than in the case of the Constant Mean–Level Volatility model. The maximum likelihood estimates of the GARCH parameters indicate the presence of volatility clustering in the standard deviation of extremes since λ_1 and λ_2 are found to be

[13] To save space we decide not to discuss the empirical results obtained from the 5% and 10% tail estimates. They are available upon request.

Table 2
Summary statistics of extremes

Federal funds rate	k	Mean	Standard deviation	Maximum	Minimum
Maxima	292	0.01662	0.00813	0.0779	0.0100
Minima	292	– 0.01479	0.00756	– 0.0087	– 0.0789

This table displays the number of extremes k ($= 0.025 \times n$), means, standard deviations, maximum, and minimum values of the extremes, which are obtained from the daily federal funds rate database. Extreme changes are defined as those more than two standard deviations away from the sample mean of daily interest rate changes, which correspond to 2.5% of the right and left tails of the empirical distribution.

highly significant for both the maximal and minimal changes. The current volatility of local maxima and minima turns out to be more sensitive to the last period's volatility than to the last period's unexpected news in the interest rate market since λ_2 is estimated to be greater than λ_1. Another important point is that volatility shocks persist longer in local minima than in local maxima because the sum of the GARCH parameters for the minimal changes ($\lambda_1 + \lambda_2 = 0.9094$) is greater than that for the maximal changes ($\lambda_1 + \lambda_2 = 0.8113$).

Table 4 displays the maximum likelihood estimates of the asymptotic distribution with time-dependent location parameter, asymptotic t-statistics, and the maximized log-likelihood values. The estimated shape parameter again turns out to be higher for the local

Table 3
Constant Mean

	μ	σ	ξ	λ_1	λ_2	γ	log$L$
Maxima GARCH Volatility	0.0125 (56.824)	0.00062 (2.0944)	0.2867 (2.8836)	0.0851 (2.2409)	0.7262 (7.2234)	–	1253.72
Minima GARCH Volatility	– 0.0113 (– 54.958)	0.00033 (2.5919)	0.2459 (3.1183)	0.0483 (2.1708)	0.8611 (15.458)	–	1281.93
Maxima Level Volatility	0.0130 (54.949)	0.0714 (1.3751)	0.2131 (2.6354)	–	–	0.6886 (3.6327)	1252.30
Minima Level Volatility	– 0.0119 (– 53.747)	0.0311 (1.2993)	0.2012 (2.9064)	–	–	0.4291 (2.4219)	1281.18
Maxima Constant Volatility	0.0130 (53.784)	0.00388 (8.9575)	0.2722 (2.9796)	0.0	0.0	0.0	1245.64
Minima Constant Volatility	– 0.0119 (– 51.886)	0.00366 (9.1241)	0.2146 (3.1108)	0.0	0.0	0.0	1277.65

This table presents the maximum likelihood estimates of the location (μ), scale (σ), sensitivity (γ), GARCH (λ_1, λ_2), and shape (ξ) parameters of the generalized Pareto distribution. Asymptotic t-statistics are given in parentheses. The last column reports the maximized log-likelihood values.

$$H(x_{t_i}) = 1 - \left[1 + \xi\left(\frac{M_{t_i} - \mu}{\sigma_{t_i}}\right)\right]^{-1/\xi} \qquad \text{where} \qquad x_{t_i} = \left(\frac{M_{t_i} - \mu}{\sigma_{t_i}}\right)$$

Constant Mean: $\mu_{t_i} = \mu$; GARCH Volatility: $\sigma_{t_i} = \sigma + \lambda_1 |\varepsilon_{t_{i-1}}| + \lambda_2 \sigma_{t_{i-1}}$; Level Volatility: $\sigma_{t_i} = \sigma M^{\gamma}_{t_{i-1}}$; Constant Volatility: $\sigma_{t_i} = \sigma$, $\lambda_1 = \lambda_2 = \gamma = 0$.

Table 4
Trend-in-Mean

	μ	σ	ξ	α	λ_1	λ_2	γ	log$L$
Maxima GARCH Volatility	0.0125 (37.661)	0.00063 (2.3198)	0.2367 (2.7334)	0.00001 (1.8697)	0.0899 (2.2035)	0.7299 (7.4055)	–	1255.43
Minima GARCH Volatility	– 0.0098 (– 40.623)	0.00032 (2.1367)	0.2387 (3.1223)	– 0.00001 (– 7.1376)	0.0627 (1.9661)	0.8331 (12.614)	–	1286.80
Maxima Level Volatility	0.0125 (28.134)	0.0775 (1.2512)	0.2278 (2.6356)	0.00001 (1.6908)	–	–	0.7121 (3.7149)	1253.77
Minima Level Volatility	– 0.0099 (– 25.730)	0.0304 (1.2206)	0.2963 (3.2692)	– 0.00001 (– 4.2491)	–	–	0.5149 (2.7081)	1283.31
Maxima Constant Volatility	0.0125 (27.688)	0.00373 (8.9413)	0.3074 (3.1592)	0.000003 (1.1803)	0.0	0.0	0.0	1246.93
Minima Constant Volatility	– 0.0112 (– 26.157)	0.00371 (9.4327)	0.1971 (3.0543)	– 0.00001 (– 3.6778)	0.0	0.0	0.0	1280.02

