EXTREME VALUE THEORY: VALUE AT RISK AND RETURNS DEPENDENCE AROUND THE WORLD

Viviana Fernandez[1]

Abstract

This paper presents two applications of Extreme Value Theory (EVT) to financial markets: computation of value at risk and assets returns dependence under extreme events (i.e. tail dependence). We use a sample comprised of the United States, Europe, Asia, and Latin America. Our main findings are the following. First, on average, EVT gives the most accurate estimates of value at risk. Second, tail dependence decreases when filtering out heteroscedasticity and serial correlation by multivariate GARCH models. Both findings are in agreement with previous research in this area for other financial markets.

JEL classification: C22, G10 Keywords: EVT, VaR, extremal dependence

I Introduction

Extreme value theory (EVT) has emerged as one of the most important statistical disciplines for the applied sciences over the last fifty years, and for other fields in recent years (e.g., finance). The distinguishing feature of EVT is to quantify the stochastic behavior of a process at unusually large or small levels. Specifically, EVT usually requires estimation of the probability of events that are more extreme than any other that has been previously observed.

This article tackles two key issues in risk management: computation of value at risk and stock market dependence using the new approach of EVT. In particular, value at risk (VaR) is a popular measure of market risk (see, for example, Jorion, 2001), whose origins date back to the late 1980's at J.P. Morgan. VaR answers the question about how much we can lose with a given probability over a certain time horizon. It became a key measure of market risk since the Basle Committee stated that banks should be able to cover losses on their trading portfolios over a ten-day horizon, 99 percent of the time. Financial firms usually use VaR for internal risk control considering a one-day horizon and a 95-percent confidence level.

More formally, VaR measures the quantile of the projected distribution of gains and losses over a given time horizon. If α is the selected confidence level, VaR is the $1-\alpha$ lower-tail level. In practical applications, computation of VaR involves choosing α, the time horizon, the frequency of the data, the cumulative distribution function of the price change of a financial position over the time horizon under consideration, and the amount of the financial position.

---

[1] Department of Industrial Engineering at the University of Chile (DII). Postal: Avenida Republica 701, Santiago-Chile. E-mail: vfernand@dii.uchile.cl. Funds provided by an institutional grant of the Hewlett Foundation to the Center for Applied Economics (CEA) at the DII are greatly acknowledged.

The assumption made about the distribution function of the price change is key to VaR calculation. Some available methods are Riskmetrics, the GARCH approach, quantile estimation, and extreme value theory (see, for example, Tsay, 2001, chapter 7). Riskmetrics assumes that the continuously compounded daily return of a portfolio follows a conditional normal distribution. The GARCH approach resorts to conditional heterocedastic models. If innovations are assumed normal, quantiles to compute VaR can be easily obtained from the standard normal distribution. Alternatively, if innovations are assumed Student-t, standardized quantiles are used. Quantile estimation in turn provides a non-parametric estimate of VaR. It does not make any assumption about the distribution of the portfolio return. There are two types of quantile methods: empirical and quantile regression. Finally, extreme value theory (EVT) has a goal to quantify the probabilistic behavior of unusually large losses, and it has arisen as a new methodology to analyze the tail behavior of stock returns (see, for example, McNeil and Frey, 2000; Zivot and Wang, 2003, chapter 5).

Traditional parametric and non-parametric methods work well in areas of the empirical distribution where there are many observations, but they provide with a poor fit to the extreme tails of the distribution. This is evidently a disadvantage because management of extreme risk calls for estimation of quantiles and tail probabilities that usually are not directly observable from the data. EVT focuses on modeling the tail behavior of a loss distribution using only extreme values rather than the whole data set. In addition, EVT offers a parametric estimate of tail distribution. This feature allows for some extrapolation beyond the range of the data. In this article, we estimate assets volatility with GARCH-type models and compute tails distributions of GARCH innovations by EVT. This makes it possible to compute conditional quantiles (i.e., VaR), and compare the EVT approach to other alternatives, such as conditional normal, t, and non-parametric quantiles.

In second term, we focus on financial markets dependence. Estimating dependence of asset returns is one of the most important subjects of portfolio theory and of many other fields of finance, such as hedging, derivatives valuation and credit analysis. One of the simplest ways to measure association is by the Pearson correlation coefficient. However, this is only appropriate for detecting linear association between two random variables. And, given that it is constructed from deviations from the mean, the weight given to extreme observations is the same as that given to all the other observations. Therefore, the Pearson correlation coefficient is not an accurate measure of dependence if extreme observations present different patterns of dependence from the rest of the sample.

An alternative approach can be found in the extreme value theory, which comes from the statistics field. EVT has been applied to financial issues only in the past years, although it has been broadly utilized in other fields, such as insurance claims, telecommunications and engineering. To date, the applications of EVT to finance have been primarily univariate (e.g., momentum of financial returns and characterization of the tails of stock returns), while multivariate applications are relatively recent (e.g., computation of the value at risk of a portfolio, co-crashes of stock and bond markets).

An important issue that arises when studying cross section dependence under EVT is that there are two types of extreme value dependence: asymptotic dependence and asymptotic independence. Both forms of dependence allow dependence between relatively large values of each variable, but the largest values from each variable can take place jointly only when the variables are asymptotically dependent (see, for example, Coles, Heffernan and Twan, 1999; Poon, Rockinger, and Tawn, 2003). The literature has generally focused on the latter.

However, if the series are asymptotically independent, such an approach will overestimate extremal dependence and, therefore, risk. The degree of overestimation will depend upon the degree of asymptotic independence. Recent research by Poon, Rockinger, and Tawn (2003) controls for asset returns heterocedasticity before testing for tail dependence. Their estimation results show that tail dependence decreases when filtering out heteroscedasticity by univariate and bivariate GARCH models. In addition, Poon et al. find that extreme value dependence is usually stronger in bear markets (left tails) than in bull markets (right tails).

First, we analyze different ways to compute value at risk for stock markets across the United States, Latin America, Europe, and Asia. We conclude that quantile estimates based on Extreme Value Theory are the best predictors. Secondly, we test the degree of extremal dependence across different financial markets. In particular, we present an application for the United States and conclude that bond markets do not exhibit extremal dependence of stock markets, and much of the extremal dependence across stock markets disappear when controlling for both serial correlation and heteroscedasticity. Some of these issues have been tackled in previous research, but our analysis presents some further insights. The topics covered in this article are relevant to portfolio management of commercial banks, insurance and re-insurance companies, and investment banks around the world, which have to asses their portfolios risk periodically.

This paper is organized as follows. Section II presents a brief overview of extreme value theory. Section III deals with computing value at risk under alternative methods for a sample of different countries around the world. Section IV focuses on the topic of extremal dependence, and presents an application for the United States. Finally, Section V summarizes our main findings.

II Theoretical Background on Extreme Value Theory

Let X_1, X_2,.., X_n be identically distributed and independent (iid) random variables representing risks or losses with unknown cumulative distribution function (cdf), $F(x)=Pr(X_i \leq x)$. Examples of random risks are negative returns on financial assets or portfolios, operational losses, catastrophic insurance claims, credit losses, natural disasters, service life of items exposed to corrosion, traffic prediction in telecommunications, etcetera (see Coles, 2001; Reiss and Thomas, 2001; McNeil and Frey 2000).

As a convention, a loss is treated as a positive number and extreme events take place when losses come from the right tale of the loss distribution F. Let $M_n=\max(X_1,X_2,\ldots, X_n)$ be the worst-case loss in a sample of n losses. For a sample of iid observations, the cdf of M_n is given by

$$\Pr(M_n \le x) = \Pr(X_1 \le x, X_2 \le x, ..., X_n \le x) = \prod_{i=1}^{n} F(x) = F^n(x) \qquad (1)$$

An asymptotic approximation to F^n (x) is based on the Fisher-Tippet (1928) theorem. Given that for x<x+, where x+ is the upper end-point of F,[2] $F^n(x) \to 0$ as $n \to \infty$, the asymptotic approximation of F^n is based on the standardized maximum value

$$Z_n = \frac{M_n - \mu_n}{\sigma_n}, \qquad \sigma_n > 0 \qquad (2)$$

where σ_n and μ_n are a scale and location parameters, respectively. The Fisher-Tippet theorem states if Z_n converges to some non-degenerate distribution function, this must be a generalized extreme value (GEV):

$$G_\zeta(z) = \begin{cases} \exp(-(1+\zeta z)^{-1/\zeta} & \zeta \ne 0, 1+\zeta z > 0 \\ \exp(-\exp(-z)) & \zeta = 0, -\infty < z < \infty \end{cases} \qquad (3)$$

The parameter ζ is a shape parameter and determines the tail behavior of $G_\zeta(z)$. If Z_n converges to $G_\zeta(z)$, then Z_n is said to be in the domain of attraction of $G_\zeta(z)$. If the tail of F declines exponentially, then $G_\zeta(z)$ is of the Gumbel type and $\zeta=0$. In this case, distributions in the domain of attraction of $G_\zeta(z)$ are of the thin-tailed type (e.g., normal, log-normal, exponential, and gamma). If the tail of F declines by a power function, then $G_\zeta(z)$ is of the *Fréchet* type and $\zeta>0$. Distributions in the domain of attraction of $G_\zeta(z)$ are called fat-tailed distributions (e.g., Pareto, Cauchy, Student-t, and mixtures models). Finally, if the tail of F is finite then $G_\zeta(z)$is of the Weibull type and $\zeta<0$. Distributions in the domain of attraction of $G_\zeta(z)$are distributions with bounded support (e.g., uniform and beta).

In practice, modeling all block maxima is wasteful if other data on extreme values are available. Therefore, a more efficient approach is to model the behavior of extreme values above a high threshold. This method receives the name of peaks over threshold (POT). An additional advantage of POT is that provides with Value-at-Risk (VaR) estimates that are easy to compute.

Let us define the excess distribution above the threshold u as the conditional probability

$$F_u(y) = \Pr(X - u \le y \mid X > u) = \frac{F(y+u) - F(u)}{1 - F(u)}, \quad y > 0 \qquad (4)$$

For those distributions F that satisfy the Fisher-Tippet theorem, it can be shown that for large enough u there exists a positive function $\beta(u)$, such that (4) is well approximated by the generalized Pareto distribution (GPD)

[2] That is, the smallest value of x such that F(x)=1

$$H_{\zeta,\beta(u)}(y) = \begin{cases} 1-\left(1+\dfrac{\zeta y}{\beta(u)}\right)^{-1/\zeta} & \zeta \neq 0 \\ 1-\exp\left(-\dfrac{y}{\beta(u)}\right) & \zeta = 0 \end{cases} \quad (5)$$

where $\beta(u)>0$, and $y\geq 0$ when $\zeta\geq 0$, and $0\leq y\leq -\beta(u)/\zeta$ when $\zeta<0$ (see, for example, Coles, 2001).

For a given value of u, the parameters ζ, μ, and σ of the GEV distribution determine the parameters ζ and $\beta(u)$. In particular, the shape parameter ζ is independent of u, and it is the same for both the GEV and GDP distributions. If $\zeta>0$, F is in the Fréchet family and $H_{\zeta,\beta(u)}$ is a Pareto distribution; if $\zeta=0$, F is in the Gumbell family and $H_{\zeta,\beta(u)}$ is an exponential distribution; and, if $\zeta<0$, F is in the Weibull family and $H_{\zeta,\beta(u)}$ is a Pareto type II distribution. In most applications of risk management, the data comes from a heavy-tailed distribution, so that $\zeta>0$.

In order to estimate the tails of the loss distribution, we resort to a theorem which establishes that, for a sufficiently high threshold u, $F_u(y)\approx H_{\zeta,\beta(u)}(y)$ (see Embrechts, Klüpperberg and Mikosch, 1997, chapter 3). By setting $x=u+y$, an approximation of F(x), for $x>u$, can be obtained from equation (4)

$$F(x) = (1-F(u))H_{\zeta,\beta(u)}(y) + F(u) \quad (6)$$

The function F(u) can be estimated non-parametrically using the empirical cdf

$$\hat{F}(u) = \frac{n-k}{n} \quad (7)$$

where k represents the number of exceedences over the threshold u. After substituting (5) and (7) into (6), we get the following estimate for F(x)

$$\hat{F}(x) = 1-\frac{k}{n}\left(1+\hat{\zeta}\,\frac{(x-u)}{\hat{\beta}}\right)^{-\frac{1}{\hat{\zeta}}} \quad (8)$$

where $\hat{\zeta}$ and $\hat{\beta}$ are estimates of ζ and β, respectively, which can be obtained by the method of maximum likelihood.

