EXTREME VALUE THEORY AND VALUE AT RISK

Viviana Fernandez[1]

Abstract

Value at Risk (VaR) is a measure of the maximum potential change in value of a portfolio of financial assets with a given probability over a given time horizon. VaR 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. A common practice is to compute VaR by assuming that changes in value of the portfolio are normally distributed, conditional on past information. However, assets returns usually come from fat-tailed distributions. Therefore, computing VaR under the assumption of conditional normality can be an important source of error. We illustrate this point for some return series key to the Chilean financial market by resorting to extreme value theory (EVT) and GARCH-type models. In addition, we show that dynamic estimation of empirical quantiles can also give more accurate VaR estimates than quantiles of a standard normal.

JEL classification: C22, G10

Keywords: GARCH, VaR, EVT

[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.

I Introduction

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.

The assumption made about the cumulative distribution function of the price change is key to VaR calculation. Some available methods are the following: 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 with υ degrees of freedom, standardized quantiles are used. Quantile estimation provides a non-parametric estimate of VaR. It does not make any assumption of 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 paper, we follow McNeil and Frey (2000) and estimate assets volatility with GARCH-type models and compute tails distributions of GARCH innovations by EVT. This allows us to compute conditional quantiles (i.e., VaR), and compare the EVT approach to other alternatives, such as conditional normal, t, and non-parametric quantiles. Our results show that EVT outdoes a GARCH model with normal innovations by far, and that it gives similar results to a GARCH model with t innovations, as long as innovations come from a relatively symmetric and fat-tailed distribution. In turn, GARCH models with non-parametric estimation of quantiles give also more accurate VaR estimates than the assumption of conditional normality. And, they have the advantage of being easy to compute.

The relevance of this paper is the following. Value at risk has recently become a subject of major importance to the Chilean financial system. In particular, last October the Superintendence of Financial Assets and Insurance of Chile (*Superintendencia de Valores y Seguros*) gave special instructions to insurance and re-insurance companies on how to asses monthly the market risk of all their financial assets and real state using VaR.[2]

[2] Excluded from this computation are Chilean peso-denominated and inflation-linked financial assets whose maturity does not exceed one year. (See *Normal de Caracter General No. 148, Superintendencia de Valores y Seguros*, available at www.svs.cl).

Consequently, a better understanding of value at risk and the drawbacks involved in the traditional ways of computing it are worth discussing. In addition, the more advanced techniques presented in this paper deserved to be taken into consideration for refinements government authorities might consider in the future. [3]

This paper is organized as follows. Section II presents a theoretical background on extreme value theory. Section III presents a description of the data and our estimation results. Finally, Section IV summarizes our main findings.

II Theoretical Background

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), $F(x)=Pr(X_i \le x)$. Examples of random risks are negative returns on financial assets or portfolios, operational losses, catastrophic insurance claims, credit losses, natural disasters such as floods, 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

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

[3] Page 11 of *Normal de Caracter General No. 148* states that all observations must be within three standard deviations from the average return in a particular month. If an observation does not meet this requirement, it must be accordingly truncated. This procedure is certainly an arbitrary way to deal with outliers.

An asymptotic approximation to $F^n(x)$ is based on the Fisher-Tippet theorem. Given that x<x+, where x+ is the upper end-point of F (that is, the smallest value of x such that F(x)=1), $F^n(x)\to 0$ as $n\to\infty$, the asymptotic approximation is based on the standardized maximum value

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

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):

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

The parameter ξ is a shape parameter and determines the tail behavior of $G_\xi(z)$. If Z_n converges to $G_\xi(z)$, then Z_n is said to be in the domain of attraction of $G_\xi(z)$. The shape parameter ξ is in turn determined by the tail behavior of the cdf of the underlying data, F. If the tail of F declines exponentially, then $G_\xi(z)$ is of the Gumbel type and $\xi=0$. In this case, distributions in the domain of attraction of $G_\xi(z)$ are of the thin-tailed type, such as the normal, log-normal, exponential, and gamma. If the tail of F declines by a power function instead, then $G_\xi(z)$ is of the *Fréchet* type and $\xi>0$. Distributions in the domain of attraction of $G_\xi(z)$ are called fat tailed distributions, which include the Pareto, Cauchy, Student-t, and mixtures models. Finally, if the tail of F is finite then $G_\xi(z)$ is of the Weibull type and $\xi<0$. Distributions in the domain of attraction of $G_\xi(z)$ are in this case distributions with bounded support, such as the 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) and expected shortfall (ES) estimates that are easy to compute. As we know, VaR (i.e., the qth quantile of F) and ES (i.e., the average loss given that VaR has been exceeded), are commonly used risk measures.