This table displays the maximum likelihood estimates of the location (μ), trend (α), scale (σ), sensitivity (γ), GARCH (λ_1, λ_2), and shape (ξ) parameters of the generalized Pareto distribution. Asymptotic t-statistics are given in parentheses. The last column reports the maximized log-likelihood values.

$$H(x_{t_i}) = 1 - \left[1 + \xi\left(\frac{M_{t_i} - \mu}{\sigma_{t_i}}\right)\right]^{-1/\xi} \quad \text{where} \quad x_{t_i} = \left(\frac{M_{t_i} - \mu}{\sigma_{t_i}}\right)$$

Trend-in-Mean: $\mu_{t_i} = \mu + \alpha t_i$; GARCH Volatility: $\sigma_{t_i} = \sigma + \lambda_1 | \varepsilon_{t_{i-1}} | + \lambda_2 \sigma_{t_{i-1}}$; Level Volatility: $\sigma_{t_i} = \sigma M^{\gamma}_{t_{i-1}}$; Constant Volatility: $\sigma_{t_i} = \sigma$, $\lambda_1 = \lambda_2 = \gamma = 0$.

maxima than for the local minima. In the Level–Volatility specification, the sensitivity parameter for the local maxima ($\gamma_{max} = 0.7121$) is estimated to be higher than that for the local minima ($\gamma_{min} = 0.5149$), indicating that the volatility of minimal changes is more sensitive to the level of extremes than the volatility of maximal changes. Similar to our earlier findings, the last period's volatility increases the current volatility of extremes more than the last period's unexpected news in the interest rate market. The shock persistence in volatility is again greater in the minimal interest rate changes than in the maximal changes. A notable point in Table 4 is that the trend parameter, α, for the local minima is estimated to be negative and significant at the 1% level with or without the level and GARCH effects in volatility. However, the coefficients on trend for the local maxima in the Constant, Level, and GARCH–Volatility models turn out to be insignificant. These results point to the presence of trend (or time-varying mean) in the local minima and also suggest the mean of the local maxima being independent of time.

Table 5 shows the maximum likelihood estimates of the location, AR(1), scale, sensitivity, GARCH, and shape parameters along with their asymptotic t-statistics and maximized log-likelihood functions. The previous empirical findings are not affected by the incorporation of an autoregressive process in the mean. The estimated sensitivity and shape parameters for the local maxima are again greater than those for the local minima. When the last period's extremes are included in the mean, the volatility shock persistence in the

Table 5
AR(1)-in-Mean

	μ	σ	ξ	β	λ_1	λ_2	γ	log*L*
Maxima	0.0109	0.00080	0.1727	0.1648	0.1096	0.6825	–	1256.58
GARCH	(17.750)	(2.0122)	(2.2225)	(4.1880)	(2.1134)	(4.9184)		
Volatility								
Minima	– 0.0099	0.00049	0.1792	0.1616	0.0575	0.8009	–	1287.24
GARCH	(– 21.074)	(2.4014)	(2.7496)	(4.8397)	(2.0120)	(9.7824)		
Volatility								
Maxima	0.0108	0.0725	0.1963	0.1667	–	–	0.6908	1255.23
Level	(15.973)	(1.3401)	(2.4843)	(3.6021)			(3.3103)	
Volatility								
Minima	– 0.0099	0.0231	0.1983	0.1618	–	–	0.4297	1285.65
Level	(– 18.423)	(1.2041)	(2.9257)	(4.2593)			(2.3623)	
Volatility								
Maxima	0.0116	0.00391	0.2630	0.0946	0.0	0.0	0.0	1247.95
Constant	(21.332)	(9.0112)	(2.9166)	(3.1536)				
Volatility								
Minima	– 0.0099	0.00366	0.2035	0.1555	0.0	0.0	0.0	1281.60
Constant	(– 19.672)	(9.4847)	(3.1011)	(4.9257)				
Volatility								

This table shows the maximum likelihood estimates of the location (μ), AR(1) (β), scale (σ), sensitivity (γ), GARCH (λ_1, λ_2), and shape (ξ) parameters of the generalized Pareto distribution. Asymptotic *t*-statistics are given in parentheses. The last column reports the maximized log-likelihood values.

$$H(x_{t_i}) = 1 - \left[1 + \xi\left(\frac{M_{t_i} - \mu}{\sigma_{t_i}}\right)\right]^{-1/\xi} \qquad \text{where} \qquad x_{t_i} = \left(\frac{M_{t_i} - \mu}{\sigma_{t_i}}\right)$$