III Application of EVT to Value at Risk

Value at Risk is usually computed for confidence levels between 95 and 99.5 percent. That is, for $0.95 \le q < 1$, VaR_q is the qth quantile of the distribution F

$$VaR_q = F^{-1}(q) \tag{9}$$

where F^{-1} is the inverse function of F. For $q > F(u)$, an estimate of (9) can be obtained from (8) by solving for x

$$\widehat{VaR}_q = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{1-q}{k/n}\right)^{-\hat{\xi}} - 1\right) \tag{10}$$

In our estimation process, we follow McNeil and Frey (2000)'s two-step estimation procedure called conditional EVT

Step 1: Fit a GARCH-type model to the return data by quasi-maximum likelihood. That is, maximize the log-likelihood function of the sample assuming normal innovations.

Step 2: Consider the standardized residuals computed in Step 1 to be realizations of a white noise process, and estimate the tails of the innovations using extreme value theory. Next, compute the quantiles of the innovations for $q \ge 0.95$.

We assume that the dynamics of log-negative returns can be represented by

$$r_t = \alpha_0 + \alpha_1 r_{t-1} + \sigma_t Z_t \tag{11}$$

where α_0 and α_1 are parameters, r_{t-1} is the lagged return, and Z_t are iid innovations with zero mean and unit variance, and marginal distribution $F_Z(z)$. For simplicity, we assume that the conditional variance σ_t^2 of the mean-adjusted series $\varepsilon_t = r_t - \alpha_0 - \alpha_1 r_{t-1}$ follows a GARCH(1,1) process

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 \tag{12}$$

where $\beta_0 > 0$, $\beta_1 > 0$, and $\gamma > 0$. Strictly stationarity is ensured by $\beta_1 + \gamma < 1$.

Under the assumption of normally distributed innovations, the log-likelihood function of a sample of m iid observations is given by[3]

[3] Even if Z_t is not truly normally distributed, the maximization of (13) still provides consistent and asymptotically normal estimates (see, for example, Engle and Gonzalez-Rivera, 1991). However, Huber/White robust standard errors must be computed (Huber, 1967; White, 1982).

$$L(\theta) = -\frac{m}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{m}\log(\sigma_t) - \frac{1}{2}\sum_{t=2}^{m}\frac{(r_t - \alpha_0 - \alpha_1 r_{t-1})^2}{\sigma_t} \qquad (13)$$

Standardized residuals can be computed after maximizing (13) with respect to the unknown parameters α_0, α_1, β_0, β_1, and γ

$$(z_{t-m+1}, z_{t-m+2,},\ldots, z_t) = \left(\frac{r_{t-m+1} - \hat{\alpha}_0 - \hat{\alpha}_1 r_{t-m}}{\hat{\sigma}_{t-m+1}}, \frac{r_{t-m+2} - \hat{\alpha}_0 - \hat{\alpha}_1 r_{t-m+1}}{\hat{\sigma}_{t-m+2}}, \ldots, \frac{r_t - \hat{\alpha}_0 - \hat{\alpha}_1 r_{t-1}}{\hat{\sigma}_t}\right) \quad (14)$$

The natural 1-step forecast for the conditional variance in t+1 is given by

$$\hat{\sigma}^2_{t+1} = \hat{\beta}_0 + \hat{\beta}_1\hat{\varepsilon}^2_t + \hat{\gamma}\hat{\sigma}^2_t \qquad (15)$$

where $\hat{\varepsilon}_t = r_t - \hat{\alpha}_0 - \hat{\alpha}_1 r_{t-1}$.

For a one-day horizon, an estimate of the value at risk is

$$\widehat{VaR}^{t}_{q} \equiv \hat{r}^t_q = \hat{\alpha}_0 + \hat{\alpha}_1 r_{t-1} + \hat{\sigma}_{t+1}\widehat{VaR}(Z)_q \qquad (16)$$

where $\widehat{VaR}_q(Z)$ is given by equation (10) applied to the negative standardized residuals.

3.1 Dynamic Backtesting

In order to asses the accuracy of the EVT approach and alternative procedures to compute VaR, we backtested each method on each return series by the following steps. Let r_1, r_2, …, r_m be a historical return series. The conditional quantile $\hat{r}^t_q$ is computed on t days in the set T={n, …, m–1) using an n-day window each time, where the large of n depends on the sample size m of each returns series. Where otherwise stated, we left the last four years of data for prediction, so that n=1,000 approximately.

The constant k, which defines the number of exceedences above a threshold u, was set so that the 90th percentile of the innovation distribution is estimated by historical simulation, as suggested by McNeil and Frey (2000).

On each day $t \in T$, we estimate a new GARCH(1,1) model and fit a new generalized Pareto distribution to losses, which are computed from the standardized residuals series. This procedure, as mentioned earlier, is called conditional EVT. In addition, we estimate the unconditional EVT quantile, which corresponds to expression (10) applied to the log-negative return series.

The conditional normal quantile of the standardized residuals is simply given by $z_q=\Phi^{-1}(q)$, where $\Phi(.)$ is the cdf of a standard normal. In turn the quantile of a Student-t distribution (scaled to have variance 1) is given by $z_q = \sqrt{(\upsilon-2)/\upsilon}\,F^{-1}_W(q)$, where W follows a t-distribution with υ degrees of freedom (υ>2). On each day t, we estimate a GARCH(1,1) model with Student-t innovations and estimate a new υ and new quantiles.[4] The value at risk is computed according to formula (16) for both the normal and t conditional cases.

The quantile estimate in t $\hat{r}_q^t$ is compared in each case with r_{t+1}, the log-negative return in t+1 for $q \in \{0.95, 0.99, 0.995\}$. A violation is said to take place whenever $r_{t+1} > \hat{r}_q^t$. We can test whether the number of violations is statistically significant. In particular, let us consider the following statistic based on the binomial distribution

$$\frac{\frac{Y}{T}-p}{\sqrt{\frac{p(1-p)}{T}}} \xrightarrow{d} N(0,1) \qquad (17)$$

where T=m–n and Y is the number of violations, so that Y/T is the actual proportion of violations in the set T. The proportion p is the expected number of violations under the assumption that $Y \equiv \sum_{t \in T} I_t$ ~B(T, p), where $I_t \equiv 1_{\{r_{t+1} > r_q^t\}} = 1_{\{Z_{t+1} > z_q\}}$ ~Be(p), and I_t and I_s are independent for $t, s \in T$, $t \neq s$.

Expression (17) is a one-tailed test that is asymptotically distributed as standard normal (see, for example, Larsen and Marx, 1986, chapter 5). If Y/T<p, we test the null hypothesis of estimating correctly the conditional quantile against the alternative that the method systematically underestimates it. Otherwise, we test the null against the alternative that the method systematically overestimates the conditional quantile.

3.2 Empirical Results

3.2.1 U.S. Market: Engle (2001)'s example revisited

In this section, we look at Engle (2001)'s example on computing the value at risk of a portfolio made up by 50 percent Nasdaq, 30 percent Dow Jones Industrial Average (DJIA) and 20 percent long bonds (10-year constant maturity Treasury bond). His sample period covers March 23, 1990 through March 23, 2000. His uses a GARCH(1,1) model to estimate the 1 percent value at risk of $1,000,000 invested on the above portfolio on a specific date (March 24, 2000).

[4] If Z_t is distributed as t with υ degrees of freedom, the log-likelihood function of a sample of m independent observations becomes

$$L(\theta) = m \log\left(\frac{\Gamma(\upsilon+1)/2}{\pi^{1/2}\Gamma(\upsilon/2)} (\upsilon-2)^{-1/2} \right) - \frac{1}{2}\sum_{t=2}^{m} \log(\sigma_t) - \frac{(\upsilon+1)}{2}\sum_{t=2}^{m} \log\left(1 + \frac{(r_t - \alpha_0 - \alpha_1 r_{t-1})^2}{\sigma_t(\upsilon-2)}\right)$$

(see, for instance, Hamilton, 1994, chapter 21).

In our exercise we extended Engle's sample period to January 1990-December 2002, and used alternative approaches to compute VaR. All computations hereafter were carried out with the Finmetrics module of S-Plus 6.1. Figure 1, panels (a) and (b), shows daily returns on the four above-mentioned series along with their corresponding histograms. Table 1 in turns shows some statistics for the return series. All series strongly reject the assumption of normality, and, according to the standard deviation and the interquartile range, the most volatile series were the Nasdaq and the DJIA. In addition, the Nasdaq is the one that exhibited the lowest and highest daily return for 1990-2002.

[Figure 1 and Table 1 about here]

A better characterization of the left tail of each asset and the portfolio is given in Figure 2. In all cases, the tail estimates from a generalized Pareto distribution (GPD) are fatter than those of a normal and t distributions. This translates, as discussed below, into underestimating potential losses when using the latter distributions to model returns innovations.

[Figure 2 about here]

In order to compute the value at risk and carry out the exercises discussed below, we used a GARCH(1,1) model with an AR(1) term. The parameter estimates were obtained by the method of quasi maximum likelihood. That is, the log-likelihood function of the data was constructed by assuming that innovations are conditionally distributed as Gaussian. Huber/White robust standard errors were computed accordingly. Specification tests carried out after estimation failed to detect serial autocorrelation and missing ARCH effects, showing that the chosen functional form is adequate for the data. The details are in Table 2.

[Table 2 about here]

Table 3 presents our backtesting results for each individual asset and for the portfolio. As a decision rule, we took a p-value less than 5 percent to be evidence against the null hypothesis. Panel (a) shows the population quantiles computed by assuming normal and t innovations, and by the methods of conditional and unconditional EVT. Out of the 12 cases analyzed, the null hypothesis is rejected 11 times under the normal assumption, 9 times under the t assumption, and 3 and 9 times under the conditional and unconditional EVT assumptions, respectively. In particular, the conditional EVT approach never rejects the null hypothesis for the 99-percent quantile. Panel (b) in turn shows the backtesting results for the quantiles obtained from the empirical distribution of the standardized residuals. Overall, this procedure shows a similar pattern to that obtained under the conditional EVT assumption: the null hypothesis is rejected only thrice and no violations of the null take place for the 99-percent quantile.

Finally, Panel (c) of Table 3 shows 99-percent value at risk estimates on March 24, 2000 for the $1-million portfolio earlier described. We used five-year and ten-year windows for prediction, and computed the value at risk for that particular day by the four methods in Panels (a) and (b). As we can see, when we use a longer- time horizon window, risk tends to be underestimated by all methods. In addition, under the conditional normal, conditional t and unconditional EVT assumptions, the potential loss at the 99 percent confidence level is always underestimated with respect to the EVT and empirical quantiles methods. For instance, under the unconditional EVT procedure, and using the five-year window, the 99-percent value at risk for March 24, 2000 is $26,516, which is over $14,000 lower than that yielded by the conditional EVT and empirical quantile procedures.

[Table 3 about here]

Figure 3 depicts the actual portfolio loss over 2000-2002 and the 99-percent value at risk predicted by the conditional and unconditional EVT procedures. (This is the case above where the first 10 years of data are used for estimation, and three years are left for prediction). The years 2000 and 2001 stand out as stress periods. The figure clearly illustrates the fact that the unconditional EVT estimate does not respond quickly to changing volatility and, therefore, it tends to be violated more often than the conditional EVT estimate.

[Figure 3 about here]

3.2.2 Other U.S. Stock Indices

We also studied the properties of our four VaR-estimation methods for other U.S. stock indices: the Standard and Poor (S&P) 500, the Wilshire 5000, and the Russell 3000. The Wilshire 5000 is the most comprehensive stock of all U.S. indexes, encompassing small-cap, mid-cap, and large-cap stocks. The index is comprised by over 7,000 stocks, and it is considered a better representation of total market performance than the S&P 500. The Russell 3000 Index is composed of 3,000 large U.S. Companies, as determined by market capitalization. This portfolio represents approximately 98 percent of the U.S. equity market. The sample periods for the S&P500, the Wilshire 5000 and the Russell 3000 are, respectively, January 1980-December 2002, January 1991-December 2002, and January 1988-December 2002. All returns series are daily.