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 that the cdf in (2) converges to (3), 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)

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

where $\beta(u)>0$, and $y\geq0$ when $\xi\geq0$, and $0\leq y\leq-\beta(u)/\xi$ when $\xi<0$.

For a given value of u, the parameters ξ, μ, and σ of the GEV distribution determine the parameters ξ and $\beta(u)$. In particular, the shape parameter ξ is independent of u, and it is the same for both the GEV and GDP distributions. If $\xi>0$, F is in the Fréchet family and $H_{\xi,\beta(u)}$ is a Pareto distribution; if $\xi=0$, F is in the Gumbell family and $H_{\xi,\beta(u)}$ is an exponential distribution; and, if $\xi<0$, F is in the Weibull family and $H_{\xi,\beta(u)}$ is a Pareto type II distribution. In most applications of risk management, the data comes from a heavy-tailed distribution, so that $\xi>0$. In this case, $E(X^k)=\infty$ for $k\geq 1/\xi$. For example, if $\xi=0.5$ the distribution of losses has an infinite variance.

Estimates of the parameters ξ and $\beta(u)$ can be obtained from expression (5) by the method of maximum likelihood (ml). In particular, let x_1, x_2, …, x_n be an iid sample of losses with unknown cdf F. For a given high threshold u, extreme values are those x_i such that $x_i-u>0$. Let us denote these values as $x_{(1)}$, $x_{(2)}$,…, $x_{(k)}$, and define the excesses as $y_i-x_{(i)}-u$, i=1, 2,.., k. If u is large enough, then y_1, y_2,…,y_k may be thought of as a random sample from a GDP distribution with unknown parameters ξ and β. (Hereafter, for simplicity the argument of β is omitted). For $\xi\neq 0$, the log-likelihood for an iid sample is given by

$$L(\xi,\beta) = -k\ln(\beta) - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{k}\ln\left(1+\frac{\xi y_i}{\beta}\right) \tag{6}$$

with $y_i\geq 0$ when $\xi>0$ and $0\leq y_i\leq -\beta/\xi$. For $\xi=0$, the log-likelihood function simplifies to

$$L(\xi,\beta) = -k\ln(\beta) - \frac{1}{\beta}\sum_{i=1}^{k} y_i \ .$$

The asymptotic properties of ml estimates apply here as usual.

In order to estimate the tails of the loss distribution, we use the result that, for a sufficiently high threshold u, $F_u(y)\approx G_{\xi,\beta(u)}(y)$. Now, by setting $x=u+y$, an approximation of F(x), for x>u, can be obtained from equation (4)

$$F(x) = (1-F(u))G_{\xi,\beta(u)}(y) + F(u) \tag{7}$$

An estimate of F(u) can be obtained non-parametrically by means of the empirical cdf

$$\hat{F}(u) = \frac{n-k}{n} \tag{8}$$

where k represents the number of exceedences over the threshold u. After substituting (7) into (8), we get the following estimate for F(x)

$$\hat{F}(x) = 1 - \frac{k}{n}\left(1 + \hat{\xi}\frac{(x-u)}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}} \tag{9}$$

where $\hat{\xi}$ and $\hat{\beta}$ are the ml estimates of ξ and β, respectively.

As mentioned earlier, two commonly used risk measures are the value at risk (VaR) and the expected shortfall (ES). Both are usually computed for confidence levels between 95 and 99.5 percent. That is, for $0.95 \leq q < 1$, VaR_q is the qth quantile of the distribution F

$$VaR_q = F^{-1}(q) \tag{10}$$

where F^{-1} is the inverse function of F. For $q > F(u)$, an estimate of (10) can be obtained from (9) by solving for x

$$\widehat{VaR}_q = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{1-q}{k/n}\right)^{-\hat{\xi}} - 1\right) \tag{11}$$

The expected shortfall is the expected loss, given that VaR_q is exceeded

$$ES_q = E(X \mid X > VaR_q) = VaR_q + E(X - VaR_q \mid X > VaR_q) \tag{12}$$

The expression $E(X-VaR_q|X>VaR_q)$ is the mean of the excess distribution over the threshold VaR_q. It can be shown that (see, for example, Coles, 2001)

$$E(X - VaR_q \mid X > VaR_q) = \frac{\beta + \xi(VaR_q - u)}{1-\xi} \tag{13}$$

provided that $\xi<1$. From equations (11) through (13), we obtain an approximation to ES_q

$$\hat{ES}_q = \frac{\hat{VaR}_q}{1-\hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1-\hat{\xi}} \tag{14}$$

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

<u>Step 1</u>: 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.