AR(1)-in-Mean: $\mu_{t_i} = \mu + \beta M_{t_{i-1}}$; GARCH Volatility: $\sigma_{t_i} = \sigma + \lambda_1 |\varepsilon_{t_{i-1}}| + \lambda_2 \sigma_{t_{i-1}}$; Level Volatility: $\sigma_{t_i} = \sigma M^{\gamma}_{t_{i-1}}$; Constant Volatility: $\sigma_{t_i} = \sigma$, $\lambda_1 = \lambda_2 = \gamma = 0$.

standard deviation of extremes slightly reduces, and the persistence is again greater for the local minima than for the local maxima. For each volatility specification, the AR(1) coefficient, β, is statistically significant at the 1% level for both the maximal and minimal changes.

We formally determine the significance of trend, the past history of extremes, and the level and GARCH effects in explaining the stochastic behavior of extreme interest rate changes. We test the hypotheses that $\alpha = 0$ for the significance of trend, and $\beta = 0$ for the significance of autoregressive process in the mean. In addition, we determine whether the level and GARCH effects should be incorporated in the volatility of local extrema by testing the hypotheses $\gamma = 0$ and $\lambda_1 = \lambda_2 = 0$. These hypotheses are tested on the basis of the likelihood ratio test (LR).[14] As presented in Table 6, the LR statistics for testing the hypotheses $\gamma = 0$ and $\lambda_1 = \lambda_2 = 0$ turn out to be well above the critical value for both the maxima and minima, indicating the presence of level and GARCH effects in the volatility of extremes. The results also point to the presence of trend and autoregressive process in the mean of local minima, since both hypotheses $\alpha = 0$ and $\beta = 0$ are strongly rejected for the minimal values. The presence of AR(1) process in the mean of local maxima is also

[14] The LR statistic is calculated as LR = – 2(logL^* – log$L$), where logL^* is the value of the log likelihood under the null hypothesis, and log$L$ is the log likelihood under the alternative.

Table 6
Likelihood ratio test

LR presence of Trend-in-Mean	LR presence of AR(1)-in-Mean	LR presence of GARCH effect in Volatility	LR presence of level effect in Volatility
Maxima			
3.42 GARCH Volatility	5.72 GARCH Volatility	17.26 AR(1)-in-Mean	14.56 AR(1)-in-Mean
2.94 Level Volatility	5.86 Level Volatility	17.00 Trend-in-Mean	13.68 Trend-in-Mean
2.58 Constant Volatility	4.62 Constant Volatility	16.16 Constant Mean	13.32 Constant Mean
Minima			
9.74 GARCH Volatility	10.62 GARCH Volatility	11.28 AR(1)-in-Mean	8.10 AR(1)-in-Mean
4.26 Level Volatility	8.94 Level Volatility	13.56 Trend-in-Mean	6.58 Trend-in-Mean
4.74 Constant Volatility	7.90 Constant Volatility	8.56 Constant Mean	7.06 Constant Mean

This table displays the likelihood ratio test statistics for testing the presence of trend and AR(1) process in mean, and the presence of level and GARCH effects in Volatility of extreme changes in the federal funds rate. The hypotheses that $\alpha=0$ for testing the significance of trend, $\beta=0$ for testing the significance of autoregressive process in mean, and that $\gamma=0$ and $\lambda_1=\lambda_2=0$ for testing the presence of level and GARCH effects in Volatility are tested on the basis of the likelihood ratio test (LR). LR statistic is calculated as $\mathrm{LR}=-2(\log L^*-\log L)$, where $\log L^*$ is the value of the log likelihood under the null hypothesis, and $\log L$ is the log likelihood under the alternative. The critical values with 1° and 2° of freedom at the 5% and 1% level of significance are $\chi^2_{(1,0.05)}=3.84$, $\chi^2_{(1,0.01)}=6.63$, $\chi^2_{(2,0.05)}=5.99$, and $\chi^2_{(2,0.01)}=9.21$.

detected. The maximal values are found to be independent of time since the coefficient on trend turns out to be insignificant and the hypothesis $\alpha=0$ cannot be rejected for the local maxima based on the likelihood ratio test.