Panel (a) of Table 4 shows again that all quantile estimators do very poorly except for the conditional EVT approach, which fails only once. The prediction errors are particularly high for the Wilshire 5000 index: only in two cases the null hypothesis is not rejected. As Figure 4 shows, this index exhibits a particularly volatile behavior from 1997 onwards. Panel (b) of Table 4 shows in turn that the empirical quantiles approach behaves quite successfully for the S&P 500 and the Russell 3000, in that the null hypothesis is never rejected. Again the poorest fit is obtained for the Wilshire index.

[Table 4 and Figure 4 about here]

3.2.3 Stock Markets outside the U.S.

In this section, we look at stock markets outside the U.S.: Latin America, Asia, and Europe. For each continent, we selected the most representative stock indices. Specifically, for Latin America we chose the BOVESPA (Brazil), the MERVAL (Argentina), the IPSA (Chile), and the IPC (Mexico). The BOVESPA is a total return index weighted by traded volume and is comprised by the most liquid stocks traded on the Sao Paulo Stock Exchange. Argentina's MERVAL is the market value of a stock portfolio, selected according to participation in the Buenos Aires Stock Exchange, number of transactions and trading value. The IPSA is composed of the 40 stocks with the highest average annual trading volume on the Santiago Stock Exchange, Chile. And, the Mexican Stock Exchange Index (IPC) is a capitalization-weighted index of the leading stocks traded on the Mexican Stock Exchange. The sample period of each index covers approximately from the early 1990's until December 2002.

For Asia, we picked the Nikkei-225 (Japan), the Hang Seng (Hong Kong), the Kuala Lumpur Stock Exchange Composite Index, KLSE (Malaysia), the Korean Composite Stock Price Index (Kospi)-200 (South Korea), and the Straits Time Index, STI (Singapore). The Nikkei-225 Stock Average is a price-weighted index of 225 top-rated Japanese companies listed in the First Section of the Tokyo Stock Exchange. The Hang Seng is a capitalization-weighted index of 33 companies that represent approximately 70 percent of the total market capitalization of the Stock Exchange of Hong Kong. The KLSE Composite Index is a broad-based capitalization-weighted index of 100 stocks designed to measure the performance of the Kuala Lumpur Stock Exchange. The KOSPI-200 is a capitalization-weighted index of 200 Korean stocks, which make up 93 percent of the total market value of the Korea Stock Exchange. And, the Straits Times Index (STI) is a modified market capitalization-weighted index comprised of the most heavily weighted and active stocks traded on the Stock Exchange of Singapore. The sample period for the Nikkei-225, the Hang Seng, and the STI covers from the late 1980's until December 2002, whereas the sample periods for the KLSE and the KOSPI-200 cover from the early 1990's until December 2002.

Finally, for Europe, we selected the DAX-30 (Germany), the CAC-40 (France), the FTSE-250 (U.K.), and the IBEX-35 (Spain). The German Stock Index (DAX) is a total return index of 30 selected German blue chip stocks traded on the Frankfurt Stock Exchange. The CAC-40 is a narrow-based, modified capitalization-weighted index of 40 companies listed on the Paris Stock Exchange. The FTSE (Financial Times Stock Exchange)-250 is a capitalization-weighted index of the 250 most highly capitalized companies, outside of the FTSE-100, traded on the London Stock Exchange. And, the IBEX 35 is the official index for the market segment of continuously traded stocks. The index sample is composed of the 35 most actively traded stocks among the securities quoted on the Joint Stock Exchange System of the four Spanish stock exchanges. Except for the FTSE 250, whose sample period covers form the mid-1980's to December 2002, the sample period for the other three indices goes from the early 1990's until December 2002 or the early months of 2003.

Table 5 shows descriptive statistics for the above indices. All return series are daily. Both Brazil and Argentina stand out among all other countries for their high volatility, as measured by the standard deviation and interquartile range of their indices returns. By contrast, Chile and Mexico show similar volatility to that of Asian and European countries. The index that exhibits the lowest volatility is the U.K.'s FTSE-250 index.

[Table 5 about here]

All returns series strongly reject the assumption of normality because of their high kurtosis (which ranges between 5.3 and 73.7) and skewness which departs from zero. In particular, the most skewed and leptokurtic series is Hong Kong's Hang Seng. Figure 4 sheds more light on the evolution of each series. In general, stock markets exhibit more volatility from 1998 onwards. This is particularly noticeable for the KOSPI-200 return series.

[Figure 4 about here]

Table 6, panels (a) through (f), shows value-at-risk estimation for our sample of countries. For Europe, we find a similar pattern to that of the U.S (panels (c) and (d)).: quantiles computed by the conditional standard normal, conditional Student-t, and unconditional EVT methods are far off the mark, while the conditional EVT and empirical quantile methods prove quite successful. (The former fails only once whereas the latter fails twice out of 12 cases). The Latin American and Asian return series exhibit a slightly different behavior. In particular, for Latin America the methods closest to the mark are conditional EVT, unconditional EVT, and the empirical quantiles (panels (a) and (b)). Unexpectedly, the unconditional EVT method outdoes the conditional EVT method in the 99.5-percent VaR for MERVAL. This is a rare case because we would expect the conditional EVT method to do better than the unconditional EVT one, given that the former adjusts more quickly to changing volatility than the latter. For Asia (panels (e) and (f)) in turn, the conditional normal method has the worst performance, like in the above cases, but the conditional-t method do as well as the conditional-EVT method. In particular, the conditional-t method never fails for those returns series that are failed-tail but relatively symmetric, namely, the KLSE, KOSPI-200, and STI. Finally, we again encounter a case where the unconditional EVT method outdoes the conditional EVT method (99.9-percent VaR for the Nikkei-225). As a whole, the unconditional EVT and empirical quantile methods work best for Asia, rejecting the null hypothesis only once.

[Table 6 about here]

To summarize our results in Sections 3.2.1 through 3.2.3, we conclude that, out of the 60 cases analyzed, the conditional-normal and conditional-t methods have the poorest performance, failing 34 and 22 times, respectively. They are followed by the unconditional-EVT and empirical-quantile methods, which fail 19 and 10 times, respectively. Finally, the conditional-EVT method fails only 6 times. Although not infallible, the latter proves to be the best procedure, among the ones analyzed, to compute value at risk.

IV Further Applications of Extreme Value Theory: Returns Dependence

This section benefits from work by Coles, Heffernan, and Twan (1999), and from recent extensions by Poon, Rockinger, and Tawn (2003). Poon et al. introduce a special case of threshold modeling connected with the generalized Pareto distribution for the Fréchet case. For this particular case, the tail of a random variable Z above a (high) threshold u can be approximated as

$$1 - F(z) = \Pr(Z < z) \sim z^{-1/\eta} L(z) \qquad \text{for } z>u \qquad (18)$$

where L(z) is a slowly varying function of z,[5] and $\eta>0$. If treated as a constant for all z>u, that is L(z)=c, and under the assumption of n independent observations, the maximum-likelihood estimators for η and c are

$$\hat{\eta} = \frac{1}{n_u}\sum_{j=1}^{n_u}\log\left(\frac{z_{(j)}}{u}\right) \qquad \hat{c} = \frac{n_u}{n}u^{1/\hat{\eta}}, \tag{19}$$

where $z_{(1)},\dots, z_{(n_u)}$, are the n_u observations above the threshold u. $\hat{\eta}$ is known as the Hill estimator

The asymptotic variance of $\hat{\eta}$ is given by $a\,\mathrm{var}(\hat{\eta}) = \frac{\eta^2}{n_u}$. The asymptotic variance of $\hat{c}$ can be obtained by the delta method, avar($\hat{c}$)=$\frac{n_u}{n^2}\frac{u^{2/\eta}\log^2(u)}{\eta^2}$.

The first step is to transform the original variables to a common marginal distribution. Let (X,Y) be bivariate returns with corresponding cumulative distributions functions F_X and F_Y. The bivariate returns are transformed to unit Fréchet marginals (S, T) using the transformation

$$S = -\frac{1}{\ln F_X(X)} \qquad T = -\frac{1}{\ln F_Y(Y)} \qquad S>0,\ T>0. \tag{20}$$

Under this transformation, $\Pr(S>s)=\Pr(T>s)\sim s^{-1}$.[6] As both S and T are on a common scale, the events {S>s} and {T>s}, for large values of s, correspond to equally extreme events for each one. Given that $\Pr(S>s)\to 0$ as $s\to\infty$, it becomes natural to consider the conditional probability $\Pr(T>s|S>s)$ for large s. If (S,T) are perfectly dependent, $\Pr(T>s|S>s)=1$. By contrast, if (S,T) are exactly independent, $\Pr(T>s|S>s)=\Pr(T>s)$, which tends to zero as $s\to\infty$. Let us define

$$\chi = \lim_{s\to\infty}\Pr(T>s \mid S>s) \qquad 0\le\chi\le 1\ . \tag{21}$$

Variables are called asymptotically dependent if $\chi>0$, and asymptotically independent if $\chi=0$. In other words, χ measures the degree of dependence that lingers in the limit. Nonetheless, random variables, which are asymptotically independent, may show different degrees of dependence for finite levels of s. Based on this fact, Coles, Heffernan and Twan (1999) proposed the following measure of dependence

[5] A function on L on (0, ∞) is slowly varying if $\lim_{x\to\infty} L(tz)/L(z)=1$ for t>0.

[6] $\Pr(S>s)=\Pr(-1/\ln F_X(X)>s)=\Pr(\ln F_X(X)>-1/s)=\Pr(F_x(X)>\exp(-1/s))$, given that the log function is monotonic. Furthermore, $F(X)\sim U(0,1)$. Therefore, $\mathrm{Prob}(S>s)=1-\exp(-1/s)=L(s)\ s^{-1}\sim s^{-1}$, where L(s) is a slow varying function of s.

$$\bar{\chi} = \lim_{s\to\infty} \frac{2\log(\Pr(S>s))}{\log(\Pr(S>s, T>s))} - 1 \qquad -1 < \bar{\chi} \leq 1. \tag{22}$$

This is a well-defined measure of asymptotic independence as it gives the rate at which Pr(T>s|S>s)→0. Values of $\bar{\chi}$>0, $\bar{\chi}$=0 and $\bar{\chi}$<0 are approximate measure of positive dependence, exact independence, and negative dependence.

The pair $(\chi, \bar{\chi})$ provide all the necessary information to characterize both the form and degree of extreme dependence. For asymptotically dependent variables, $\bar{\chi}$=1 and the degree of dependence is measured by χ>0. For asymptotic independent variables, χ=0 and the degree of dependence is measured by $\bar{\chi}$. Therefore, one should first test if $\bar{\chi}$=1 before reaching any conclusion about the dependence based on χ.

It can be shown that

$$\Pr(S>s, T>s) \sim L(s)s^{-1/\xi} \qquad \text{as } s\to\infty,$$

where 0<ξ≤1 and L(s) is a slowly varying function. Given that $\Pr(S>s) \sim s^{-1}$, $\bar{\chi}$ boils down to $\bar{\chi} = 2\xi - 1$.

The test of dependence is implemented by letting Z=min(S,T), and noting that

$$\begin{aligned} \Pr(Z>z) &= \Pr\{\min(S,T)>z\} \\ &= \Pr(S>z, T>z) \\ &= L(z)\, z^{-1/\xi} \\ &= d\, z^{-1/\xi} \qquad \text{for } z>u, \end{aligned} \tag{23}$$

for some high threshold u. The above equation shows that ξ is the tail index of the univariate random variable Z. Therefore, it can be obtained by the Hill estimator, constrained to the interval (0, 1], and d, the scale parameter, can be computed as explained earlier.

Under the assumption of independent observations on Z, we have

$$\hat{\bar{\chi}} = \frac{2}{n_u}\left(\sum_{j=1}^{n_u} \log\left(\frac{z_{(j)}}{u}\right)\right) - 1 \qquad \mathrm{Var}(\hat{\bar{\chi}}) = \frac{(\hat{\bar{\chi}}+1)^2}{n_u}, \tag{24}$$

where $\hat{\bar{\chi}}$ is asymptotically normal.