6. Risk management performance of the extreme value approach

A value at risk (VaR) model measures market risk by determining how much the value of a portfolio could decline over a given probability as a result of changes in market prices or rates.[15] The most commonly used VaR models assume that the probability distribution of daily changes in market variables is normal, an assumption that is far from perfect. The daily changes in speculative prices exhibit significant amounts of excess kurtosis. This means that the probability distributions have "fat tails" so that the extreme outcomes happen much more frequently than would be predicted by the normal distribution. Under these conditions, one may think that an alternative approach that approximates the tail areas asymptotically may be more appropriate than the standard approach of using the normal distribution. In this section, we compare the risk management performance of the normal and generalized Pareto distributions. The extreme value approach to estimating VaR allows the user to consider the distribution of extreme returns instead of the distribution of all returns. This new approach originally considered by Neftci (2000), Longin (2000), and McNeil and Frey (2000) provides a reliable risk measurement

[15] For example, if the given period of time is 1 day and the given probability is 1%, the VaR measure would be an estimate of the decline in the portfolio value that could occur with a 1% probability over the next trading day. In other words, if the VaR measure is accurate, losses greater than the VaR measure should occur less than 1% of the time.

technique that models the extreme tails of the distribution, which correspond to unusual market conditions during extraordinary periods.[16]

VaR calculations are performed in an environment where the stochastic process R_t represents a vector of risk factors such as interest rates, exchange rates, equity returns, and commodity prices. We let the arbitrage-free price of a financial asset A_t be a known function of R_t, time t, and of the parameter φ:

$$A_t = A(R_t, t; \varphi). \tag{29}$$

The stochastic variation in A during an infinitesimal interval $\mathrm{d}t$ is then given by Ito's Lemma:

$$\mathrm{d}A_t = A_{\mathrm{R}}\mathrm{d}R_t + A_t\mathrm{d}t + \frac{1}{2}A_{\mathrm{RR}}\sigma_t^2\mathrm{d}t \tag{30}$$

where $A_{\mathrm{R}} = \frac{\partial A(R_t,t)}{\partial R_t}$ and $A_{\mathrm{RR}} = \frac{\partial^2 A(R_t,t)}{\partial R_t^2}$ are the delta and gamma of the asset, respectively.

In order to calculate VaR, one imposes a model on the stochastic differential $\mathrm{d}R_t$. For most risk factors, researchers choose the stochastic differential equation of the form:

$$\mathrm{d}R_t = \mu_t\mathrm{d}t + \sigma_t\mathrm{d}W_t. \tag{31}$$

Assuming that Δt denotes the length of time interval, the discrete time approximation of the stochastic process in Eq. (31) can be written as:

$$\Delta R_t = \mu_t\Delta t + \sigma_t\Delta W_t. \tag{32}$$

The critical step in calculating VaR is the estimation of the threshold point that will define what variation in returns R_t is to be considered "extreme." We let this threshold be denoted by $\Im$, with Φ, the probability that ΔR_t exceeding the threshold $\Im$ will occur. The threshold $\Im$ is defined by:

$$P(\Delta R_t \geq \Im\sqrt{\Delta t}) = \Phi \tag{33}$$

where $P(.)$ is the underlying probability distribution, assumed to be known. In commonly used VaR models, $\Im$ is defined as:

$$\Im_{\mathrm{Normal}} = 2.326\sigma_t \tag{34}$$

where $P(.)$ is assumed to be the normal distribution with Φ equal 1%. The VaR at time t is then obtained from Eq. (30) by letting $A_t = A_{\mathrm{RR}} = 0$:

$$\mathrm{VaR}(A, \Phi, \Delta t) = 2.326\sigma_t A_{\mathrm{R}}\sqrt{\Delta t}. \tag{35}$$

The risk manager who has exposure to a risk factor R_t, which changes by discrete increments of ΔR_t, needs to know how much capital to put aside to cover at least the

[16] Fully parametric approaches with the assumption of joint normality of risk factors are widely used by practitioners. These parametric approaches misspecify the tails of the empirical distribution and underestimate the actual VaR thresholds. Fully nonparametric approaches have been proposed by Harrel and Davis (1982), Falk (1985), and Gourieroux et al. (2000). Recently, semi-parametric approaches have been proposed by Embrechts et al. (1998) based on the extreme value distributions and by Gourieroux and Jasiak (1999) based on the local likelihood methods.

fraction, $1-\Phi$, of daily losses during a year. In order to do this, the risk manager must first determine a threshold $\Im$ so that the event $(\Delta R_t \geq \Im)$ has a probability Φ under $P(.)$. The standard approach does this by using an explicit distribution that is, in general, the normal distribution. The alternative provided in the literature is to work with the extreme value distribution $H(\Im)$ (e.g., GPD) instead of $P(.)$, and then determine the threshold level $\Im$ by going backward from Φ to $\Im$ by solving:

$$H(\Im) = 1 - \Phi \tag{36}$$

given the value of Φ. As shown in Neftci (2000), the generalized Pareto distribution yields the following VaR threshold:

$$\Im_{\mathrm{GPD}} = \mu + \left(\frac{\sigma}{\xi}\right)\left[\left(\frac{\Phi n}{k}\right)^{-\xi} - 1\right] \tag{37}$$

where k and n are the number of extremes and the number of total data points, respectively. Once the location (μ), scale (σ), and shape (ξ) parameters of the GPD are estimated, one can find the VaR threshold, $\Im_{\mathrm{GPD}}$, based on the choice of confidence level (Φ).

As discussed earlier, there is substantial empirical evidence that the distribution of interest rate changes, and other assets is typically leptokurtic, that is, the unconditional return distribution shows high peaks and fat tails. This implies that extreme events are much more likely to occur in practice than would be predicted by the thin-tailed normal distribution. This also suggests that the normality assumption can produce VaR numbers that are inappropriate measures of the true risk faced by financial institutions. In order to overcome the drawbacks of the normal distribution, we use the fat-tailed Student's t distribution with a GARCH(1,1) process that takes into account time-varying volatility characterized by persistence, considers the conditional non-normality of returns, and deals with events that are relatively infrequent.

In most empirical studies, the normal density is used even though the standardized residuals obtained from ARCH-type models, which assume normality, remain leptokurtic. In light of the empirical evidence of fat-tailed errors, we use the heavy-tailed standardized Student's t distribution. We let the conditional distribution of R_t be standardized t with mean $\mu_{t|t-1} = \omega_0 + \omega_1 R_{t-1}$, variance $\sigma^2_{t|t-1}$ and degrees of freedom $v>2$, i.e.,

$$R_t = \mu_{t|t-1} + \varepsilon_t, \qquad \varepsilon_t \mid \Im_{t-1} \sim f_v(\varepsilon_t \mid \Im_{t-1}) \tag{38}$$

$$E(\varepsilon_t^2 \mid \Im_{t-1}) = \sigma^2_{t|t-1} = \beta_0 + \beta_1 \varepsilon^2_{t-1} + \beta_2 \sigma^2_{t-1} \tag{39}$$

$$f_v(\varepsilon_t \mid \Im_{t-1}) = \Gamma\left(\frac{v+1}{2}\right)\Gamma\left(\frac{v}{2}\right)^{-1}\left[(v-2)\sigma^2_{t|t-1}\right]^{-1/2}\left[1 + \frac{\varepsilon_t^2}{(v-2)\sigma^2_{t|t-1}}\right]^{-(v+1)/2} \tag{40}$$

where $\Im_{t-1}$ denotes the sigma-field generated by all the information up to time $t-1$, and $f_v(\varepsilon_t|\Im_{t-1})$ is the conditional density function for the error term ε_t. It is well known that for $1/v \rightarrow 0$ the t-distribution approaches a normal distribution with variance $\sigma^2_{t|t-1}$, but for $1/v>0$ the t-distribution has fatter tails than the corresponding normal distribution.[17]

Note that we estimate the parameters of the generalized Pareto distribution using the 2.5%, 5%, and 10% tail observations, and then calculate the 0.5%, 1%, 1.5%, 2%, 2.5%, and 5% VaRs. If the new approach is superior to the current alternatives, the standard VaR model should actually leave more than x% of the observations beyond the threshold, while the extreme value method should capture x% of the observations more closely where $x\% = 0.5\%, \ldots, 5\%$. Table 7 compares the relative performance of the normal and generalized Pareto distributions to calculating value at risk. The results show the normal distribution to be quite inadequate for determining a 0.5%, 1%, 1.5%, 2%, 2.5%, and 5% VaR thresholds for the federal funds rate.

Given that there are 11,697 daily changes in the federal funds rate from July 1, 1954 to December 29, 2000, we would expect approximately 59, 117, 175, 234, 292, and 585 observations to fall into each 0.5%, 1%, 1.5%, 2%, 2.5%, and 5% tail. However, on average, the VaR measures obtained from the normal distribution include 180 observations for the 0.5% VaRs, 210 observations for the 1% VaRs, 222 observations for the 1.5% VaRs, 284 observations for the 2% VaRs, 298 observations for the 2.5% VaRs, and 370 observations for the 5% VaRs. The results clearly indicate that the actual VaR thresholds are generally underestimated by the normal distribution for the extreme VaR thresholds and overestimated for the 5% VaR. These VaR numbers imply that the extreme tails of the actual distribution cannot be approximated adequately by the normal distribution.

As shown in Table 7, the 5% tails of the generalized Pareto distribution outperform the 2.5% and 10% tails estimates of the value at risk and provide the best predictions of catastrophic market risks during extraordinary periods. The 5% tails of the GPD capture approximately 62 observations for the 0.5% VaRs, 118 observations for the 1% VaRs, 165 observations for the 1.5% VaRs, 237 observations for the 2% VaRs, 299 observations for the 2.5% VaRs, and 583 observations for the 5% VaRs. The results imply the mean absolute percentage error (MA%E) of 4.17% for the 2.5% tails of the GPD, 3.96% for the 5% tails of the GPD, and 4.56% for the 10% tails of the GPD. Although the 2.5% and 10% tails of the generalized Pareto distribution turn out to be slightly inferior to the 5% tails of the GPD, they perform much better than the normal distribution in capturing the asymptotic behavior of interest rates. Overall, the extreme value theory yields a more accurate and robust approach to calculating VaR because the normal distribution implies MA%E of 63.29%.