The decision rule is: if $\hat{\bar{\chi}}$ is significantly less than 1, that is, if $\hat{\bar{\chi}} + 1.96\sqrt{\mathrm{Var}(\hat{\bar{\chi}})} < 1$, we conclude that the variables are asymptotically independent and take χ=0. In case there is no enough evidence to reject the null hypothesis $\bar{\chi} = 1$, we estimate χ under the assumption that $\bar{\chi} = \xi = 1$. In such case, $\hat{\chi} = \frac{un_u}{n}$ and $\mathrm{Var}(\hat{\chi}) = \frac{un_u(n - n_u)}{n^3}$.

The application of this section deals with tail dependence of the DJIA, and the 10-year T-Bill and the Nasdaq returns data, which were analyzed in Section 3.2.1. To start with, we draw scatter plots of the DJIA and the T-Bill and the DJIA and the Nasdaq. The results are depicted in Figure 7. Simple inspection of the plots shows that the correlation between the DJIA and the T-Bill is very low (equal to 0.04), and that this will translate into asymptotic independence of both negative and positive extremes. By contrast, the DJIA and the Nasdaq exhibits a much higher correlation coefficient (equal to 0.68), and the dependence of their returns is persistent for both negative and positive extremes. However, as discussed below, tail dependence is quite sensitive to filtering out the data for both heterocedasticity and serial correlation.

[Figure 7 about here]

In order to control for both heteroscedasticity and serial correlation, we use a diagonal VEC model or DVEC (1, 1)

$$\mathbf{r}_t = \mathbf{c} + \boldsymbol{\beta}\mathbf{r}_{t-1} + \boldsymbol{\varepsilon}_t \qquad t=2, \ldots, T, \tag{25}$$

where $\mathbf{r}_t$ is a k x 1 vector of asset returns, $\mathbf{c}$ is a k x 1 vector of constant terms, $\mathbf{r}_{t-1}$ is a k x 1 vector containing the first lag of $\mathbf{r}_t$, $\boldsymbol{\beta}$ is a k x 1 vector, and $\boldsymbol{\varepsilon}_t$ is a k x 1 white noise vector with zero mean. The matrix variance-covariance of $\boldsymbol{\varepsilon}_t$ is given in this case by

$$\boldsymbol{\Sigma}_t = \mathbf{A}_0 + \mathbf{A}_1 \otimes (\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}') + \mathbf{B} \otimes \boldsymbol{\Sigma}_{t-1} \qquad t=2, \ldots, T, \tag{26}$$

where $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{B}$, $\boldsymbol{\Sigma}_t$ and $\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}'$ are k x k matrices, for t=2, …,T, and $\otimes$ denotes the Hadamar product (e.g., Bollerslev, Engle, and Wooldridge, 1988; Zivot and Wang, 2003, chapter 13). In order to obtain the elements of $\boldsymbol{\Sigma}_t$,, only the lower part of the system in (26) is considered. For our application, we have $\mathbf{r}_t = (r_{DJIA}, r_{T\text{-}Bill}, r_{Nasdaq})_t'$. The choice of model (25)-(26) is based on its good fit to the data.

Let us now consider independent observations of negative returns (X_t, Y_t), t=2,…,T, with unknown distribution F. It can be shown that the random variables $u_t = F_X(X_t)$ and $v_t = F_Y(Y_t)$ are both distributed as Uniform, where F_X and F_Y are the marginal distribution functions. An informal procedure to detect extremal dependence (in the left tail, in this case) consists of examining the large values of u_t and v_t (see, for example, Coles, Heffernan, and Tawn, 1999). Since F_X and F_Y are unknown, estimates are obtained from the empirical distribution functions.

Panel (a) of Figure 7 shows left-tail dependence for the unfiltered data of the Nasdaq and the DJIA returns and the T-Bill and the DJIA returns. Evidence of rather strong dependence of the Nasdaq and the DJIA in bear markets is shown, while little evidence is found for the T-Bill and the DJIA. Panel (b) in turn shows how extremal dependence decreases when filtering out the data for both heteroscedasticity and serial correlation. Specifically, the corresponding pairs of negative standardized residuals obtained from the DVEC(1,1) model are shown. The extremal dependence of the Nasdaq and the DJIA is substantially reduced, while the T-Bill and the DJIA show no pattern of extremal dependence as before.

[Figure 7 about here]

The next step consists of testing tail dependence formally by using the machinery described earlier. In order to do so, one has to choose an appropriate threshold u to compute the Hill estimator. The simplest approach is to plot it against u and find a proper u, such that the Hill estimator appears to be stable (see, for instance, Tsay, 2001, chapter 7). However, in some cases such stability is not easy to visualize. That is why formal approaches to find u have been designed.

In a recent article, Matthys and Beirlant (2000) overview several methods of adaptive threshold selection that have been developed in recent years. The authors distinguish two approaches to estimating the optimal threshold. One consists of constructing an estimator for the asymptotic mean-squared error (AMSE) of the Hill estimator, and choosing the threshold that minimizes it. This approach includes a bootstrap method (the one used by Poon et al.) and an exponential regression model. The latter is studied in detailed in Beirlant, Diercks, Goegebeur, and Matthys (1999). The second approach derives estimators directly for u, based on the representation of the AMSE of the Hill estimator.

Given that the exponential regression approach is easy to implement, it is our choice to find the optimal threshold. Feuerverger and Hall (1999) and Beirlant et al (1999) derived an exponential regression model for the log-spacings of upper statistics

$$j(\log(X_{n-j+1,n}) - \log(X_{n-j,n}) \sim \left(\gamma + b_{n,k}\left(\frac{j}{k+1}\right)^{-\rho}\right) f_j \qquad 1 \le j \le k \qquad (27)$$

where $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$, $b_{n,k} = b\left(\frac{n+1}{k+1}\right)$, $1 \le k \le n-1$, $(f_1, f_2, \ldots, f_k)$ is a vector of independent standard exponential random variables, and $\rho \le 0$ is a real constant.

If we fix the threshold u at the $(k+1)^{th}$ largest observation, the Hill estimator is given by $H_{k,n} = \frac{1}{k}\sum_{j=1}^{k} \log(X_{n-j+1,n}) - \log(X_{n-k,n})$. In turn this can be rewritten as

$$H_{k,n} = \frac{1}{k}\Big((\log(X_{n,n}) - \log(X_{n-1,n})) + 2(\log(X_{n-1,n}) - \log(X_{n-2,n})) + \ldots + k(\log(X_{n-k+1,n}) - \log(X_{n-k,n}))\Big)$$

$$= \frac{1}{k} \sum_{j=1}^{k} j(\log(X_{n-j+1,n}) - \log(X_{n-j,n})) \,. \quad (28)$$

The Hill estimator written this way is the maximum likelihood estimator of γ in the reduced model

$$j(\log(X_{n-j+1,n}) - \log(X_{n-j,n})) \sim \gamma f_j \qquad 1 \le j \le k.$$

Given that the Hill estimator is an average of independent exponential random variables, its variance can be approximated by

$$\mathrm{Var}(H_{k,n}) \sim \frac{\gamma^2}{k} \,, \quad (29)$$

while its bias arises from neglecting the second term in the right-hand side of equation (27)

$$E(H_{n,k} - \gamma) \sim \frac{b_{n,k}}{1-\rho} \,. \quad (30)$$

For $n \to \infty$, $k \to \infty$, and $k/n \to 0$, the Hill estimator is asymptotically normal

$$\sqrt{k}\left(H_{k,n} - \gamma - \frac{b_{n,k}}{1-\rho} \right) \xrightarrow{d} N(0, \gamma^2)$$

From above, the AMSE of the Hill estimator is given by

$$\mathrm{AMSE}\, H_{k,n} = \left(\frac{b_{n,k}}{1-\rho} \right)^2 + \frac{\gamma^2}{k} \quad . \quad (31)$$

Therefore, the optimal threshold k_n^{opt} is defined as the one that minimizes (31)

$$k_n^{opt} \equiv \arg\min_k (\mathrm{AMSE}\, H_{k,n}) = \arg\min_k \left(\left(\frac{b_{n,k}}{1-\rho} \right)^2 + \frac{\gamma^2}{k} \right). \quad (32)$$

The algorithm for the exponential regression goes as follows

- In model (26) fix ρ at $\rho_0 = -1$ and calculate least-squares estimates $\hat{\gamma}_k$ and $\hat{b}_{n,k}$ for each $k \in \{3, \ldots, n\}$.

- Determine $\overline{\text{AMSE } H_{k,n}} = \left(\frac{\hat{b}_{n,k}}{1-\hat{\rho}_k}\right)^2 + \frac{\hat{\gamma}_k^2}{k}$ for $k \in \{3,\ldots,n\}$, with $\hat{\rho}_k \equiv \rho_0$. [7]
- Determine $\hat{k}_n^{opt} = \arg\min_{3\leq k\leq n}(\overline{\text{AMSE}\,H_{k,n}})$ and estimate γ by $H_{\hat{k}_n^{opt}}$.

The first step of the algorithm boils down to running a linear regression of $j(\log(X_{n-j+1,n}) - \log(X_{n-j,n}))$ on a constant term and $\frac{j(n+1)}{(k+1)^2}$ for each $k \in \{3,\ldots,n\}$.

Figure 8 shows the Hill estimator for the tail index ξ in equation (23) for return pairs, evaluated at different values of the threshold u. Panel (a) depicts right- and left-tail dependence of the raw returns (T-Bill/DJIA and Nasdaq/DJIA), while Panel (b) depicts the right- and left-tail dependence of the corresponding standardized residuals. Simple inspections of the graphs do not shed much light on the optimal threshold to be selected in each case.

[Figure 8 about here]

Therefore, we resorted to Beirlant et al.'s procedure to determine k*, the optimal threshold. The results are reported in Table 7. The extremal dependence of the T-Bill and the DJIA returns in the left and right tails is low, as Panel (a) shows. (Our graphical analysis had already suggested the low dependence in the left tail for this pair). The Nasdaq and the Dow Jones, by contrast, show asymptotic independence in the left tail (bear markets) but asymptotic dependence in the right tail (bull markets). However, such finding does not longer hold after filtering the data. In fact, for the filtered data, we reached the same conclusion as Poon et al.: tail dependence tends to be stronger in bear markets than in bull markets for both pairs: T-Bill/DJIA and Nasdaq/T-Bill. In addition, for all cases, asymptotic independence cannot be rejected (Panel (b)).

[Table 7 about here]

V Conclusions

Extreme value theory (EVT) has emerged as one of the most important statistical disciplines for the applied sciences over the last fifty years, and for other fields in recent years (e.g., finance). The distinguishing feature of EVT is to quantify the stochastic behavior of a process at unusually large or small levels. Specifically, EVT usually requires estimation of the probability of events that are more extreme than any other that has been previously observed.

This article has tackled two key issues in risk management: computation of value at risk (VaR) and stock market dependence using the new approach of EVT. First, We analyzed different ways to compute value at risk for stock markets across the United States,

[7] Matthys and Beirlant point out that for many distributions the exponential regression method works better, in MSE-sense, if the nuisance parameter ρ is fixed at some value ρ_0 rather than estimated.

Latin America, Europe, and Asia. We concluded that quantile estimates based on EVT are best. Secondly, we tested the degree of extremal dependence across different financial markets in the United States. We concluded that bond markets do not exhibit extremal dependence of stock markets, and much of the extremal dependence across stock markets disappear when controlling for both serial correlation and heteroscedasticity.

References

Beirlant, J., G. Dierckx, Y. Goegebeur, and G. Matthys (2000), "Tail Index Estimation and an Exponential Regression Model". *Extremes* 2:2, 177-200.

Bollerslev, T., R.F. Engle, and J.M. Wooldridge (1998), "A Capital-Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy* 96, 116-131.

Coles S., J. Heffernan, and J. Tawn (1999), "Dependence Measures for Extreme Value Analyses," *Extremes* 2(4), 339-365.

Embrechts P., C. Klüppelberg, and T. Mikosch (1997), *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag Berlin Heidelberg.

Engle R., and G. Gonzalez-Rivera (1991), "Semiparametric ARCH models," *Journal of Business and Economic Statistics* 9(4): 345-59.

Engle, R. (2001), "GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives*, Volume 15(4), pp. 157-168.

Fisher, R. and L. Tippet (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.

Huber, P.J. (1967), "The Behavior of Maximum Likelihood Estimates under Non-Standard Assumptions," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* 1(1), pp. 221-233.

Larsen R. and M. Marx (1986), *An Introduction to Mathematical Statistics and Its Applications*. Second edition. Prentice-Hall, Englewoods Cliffs, New Jersey.

Matthys G., and J. Beirlant (2000), "Adaptive Threshold Selection in Tail Index Estimation," Working paper University Center for Statistics at the Catholic University Leuven.

McNeil A. and R. Frey (2000), "Estimation of Tail-Related Risk Measures for Heterocedastic Financial Times Series: an Extreme Value Approach," *Journal of Empirical Finance* 7, pp. 271-300.

Poon, S., M. Rockinger, and J. Tawn (2003), "Nonparametric Extreme Value Dependence Measures and Finance Applications," *Statistica Sinica*. Forthcoming July.

Tsay, R. (2001), *Analysis of Financial Time Series*. John Wiley & Sons; 1st edition.

White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models," *Econometrica* 50(1), pp. 1-26.

Zivot, E., and J. Wang (2003), *Modeling Financial Times Series with S-Plus*. Insightful Corporation.

TABLES

Table 1 Descriptive Statistics of Daily Returns for the U.S.

	T-Bill	DJIA	Nasdaq	Portfolio
Mean	0.000	0.000	0.000	0.000
Median	0.000	0.000	0.001	0.001
Std. Dev.	0.006	0.010	0.016	0.011
IQ range	0.007	0.011	0.014	0.010
Minimum	–0.036	–0.075	–0.102	–0.067
Maximum	0.022	0.062	0.133	0.070
Kurtosis	5.228	7.546	8.500	7.107
Skewness	–0.379	–0.251	–0.008	–0.067
Jarque-Bera	748.9	2,829.7	4,090.9	2,284.1
p-value	0.000	0.000	0.000	0.000
Observations	3,246	3,246	3,246	3,246
Sample period	Jan 90-Dec 02	Jan 90-Dec 02	Jan 90-Dec 02	Jan 90-Dec 02

Table 2 GARCH(1,1) model for the Portfolio Return (r_t)

Coefficient	Value	St. Error	Z-test	P-value
Constant	5.57E-04	1.40E-04	3.971	0.000
r_{t-1}	1.22E-01	0.020	5.984	0.000
Variance equation				
Constant	1.33E-06	7.11E-07	1.868	0.062
ARCH(1)	0.092	0.028	3.258	0.001
GARCH(1)	0.894	0.034	26.160	0.000

Notes: (1) Parameter estimates are obtained by assuming that the conditional distribution of the innovations is gaussian. (2) Robust standard errors are computed by Huber/White's method. (3) Specification tests for standardized residuals: i) Normality test (Jarque-Bera)= 596.2, p-value=0.000; ii) Ljung-Box test for standardized residuals (12 lags) =13.55, p-value= 0.330, iii) Ljung-Box test for squared standardized residuals (12 lags)=9.443, p-value=0.665, iv) ARCH effects test (12 lags)=9.19, p-value=0.686. 4) Data source: Bloomberg.

Table 3 Backtesting Results

(a) Population Quantiles

	95% cond. normal	95% cond. t	95% cond.. EVT	95% unc. EVT	99% cond. normal	99% cond. t	99% cond. EVT	99% unc. EVT	99.5% cond. normal	99.5% cond. t	99.5% cond. EVT	99.5% unc. EVT
Portfolio												
% error	7.31%	7.76%	6.46%	9.41%	2.05%	1.55%	1.20%	1.95%	1.50%	1.05%	0.55%	0.95%
binomial test	4.74	5.66	2.99	9.05	4.73	2.48	0.91	4.28	6.35	3.49	0.32	2.86
p-value	0.00	0.00	0.00	0.00	0.00	0.01	0.18	0.00	0.00	0.00	0.37	0.00
Nasdaq												
% error	7.66%	7.81%	6.61%	10.27%	1.80%	1.40%	0.80%	2.50%	1.25%	0.75%	0.65%	1.25%
binomial test	5.46	5.77	3.30	10.80	3.61	1.81	–0.89	6.75	4.76	1.59	0.96	4.76
p-value	0.00	0.00	0.00	0.00	0.00	0.04	0.19	0.00	0.00	0.06	0.17	0.00
Rate												
% error	4.91%	5.21%	5.11%	5.51%	2.00%	1.50%	1.20%	1.20%	1.55%	0.70%	0.65%	0.70%
binomial test	–0.19	0.43	0.22	1.04	4.50	2.26	0.91	0.91	6.67	1.27	0.96	1.27
p-value	0.42	0.34	0.41	0.15	0.00	0.01	0.18	0.18	0.00	0.10	0.17	0.10
DJIA												
% error	5.86%	6.31%	5.76%	8.06%	2.05%	1.55%	1.10%	1.85%	1.65%	0.95%	0.80%	1.10%
binomial test	1.76	2.68	1.56	6.28	4.73	2.48	0.46	3.83	7.30	2.86	1.91	3.81
p-value	0.04	0.00	0.06	0.00	0.00	0.01	0.68	0.00	0.00	0.00	0.03	0.00
Rejection null by quantile	3	3	2	3	4	4	0	3	4	2	1	3

(b) Empirical Quantiles

	95%	99%	99.5%
Portfolio			
% error	6.36%	1.05%	0.60%
binomial test	2.79	0.23	0.64
p-value	0.00	0.41	0.26
Nasdaq			
% error	6.51%	0.95%	0.60%
binomial test	3.10	–0.22	0.64
p-value	0.00	0.41	0.26
Rate			
% error	5.16%	1.15%	0.90%
binomial test	0.32	0.68	2.54
p-value	0.37	0.25	0.01
DJIA			
% error	5.66%	1.35%	0.75%
binomial test	1.35	1.58	1.59
p-value	0.09	0.06	0.06
Rejection of null by quantile	2	0	1

Table 3 Continued

(c) Value at risk for the portfolio on March 24, 2000 in US dollars

	99-percent Value at risk				
	conditional normal	conditional t	conditional EVT	unconditional EVT	empirical
5- year window data	$32,234.8	$35,626.4	$41,018.9	$26,516.0	$41,092.6
10- year window data	$27,056.8	$30,918.6	$33,078.6	$22,716.5	$34,041.1
		Actual losses			
		March, 24	$190.0		
		March, 27	$2,993.2		
		April, 14	$66,607.4		

Notes: (1) In Panels (a) and (b), backtesting was carried out by leaving the last five years of data for prediction. A p-value less than 5 percent is taken as evidence against the null hypothesis.

Table 4 Other U.S Indices

(a) Population quantiles

	95% **cond. normal**	95% **cond. t**	95% **cond EVT**	95% **unc. EVT**	99% **cond. normal**	99% **cond. t**	99% **cond EVT**	99% **unc. EVT**	99.5% **cond. normal**	99.5% **cond. t**	99.5% **cond EVT**	99.5% **unc. EVT**
S&P 500												
% error	4.92%	5.55%	5.19%	6.05%	1.86%	1.36%	1.08%	1.25%	1.29%	0.67%	0.52%	0.73%
binomial test	–0.25	1.74	0.62	3.33	5.96	2.48	0.59	1.75	7.79	1.64	0.21	2.26
p-value	0.60	0.04	0.27	0.00	0.00	0.01	0.28	0.04	0.00	0.05	0.58	0.01
Wilshire 5000												
% error	6.88%	7.25%	6.37%	9.01%	2.54%	1.76%	1.29%	1.81%	1.86%	1.04%	0.62%	0.88%
binomial test	3.80	4.53	2.76	8.08	6.79	3.36	1.30	3.59	8.50	3.34	0.75	2.37
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.23	0.01
Russell 3000												
% error	5.56%	6.46%	5.52%	6.96%	2.16%	1.55%	1.30%	1.41%	1.62%	0.90%	0.47%	0.58%
binomial test	1.34	3.52	1.25	4.74	6.16	2.92	1.58	2.15	8.39	3.00	–0.23	0.58
p-value	0.09	0.00	0.10	0.00	0.00	0.00	0.06	0.02	0.00	0.00	0.41	0.28
Total rejections by quantile	1	3	1	3	3	3	0	3	3	2	0	2

(b) Empirical quantiles

	95%	99%	99.5%
S&P 500			
% error	5.17%	1.21%	0.67%
binomial test	0.55	1.46	1.64
p-value	0.29	0.07	0.05
Wilshire 5000			
% error	6.47%	1.45%	0.78%
binomial test	2.96	1.98	1.72
p-value	0.00	0.02	0.04
Russell 3000			
% error	5.52%	1.26%	0.54%
binomial test	1.25	1.39	0.31
p-value	0.10	0.08	0.38
Total rejections null by quantile	1	1	1

Notes: 1) The sample periods for the S&P 500, the Wilshire 5000, and the Russell 3000 are, respectively, January 1980-December 2002, January 1991-December 2002, and January 1988-December 2002. All returns series are daily. 2) The first four years of the data are left for backtesting in each case. A p-value less than 5 percent is taken as evidence against the null hypothesis. 3) Data source: Bloomberg.

Table 5 Descriptive Statistics of Daily Stock Returns outside the U.S.

(a) Latin America

	BOVESPA	**IPSA**	**MERVAL**	**IPC**
Mean	0.004	0.001	0.001	0.001
Median	0.003	0.000	0.000	0.000
Std. Dev.	0.035	0.013	0.031	0.018
IQ range	0.032	0.014	0.028	0.019
Minimum	–0.395	–0.085	–0.185	–0.143
Maximum	0.345	0.090	0.262	0.122
Kurtosis	24.441	7.507	11.053	7.965
Skewness	–0.129	0.212	0.892	–0.030
p-value JB test	0.000	0.000	0.000	0.000
Observations	2,710	3,238	3,207	2,741
Sample period	Jan 92-Dec 02	Jan 90-Dec 02	Jan 90-Dec 02	Jan 92-Dec 02

(b) Asia

	Nikkei-225	**Hang Seng**	**KLSE Comp.**	**KOSPI-200**	**STI**
Mean	0.000	0.000	0.000	0.000	0.000
Median	0.000	0.001	0.000	–0.001	0.000
Std. Dev.	0.014	0.019	0.019	0.022	0.013
IQ range	0.014	0.016	0.015	0.021	0.012
Minimum	–0.161	–0.405	–0.242	–0.151	–0.160
Maximum	0.124	0.172	0.208	0.108	0.129
Kurtosis	11.106	73.705	33.406	6.570	16.997
Skewness	–0.077	–3.333	0.412	0.050	–0.406
p-value JB test	0.000	0.000	0.000	0.000	0.000
Observations	4,676	3,957	2,214	3,174	3,730
Sample period	Jan 84-Dec 02	Jan 87-Dec 02	Jan 94-Dec 02	Jan 90-Dec 02	Jan 88-Dec 02

(c) Europe

	DAX-30	**CAC-40**	**FTSE-250**	**IBEX-35**
Mean	0.000	0.000	0.000	0.000
Median	0.001	0.000	0.001	0.001
Std. Dev.	0.014	0.014	0.008	0.014
IQ range	0.015	0.016	0.007	0.016
Minimum	–0.099	–0.077	–0.120	–0.073
Maximum	0.076	0.068	0.071	0.063
Kurtosis	6.997	5.371	30.653	5.313
Skewness	–0.278	–0.121	–1.951	–0.160
p-value JB test	0.000	0.000	0.000	0.000
Observations	3,009	3,260	4,286	2,760
Sample period	Jan 91-Dec 02	Mar 90-Mar 03	Jan86-Dec 02	Feb 92-Feb 03

Notes: 1) JB test stands for Jarque-Bera normality test. 2) Data source: Bloomberg.

Table 6 Backtesting Results of Stock Markets outside the U.S.