We extend the *unconditional* extreme value approach proposed in the current literature by modeling the dynamic behavior of interest rate volatility in extreme values. The existing EVT-based methods utilize a sound statistical theory and offer a parametric form

[17] We do not present the estimation results from the Student's t GARCH model in order to preserve space. They are available upon request.

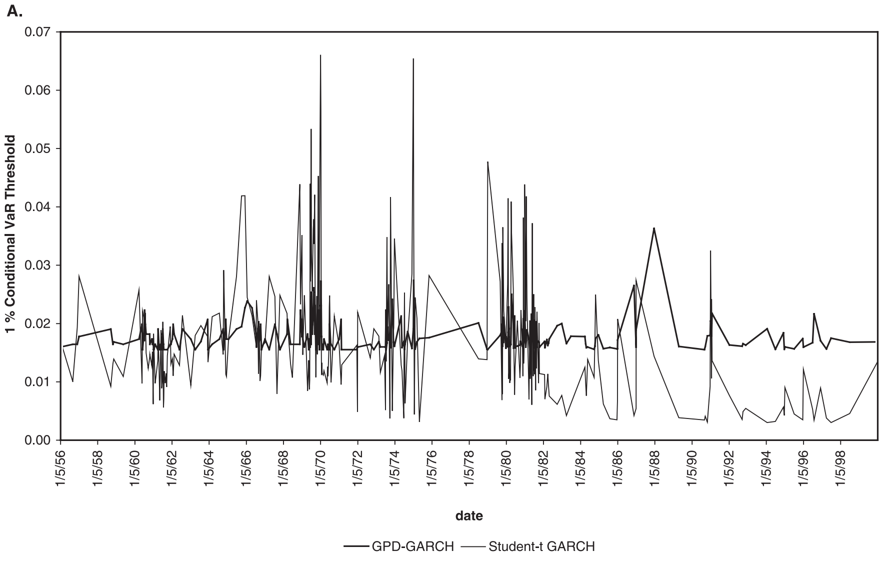
Table 7
Risk management performance of the GPD and normal distributions

Φ (%)	Actual	GPD 2.5% tails	GPD 5% tails	GPD 10% tails	Normal distribution
Maxima					
0.5	59	60	60	62	199
1	117	111	110	111	237
1.5	175	165	165	166	254
2	234	235	232	232	315
2.5	292	315	310	310	331
5	585	616	587	565	349
Minima					
0.5	59	63	64	64	160
1	117	124	125	125	182
1.5	175	173	165	165	190
2	234	245	241	241	253
2.5	292	296	287	287	264
5	585	608	579	566	391
Average MA%E	–	4.17	3.96	4.56	63.29

This table compares the relative performance of the generalized Pareto and normal distributions to calculating value at risk. The number of observations that fall in the 0.5%, 1%, 1.5%, 2%, 2.5%, and 5% tails of the distribution are presented below. Given that there are 11,697 daily changes in the federal funds rate from July 1, 1954 to December 29, 2000 we would expect approximately 59, 117, 175, 234, 292, and 585 observations to fall into each 0.5%, 1%, 1.5%, 2%, 2.5%, and 5% tail. The VaR estimates based on the extreme returns are obtained from the 2.5%, 5%, and 10% of the right and left tails of the empirical distribution. The VaR measures obtained from the normal distribution are displayed in the last column. The results imply the mean absolute percentage error (MA%E) of 4.17% for the 2.5% tail estimates, 3.96% for the 5% tail estimates, and 4.56% for the 10% tail estimates. The extreme value theory yields a more accurate and robust approach to estimating value at risk because the normal distribution implies MA%E of 63.29%.

for the tail of a distribution, but do not yield VaR measures, which reflect the current volatility background. Given the conditional heteroscedasticity of most financial data, we believe using an unconditional volatility to be a major drawback of any kind of VaR-estimator. In order to improve the existing VaR models, we propose a *conditional* extreme value approach by specifying the location and scale parameters of the generalized Pareto distribution as a function of past information as in Eqs. (21) and (23). We then compare the time-varying conditional VaR thresholds obtained from the GPD–GARCH and Student's *t* GARCH models.