(a) Latin America's population quantiles

	95% cond. normal	**95% cond. t**	**95% cond EVT**	**95% unc. EVT**	**99% cond. normal**	**99% cond. t**	**99% cond. EVT**	**99% unc. EVT**	**99.5% cond. normal**	**99.5% cond. t**	**99.5% cond. EVT**	**99.5% unc. EVT**
BOVESPA												
% error	5.74%	6.03%	5.68%	3.59%	1.80%	1.10%	1.10%	0.75%	1.04%	0.75%	0.64%	0.41%
binomial test	1.40	1.95	1.29	–2.68	3.32	0.42	0.42	–1.03	3.20	1.49	0.81	–0.56
p-value	0.08	0.03	0.10	0.00	0.00	0.34	0.34	0.15	0.00	0.07	0.21	0.29
IPSA												
% error	4.86%	5.04%	5.93%	4.90%	1.25%	0.98%	1.03%	1.03%	0.98%	0.67%	0.62%	0.58%
binomial test	–0.31	0.08	2.02	–0.21	1.18	–0.09	0.12	0.12	3.23	1.13	0.83	0.53
p-value	0.38	0.47	0.02	0.42	0.12	0.46	0.45	0.45	0.00	0.13	0.20	0.30
MERVAL												
% error	6.32%	6.81%	5.78%	4.65%	2.53%	1.90%	1.31%	0.81%	1.99%	1.04%	0.86%	0.41%
binomial test	2.85	3.92	1.68	-0.76	7.22	4.24	1.46	–0.89	9.91	3.59	2.39	–0.63
p-value	0.00	0.00	0.05	0.22	0.00	0.00	0.07	0.19	0.00	0.00	0.01	0.27
IPC												
% error	4.92%	4.86%	5.60%	5.03%	1.43%	1.03%	1.03%	0.80%	1.03%	0.69%	0.57%	0.40%
binomial test	–0.16	–0.27	1.16	0.06	1.80	0.12	0.12	–0.84	3.14	1.10	0.43	–0.59
p-value	0.44	0.39	0.12	0.48	0.04	0.45	0.45	0.20	0.00	0.13	0.34	0.28
Total rejections null by quantile	1	2	2	1	3	1	0	0	4	1	1	0

(b) Latin America's empirical quantiles

	95%	**99%**	**99.5%**
BOVESPA			
% error	5.56%	1.16%	0.70%
binomial test	1.07	0.66	1.15
p-value	0.14	0.25	0.13
IPSA			
% error	5.93%	1.03%	0.62%
binomial test	2.02	0.12	0.83
p-value	0.02	0.55	0.80
MERVAL			
% error	6.18%	1.13%	0.68%
binomial test	2.55	0.61	1.18
p-value	0.01	0.27	0.12
IPC			
% error	5.43%	1.09%	0.74%
binomial test	0.83	0.36	1.44
p-value	0.20	0.36	0.07
violations	0	0	0
Total rejections null by quantile	2	0	0

(c) Europe's population quantiles

	95%	95%	95%	95%	99%	99%	99%	99%	99.5%	99.5%	99.5%	99.5%
	cond. normal	**cond. t**	**cond. EVT**	**unc. EVT**	**cond. normal**	**cond. t**	**cond. EVT**	**unc. EVT**	**cond. normal**	**cond. t**	**cond. EVT**	**unc. EVT**
DAX-30												
% error	6.38%	6.83%	5.93%	7.93%	1.65%	1.15%	0.90%	2.09%	1.00%	0.75%	0.50%	1.15%
binomial test	2.84	3.76	1.92	6.01	2.90	0.66	–0.46	4.92	3.16	1.57	–0.01	4.11
p-value	0.00	0.00	0.03	0.00	0.00	0.25	0.32	0.00	0.00	0.06	0.50	0.00
CAC-40												
% error	5.75%	5.97%	5.35%	6.41%	1.72%	1.37%	0.93%	1.81%	0.93%	0.66%	0.62%	0.97%
binomial test	1.63	2.11	0.76	3.08	3.46	1.77	–0.34	3.88	2.89	1.10	0.80	3.19
p-value	0.05	0.02	0.22	0.00	0.00	0.04	0.37	0.00	0.00	0.14	0.21	0.00
IBEX-35												
% error	6.39%	6.74%	6.05%	7.48%	1.94%	1.54%	1.26%	1.60%	1.26%	0.86%	0.57%	0.91%
binomial test	2.89	1.10	0.80	3.19	3.96	2.28	1.08	2.52	4.48	2.11	0.42	2.45
p-value	0.00	0.14	0.21	0.00	0.00	0.01	0.14	0.01	0.00	0.02	0.34	0.01
FTSE-250												
% error	5.59%	6.38%	5.49%	5.43%	2.05%	1.53%	1.10%	1.19%	1.53%	0.85%	0.55%	0.61%
binomial test	1.54	3.62	1.30	1.14	6.01	3.03	0.57	1.10	8.33	2.88	0.40	0.90
p-value	0.06	0.00	0.10	0.13	0.00	0.00	0.28	0.14	0.00	0.00	0.34	0.18
Total rejection null by quantile	2	3	1	3	4	3	0	3	4	2	0	3

(d) Europe's empirical quantiles

	95%	**99%**	**99.5%**
DAX-30			
% error	5.78%	1.00%	0.50%
binomial test	1.61	–0.01	–0.01
p-value	0.05	0.49	0.50
CAC-40			
% error	5.26%	0.97%	0.57%
binomial test	0.57	–0.13	0.50
p-value	0.28	0.45	0.31
IBEX-35			
% error	5.82%	1.66%	0.80%
binomial test	0.50	2.76	1.77
p-value	0.31	0.00	0.04
FTSE-250			
% error	5.49%	1.10%	0.49%
binomial test	1.30	0.57	–0.09
p-value	0.10	0.28	0.46
Total rejections null by quantile	0	1	1

(e) Asia's population quantiles

	95% cond. normal	95% cond. t	95% cond. EVT	95% unc. EVT	99% cond. normal	99% cond. t	99% cond. EVT	99% unc. EVT	99.5% cond. normal	99.5% cond. t	99.5% cond. EVT	99.5% unc. EVT
Nikkei-225												
% error	5.27%	6.19%	5.41%	5.84%	1.57%	1.08%	0.78%	0.97%	0.89%	0.38%	0.27%	0.54%
binomial test	0.76	3.33	1.14	2.35	3.47	0.50	–1.32	–0.16	3.38	–1.05	–1.98	0.35
p-value	0.22	0.00	0.13	0.01	0.00	0.31	0.09	0.44	0.00	0.15	0.02	0.36
Hang Seng												
% error	5.46%	5.89%	5.96%	5.59%	1.89%	1.25%	0.98%	1.01%	1.28%	0.71%	0.44%	0.61%
binomial test	1.14	2.24	2.40	1.48	4.85	1.35	–0.13	0.06	6.02	1.60	–0.48	0.82
p-value	0.13	0.01	0.01	0.07	0.00	0.09	0.45	0.52	0.00	0.05	0.32	0.21
KLSE Comp.												
% error	3.75%	4.40%	4.08%	4.81%	1.31%	0.65%	0.65%	0.73%	0.90%	0.41%	0.49%	0.41%
binomial test	–2.00	–0.96	–1.48	–0.30	1.07	–1.22	–1.22	–0.94	1.97	–0.46	–0.05	–0.46
p-value	0.02	0.17	0.07	0.38	0.14	0.11	0.11	0.17	0.02	0.32	0.48	0.32
KOSPI-200												
% error	5.29%	5.65%	6.15%	6.93%	1.60%	1.23%	1.23%	2.05%	1.05%	0.59%	0.68%	0.87%
binomial test	0.62	1.40	2.48	4.14	2.80	1.09	1.09	4.95	3.64	0.61	1.22	2.43
p-value	0.27	0.08	0.01	0.00	0.00	0.14	0.14	0.00	0.00	0.27	0.11	0.01
STI												
% error	4.26%	5.27%	5.60%	5.57%	1.38%	1.02%	1.05%	0.95%	0.95%	0.51%	0.47%	0.58%
binomial test	–1.79	0.66	1.45	1.36	2.01	0.10	0.29	–0.29	3.31	0.07	–0.20	0.61
p-value	0.04	0.25	0.07	0.09	0.02	0.46	0.39	0.39	0.00	0.47	0.42	0.27
Total rejections null by quantile	2	2	1	1	3	0	0	0	4	0	1	0

(f) Asia’s empirical quantiles

	95%	99%	99.5%
Nikkei-225			
% error	5.19%	1.00%	0.41%
binomial test	0.54	0.00	–-0.81
p-value	0.30	0.50	0.21
Hang Seng			
% error	5.83%	1.28%	0.61%
binomial test	2.07	1.53	0.82
p-value	0.02	0.06	0.21
KLSE Composite			
% error	4.08%	0.90%	0.49%
binomial test	–1.48	–0.36	–0.05
p-value	0.07	0.36	0.48
KOSPI-200			
% error	6.20%	1.37%	0.77%
binomial test	2.58	1.73	1.77
p-value	0.00	0.04	0.04
STI			
% error	5.67%	1.20%	0.69%
binomial test	1.62	1.06	1.42
p-value	0.05	0.15	0.08
violations	0	0	0
Total rejections null by quantile	1	0	0

Notes: In Panels (a) through (f), the first four years of the data are left for backtesting in each case. A p-value less than 5 percent is taken as evidence against the null hypothesis

Table 7 Tail Dependence for Daily Returns

(a) Raw Data

	TBill/DJIA	Nasdaq/DJIA	TBill/DJIA	Nasdaq/DJIA		Nasdaq/DJIA
	Left tail		**Right tail**			**Right tail**
k*	353	204	398	478		
$\bar{\chi}$	0.200	0.522	0.249	0.868	χ	0.520
s.e.	0.064	0.107	0.062	0.085	s.e	0.219
t test	–12.525	–4.486	–11.997	–0.539		
p value	0.000	0.000	0.000	0.062		

(b) Filtered data

Left tail

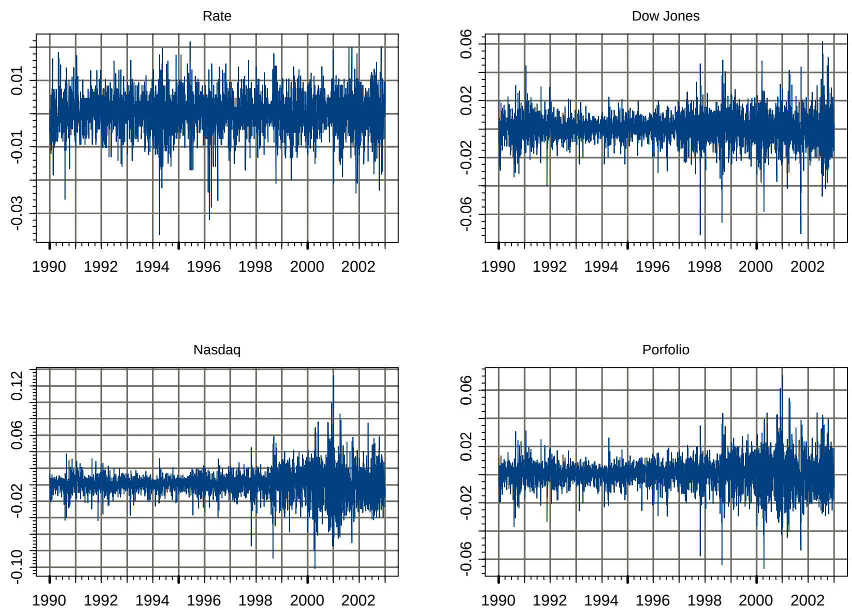
	TBill/DJIA	Nasdaq/DJIA	TBill/DJIA	Nasdaq/DJIA
	Left tail		**Right tail**	
k*	390	387	193	419
$\bar{\chi}$	0.184	0.361	0.132	0.184
s.e.	0.059	0.069	0.081	0.058
t test	–13.616	–9.237	–10.657	–14.115
p value	0.000	0.000	0.000	0.000

Notes: The sample period is 1990-2002. $\bar{\chi}$ and χ are computed based on tail index estimation of Fréchet transformed margins of daily co-exceedances of return pairs, Z=min(S,T). In those cases in which asymptotic dependence cannot be rejected ($\bar{\chi} = 1$), χ is computed under the assumption that $\bar{\chi} = 1$.

FIGURES

Figure 1 Engle's portfolio

(a) Return series



(b) Histograms

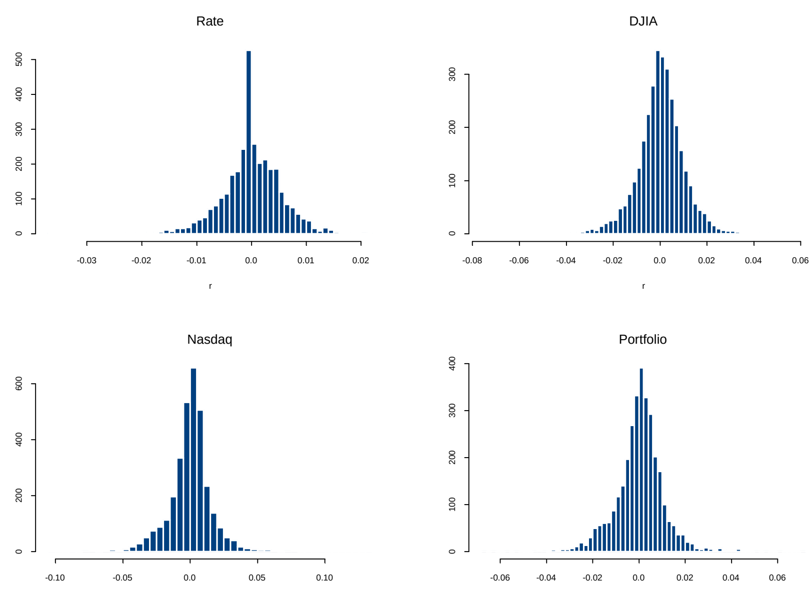


Figure 2 Behavior of Losses

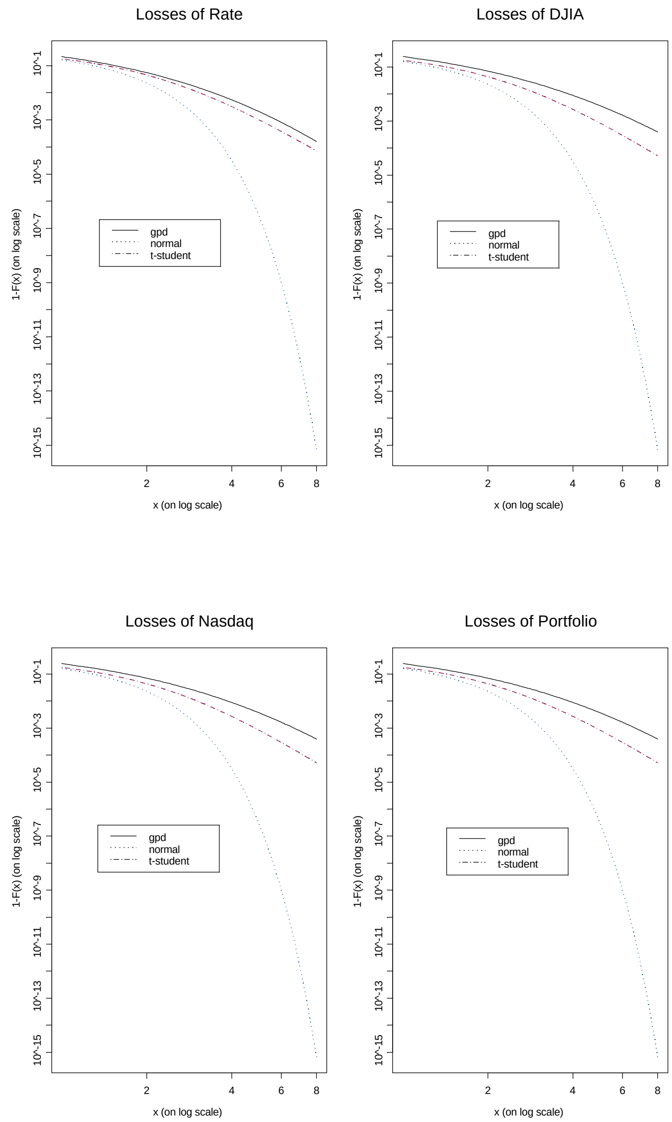
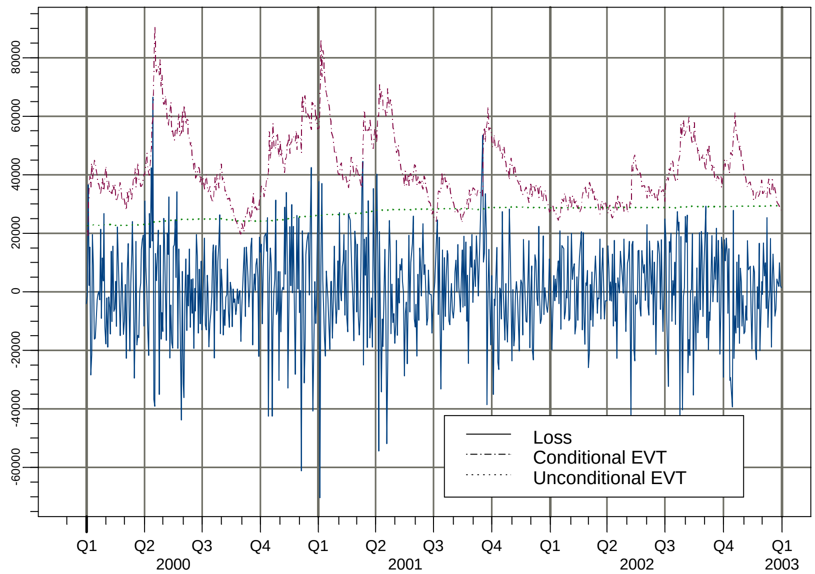


Figure 3 99% Value at Risk and Portfolio Losses: 2000-2002



Note: The last three years of data are left for prediction.

Figure 4 Evolution of the Wilshire 5000 (1991-2002) and Russell 3000 (1988-2002) Returns

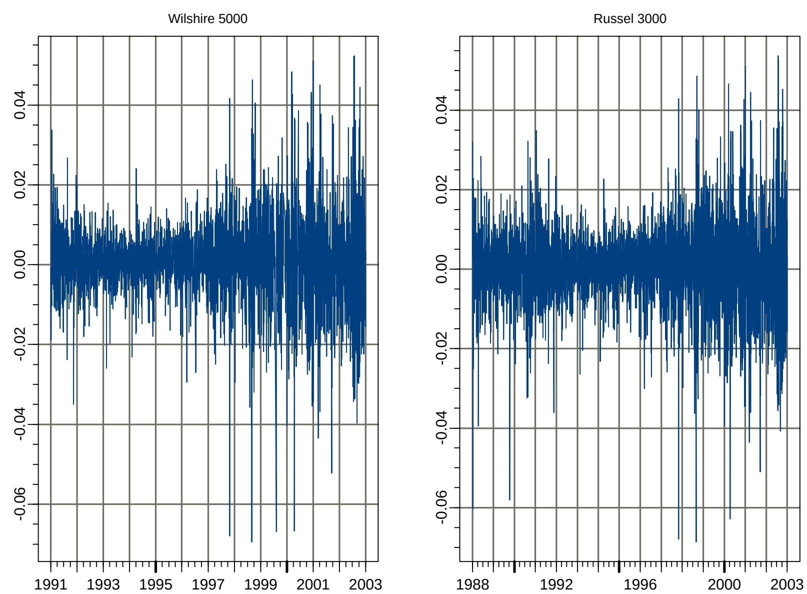
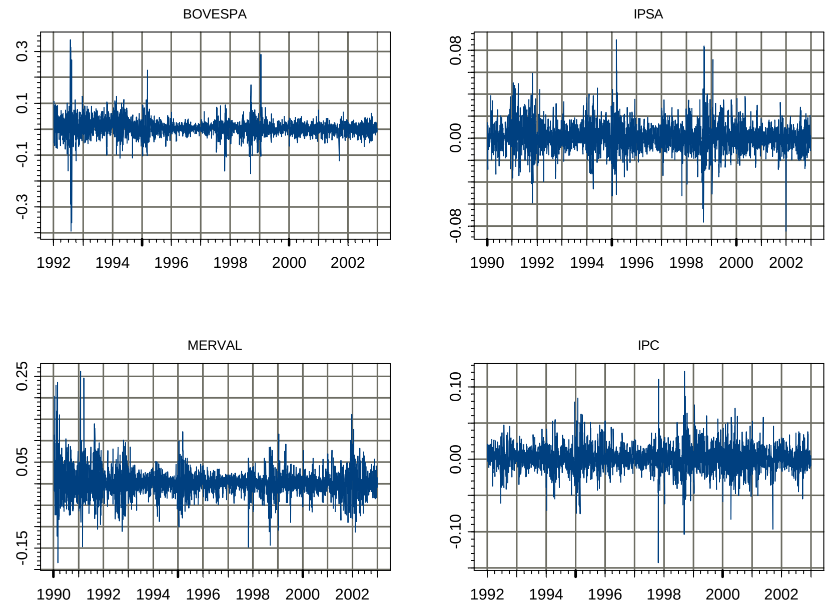


Figure 5 Evolutions of Stock Indices Returns outside the U.S.