The entire sample for the federal funds rate is used to estimate time-varying volatility of interest rate changes for the Student's *t* GARCH model. The *conditional* standard deviation of *extremes* is estimated using the maximal and minimal changes that correspond to 2.5% of the right and left tails of the distribution. To match the dates of the extreme and GARCH volatilities (i.e., GPD–GARCH and Student's *t* GARCH volatilities), once the conditional variances are obtained from the Student's *t* GARCH model, the relevant GARCH volatilities are extracted from the entire column. Fig. 1A and B plot the 1% time-varying conditional VaR thresholds of the GPD–GARCH and Student's *t* GARCH models for the maximal and minimal changes. The



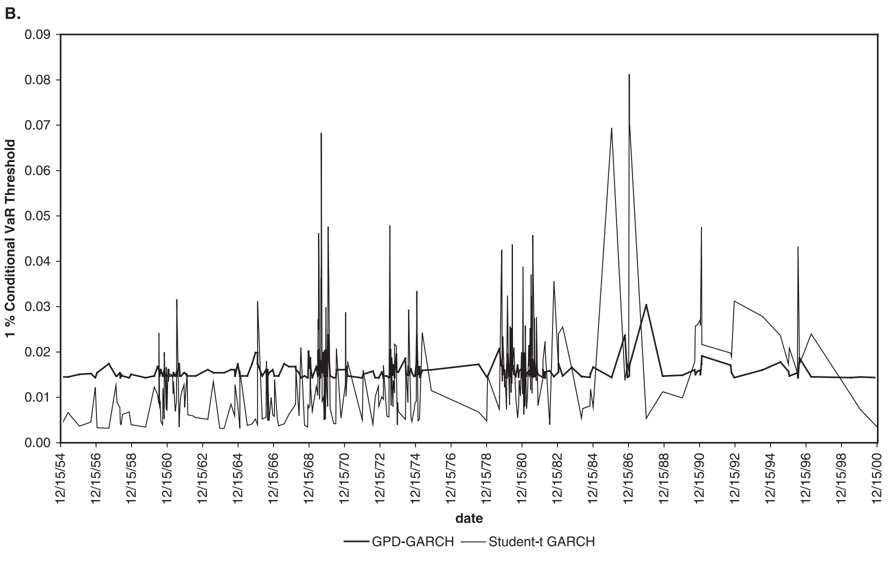


Fig. 1. (A) One percent conditional VaR threshold for the maximal changes in federal funds rate. (B) One percent conditional VaR threshold for the minimal changes in federal funds rate.

figures clearly indicate that the conditional thresholds are generally underestimated by the Student's *t* GARCH model during relatively tranquil periods, whereas the VaR thresholds are overestimated by the Student's *t* GARCH model during extremely volatile periods.

7. Conclusions

In this study, we first assume that the volatility of local maxima and local minima values of daily interest rate changes is constant, and then estimate the means of the extremes under the assumption that the location parameter is constant, or it depends on time, or it is explained as the last local extrema. We then test whether the extremes are time-dependent or if they follow an autoregressive process. Second, we incorporate the level effect and volatility clustering in the standard deviation of extremes and then reestimate the mean, which is first assumed to be constant, then is allowed to follow a trend, and is finally parameterized as a linear AR(1) process. The empirical evidence on the daily federal funds rate indicates the presence of level and GARCH effects in the volatility of both local maxima and local minima values. The results also point to the presence of an autoregressive process in the means of the local extrema. But the existence of trend in the extremes is not clear since the trend on the local maxima is not detected, although the mean of the local minima values is found to be time-dependent.

This study clearly indicates that the tails of the empirical distribution are much thicker than the tails of the normal distribution. Therefore, the extreme tails of the actual distribution cannot be approximated adequately by the normal distribution. The results imply that the VaRs calculated using the tails of the generalized Pareto distribution are significantly more precise than those estimated by the normal distribution. The 5% tails of the GPD perform slightly better than the 2.5% and 10% tails in calculating value at risk, and provide the best predictions of catastrophic market risks during extraordinary periods. The paper proposes a conditional extreme value approach to estimating VaR by specifying the mean and volatility parameters of the GPD as a function of past information. The VaR estimates show that the time-varying conditional thresholds are generally underestimated by the Student's *t* GARCH model during relatively tranquil periods, whereas the thresholds are overestimated by the Student's *t* GARCH model during extremely volatile periods. Overall, the VaR tails calculated by the generalized Pareto distribution are found to be more robust and yield more accurate estimates of the rate of occurrence and the size of extreme observations.

Acknowledgements

We would like to thank Christian Wolff (the associate editor) and an anonymous referee of this journal for their helpful comments and suggestions. We also benefited from the discussions with seminar participants at the 2001 Annual Meeting of the Financial Management Association. Turan Bali acknowledges financial support from the PSC-CUNY Research Foundation of the City University of New York.

References

Ait-Sahalia, Y., 1996a. Testing continuous-time models of the spot interest rate. Review of Financial Studies 9, 385–426.

Ait-Sahalia, Y., 1996b. Nonparametric pricing of interest rate derivatives. Econometrica 64, 527–560.

Andersen, T.G., Lund, J., 1997. Estimating continuous-time stochastic volatility models of the short-term interest rate. Journal of Econometrics 77, 343–377.

Bali, T.G., 2000. Testing the empirical performance of stochastic volatility models of the short term interest rate. Journal of Financial and Quantitative Analysis 35, 191–215.

Bali, T.G., 2001. An extreme value approach to estimating volatility and value at risk. Journal of Business (in press).

Brenner, R.J., Harjes, R.H., Kroner, K.F., 1996. Another look at models of the short term interest rate. Journal of Financial and Quantitative Analysis 31, 85–107.

Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B., 1992. An empirical comparison of alternative models of the short-term interest rate. Journal of Finance 47, 1209–1227.

Conley, T.G., Hansen, L.P., Luttmer, E.G.Z., Scheinkman, J.A., 1997. Short-term interest rates as subordinated diffusions. Review of Financial Studies 10, 525–577.

Danielsson, J., de Vries, C.G., 1997a. Tail index and quantile estimation with very high frequency data. Journal of Empirical Finance 4, 241–257.

Danielsson, J., de Vries, C.G., 1997b. Value at risk and extreme returns. FMG-Discussion Paper No. 273, Financial Markets Group, London School of Economics.

Embrechts, P., Kluppelberg, C., Mikosch, T., 1997. Modeling Extremal Events. Springer, Berlin.

Embrechts, P., Resnick, S., Samorodnitsky, G., 1998. Living on the edge. Risk 11, 96–100.

Falk, M., 1985. Asymptotic normality of the kernel quantile estimator. Annals of Statistics 13, 428–433.

Fisher, R.A., Tippett, L.H.C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society 24, 180–190.

Gollier, C., Koehl, P.F., Rochet, J.C., 1996. Risk-Taking Behavior with Limited Liability and Risk Aversion. Wharton Financial Institutions Center. DP 96-13.

Gourieroux, C., Jasiak, J., 1999. Truncated Local Likelihood and Nonparametric Tail Analysis. CREST. DP 99.

Gourieroux, C., Laurent, J.P., Scaillet, O., 2000. Sensitivity analysis of values at risk. Journal of Empirical Finance 7, 225–245.

Gradshteyn, I.S., Ryzhik, I.M., 1994. In: Jeffrey, A. (Ed.), Tables of Integrals, Series, and Products, 5th ed. Alan Jeffrey Academic Press, New York.

Harrel, F., Davis, C., 1982. A new distribution free quantile estimation. Biometrica 69, 635–640.

Jansen, D.W., de Vries, C.G., 1991. On the frequency of large stock returns: putting booms and busts into perspectives. Review of Economics and Statistics 73, 18–24.

Jansen, D.W., Koedijk, K.G., de Vries, C.G., 2000. Portfolio selection with downside limited risk. Journal of Empirical Finance 7, 247–269.

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.

Jenkinson, A.F., 1969. Statistics of extremes. In: Estimation of Maximum Flood, pp. 183–227. Technical Note 98. World Meteorological Organization, Geneva.

Kearns, P., Pagan, A., 1997. Estimating the density tail index for financial time series. Review of Economics and Statistics 79, 171–175.

Koedijk, K.G., Nissen, F.G.J.A., Schotman, P.C., Wolff, C.C.P., 1997. The dynamics of short-term interest rate volatility reconsidered. European Finance Review 1, 105–130.

Leadbetter, M.R., Lindgren, G., Rootzen, H., 1983. Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, New York.

Longin, F.M., 1996. The asymptotic distribution of extreme stock market returns. Journal of Business 69, 383–408.

Longin, F.M., 2000. From value at risk to stress testing: the extreme value approach. Journal of Banking and Finance 24, 1097–1130.

Longstaff, F.A., Schwartz, E.S., 1992. Interest rate volatility and the term structure: a two-factor general equilibrium model. Journal of Finance 47, 1259–1282.

McCulloch, J.H., 1997. Measuring tail thickness to estimate the stable index α: a critique. Journal of Business and Economic Statistics 15, 74–81.

McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7, 271–300.

Neftci, S.N., 1985. A note on the use of local maxima to predict turning points in related series. Journal of the American Statistical Association 80, 553–557.

Neftci, S.N., 2000. Extremes, tail estimation and value at risk calculations. Journal of Derivatives, 23–37 (Spring).

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of returns. Journal of Business 53, 61–65.

Pickands, J., 1975. Statistical inference using extreme order statistics. Annals of Statistics 3, 119–131.

Prescott, P., Walden, A.T., 1980. Maximum-likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution. Biometrica 67, 723–724.

Rothschild, M., Stiglitz, J.E., 1970. Increasing risk: I. Definition. Journal of Economic Theory 2, 225–243.

Santomero, A., Babbel, D., 1996. Risk Management by Insurers: An Analysis of the Process. Wharton Financial Institutions Center. DP 96-16.

Stanton, R., 1997. A nonparametric model of term structure dynamics and the market price of interest rate risk. Journal of Finance 52, 1973–2002.