(a) Latin America



(b) Europe

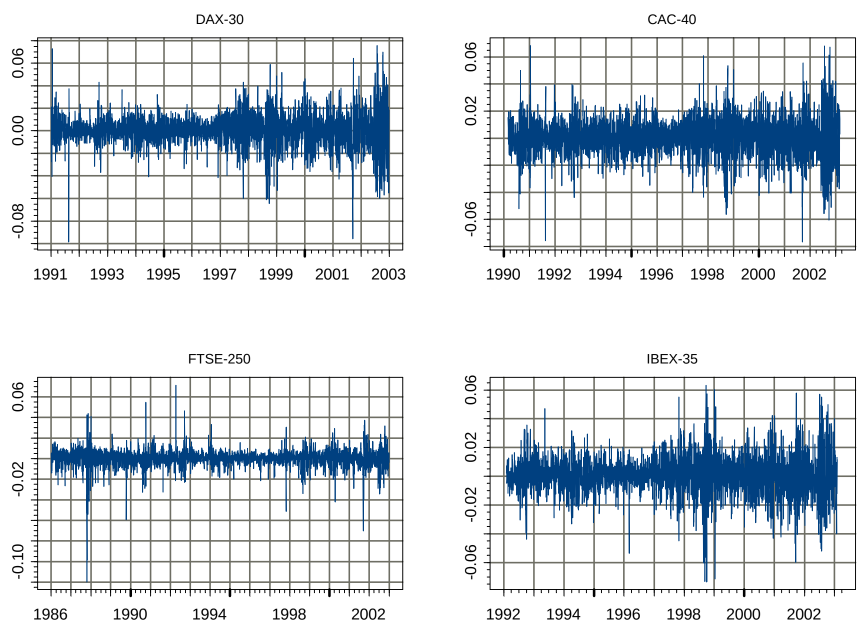


Figure 5 Continued

(c) Asia

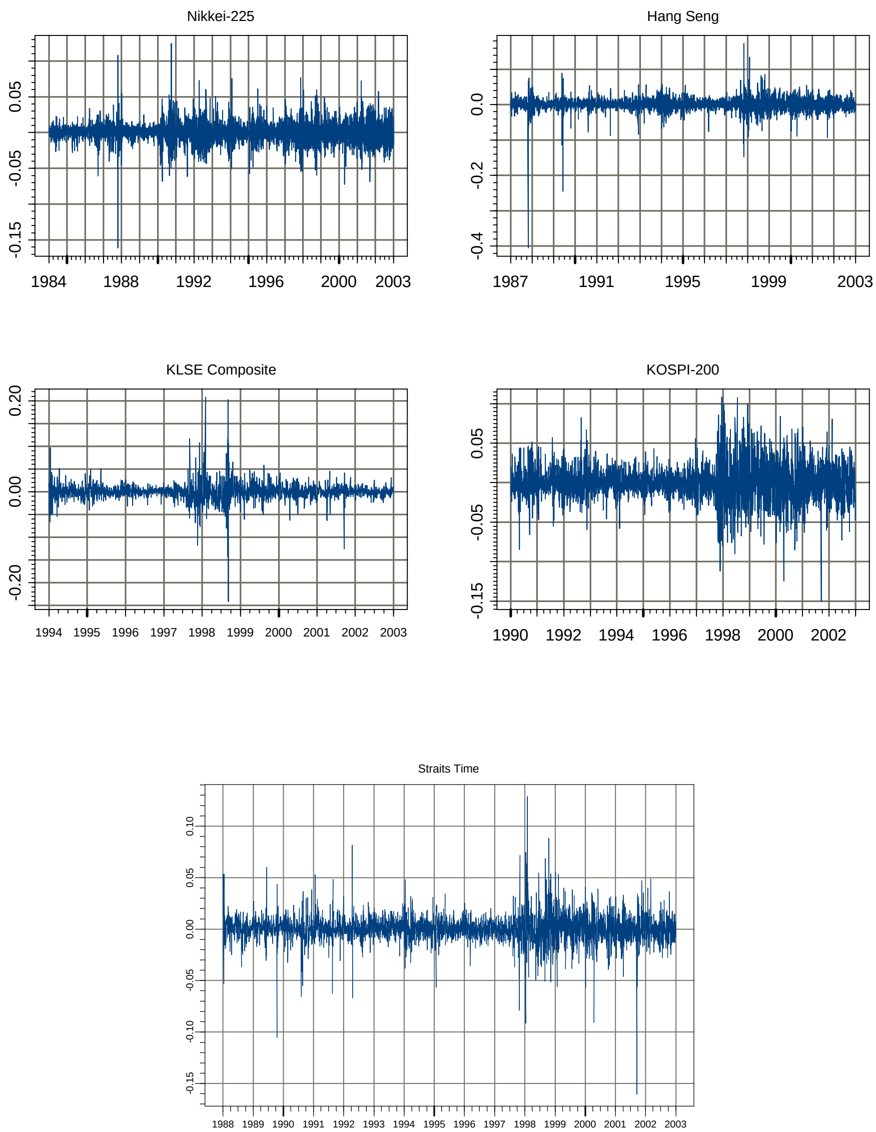


Figure 6 Scatter Plots of Returns: DJIA, T-Bill and Nasdaq

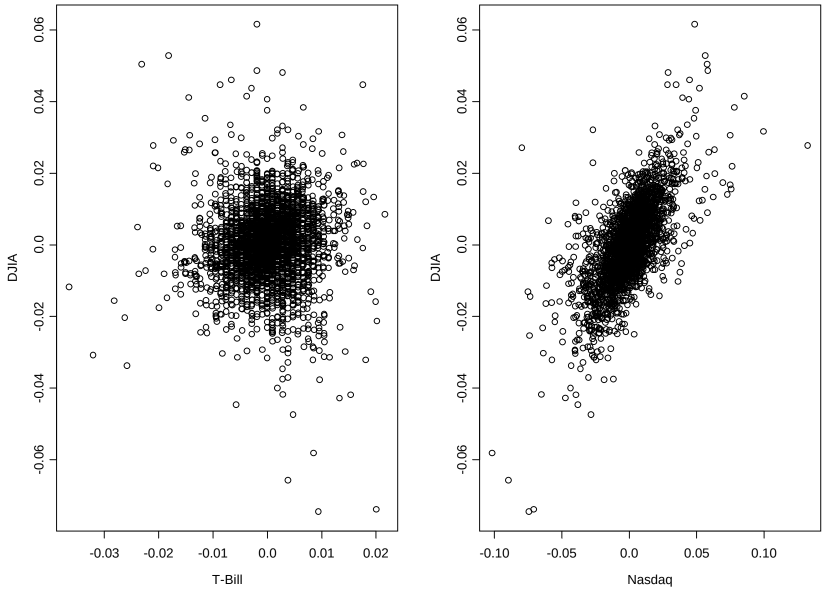
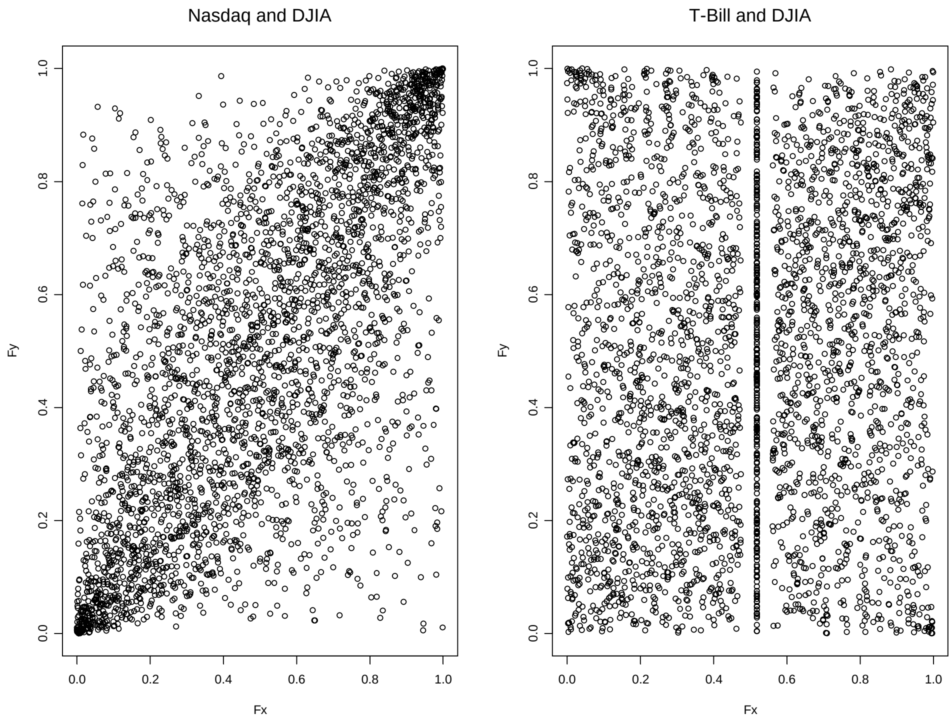


Figure 7 Tail dependence

(a) Negative returns



(c) Negative standardized residuals

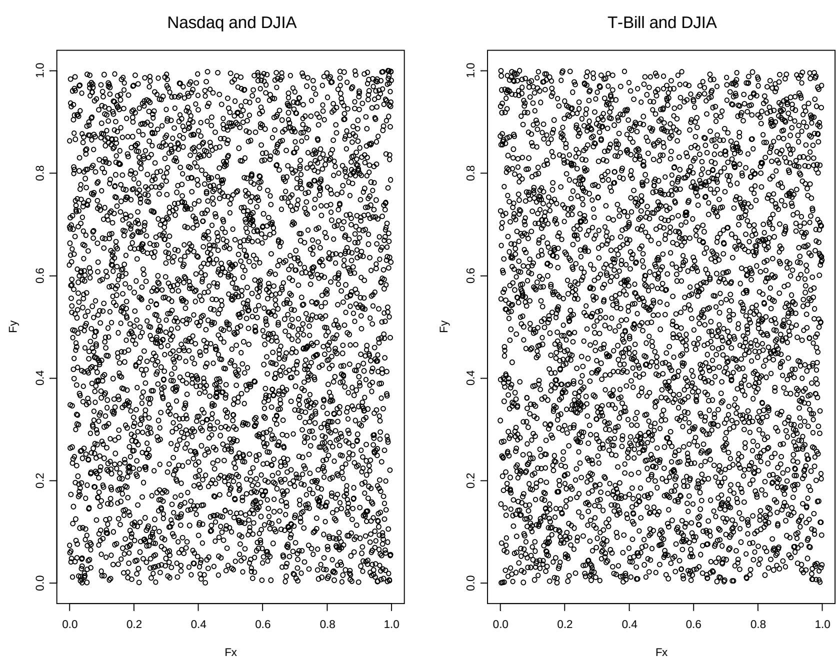
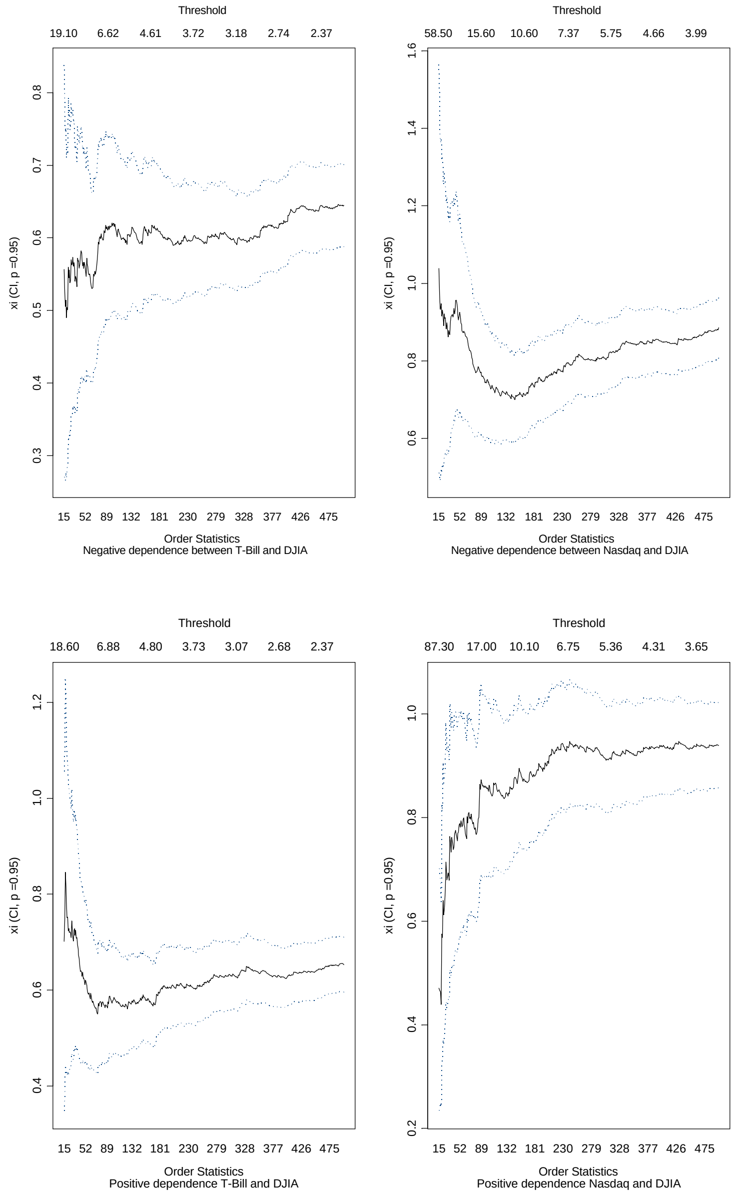
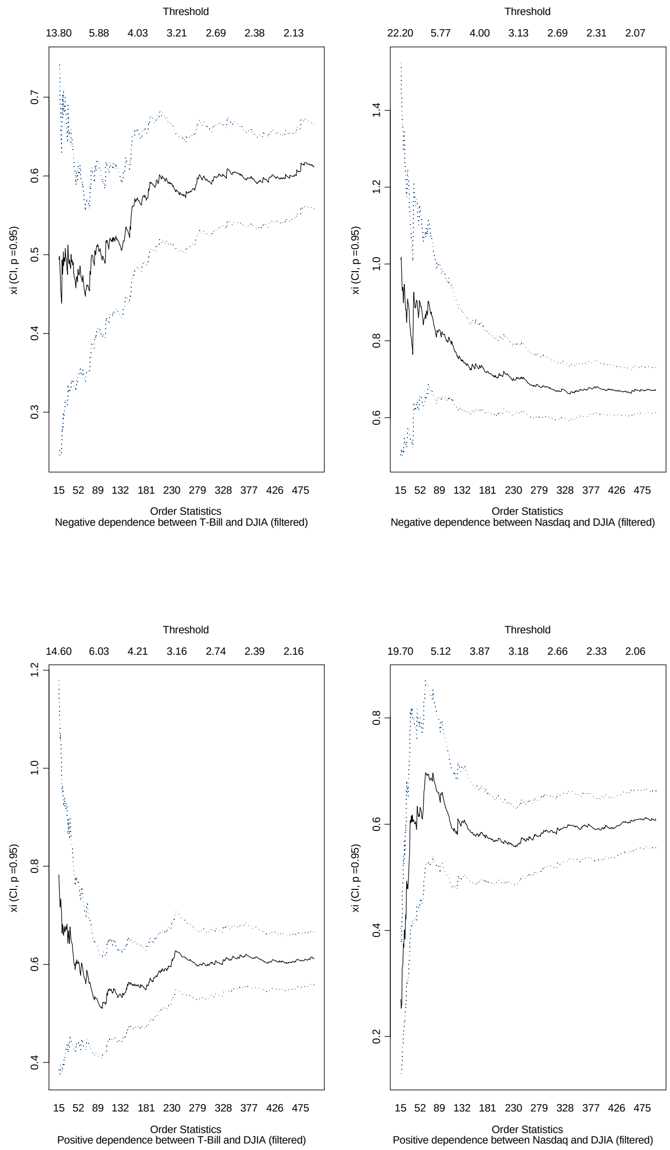


Figure 8 Hill Estimator for Tail Dependence

(a) Raw data



(b) Standardized residuals



Notes: The dotted lines are 95-percent confidence bands.