<u>Step 2</u>: 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 $q\geq0.95$.

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

$$r_t = \mu + \sigma_t Z_t \tag{15}$$

where μ is a constant term and Z_t are iid innovations with zero mean and unit variance, and marginal distribution $F_Z(z)$.

The conditional variance of the mean-adjusted series $\varepsilon_t = r_t - \mu$ follows a GARCH(1,1) process

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \gamma \sigma_{t-1}^2 \tag{16}$$

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 iid observations becomes

$$L(\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\log(\sigma_t) - \frac{1}{2}\sum_{t=1}^{n}\frac{(r_t - \mu)^2}{\sigma_t} \tag{17}$$

Standardized residuals can be computed after maximizing (17) with respect to the unknown parameters μ, β_0, β_1, and γ

$$(z_{t-n+1}, z_{t-n+2,}, \ldots, z_t) = \left(\frac{r_{t-n+1} - \hat{\mu}}{\hat{\sigma}_{t-n+1}}, \frac{r_{t-n+2} - \hat{\mu}}{\hat{\sigma}_{t-n+2}}, \ldots, \frac{r_t - \hat{\mu}}{\hat{\sigma}_t}\right) \tag{18}$$

where $\hat{\mu}$ and $\{\hat{\sigma}_{t-n+1}, \hat{\sigma}_{t-n+2}, \ldots, \hat{\sigma}_t\}$ are the pseudo-maximum likelihood estimates.

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

$$\hat{\sigma}_{t+1}^2 = \hat{\beta}_0 + \hat{\beta}_1 \hat{\varepsilon}_t^2 + \hat{\gamma} \hat{\sigma}_t^2 \tag{19}$$

where $\hat{\varepsilon}_t = r_t - \hat{\mu}$.

For a one-day horizon, estimates of the dynamic risk measures are

$$\begin{aligned} \widehat{\mathrm{VaR}}^t_q &= \hat{\mu} + \hat{\sigma}_{t+1}\, \widehat{\mathrm{VaR}}(Z)_q \\ \widehat{\mathrm{ES}}^t_q &= \hat{\mu} + \hat{\sigma}_{t+1}\, \widehat{\mathrm{ES}}(Z)_q \end{aligned} \tag{20}$$

where $\widehat{\mathrm{VaR}}_q$ and $\widehat{\mathrm{ES}}_q$ are given by equations (11) and (14), respectively, and $\hat{\sigma}_{t+1} = \sqrt{\hat{\beta}_0 + \hat{\beta}_1 \hat{\varepsilon}_t^2 + \hat{\gamma}\hat{\sigma}_t^2}$.

It is important to mention that, even if Z_t is not truly normally distributed, the maximization of (17) still provides consistent and asymptotically normal estimates (see, for example, Engle and Gonzalez-Rivera, 1991). This result is the one on which McNeil and Frey's approach relies upon.

III Data

3.1 Descriptive Statistics and Preliminary Results

We worked with the following series at a daily frequency, which are particularly relevant to the Chilean financial market: the Index Price of Selective Stocks (IPSA), which gathers the forty most traded shares on the Santiago Exchange, Chile (sample period: 1990-November 2002); the Chilean peso/US dollar exchange rate (sample period: 1988-2002); the spot price of copper (sample period: 1998-2002); and, a proxy for a one-year zero coupon bond traded domestically (sample period: 1993-2001).

Our proxy of a 1-year zero-coupon bond was constructed as follows. Daily data of the average interest rate paid on 1-year bank (inflation-linked) deposits are available from Bloomberg since 1993 approximately. Taking this as a reasonable proxy for a 1-year riskless rate, the price on day t of a zero-coupon bond that pays $1 in one year is

$P_t = \frac{1}{1+r_t}$, where r_t is the annualized 1-year rate on day t. The return on the zero-coupon bond on day t is defined as $(P_t/P_{t-1}-1)*100$. Data on zero-coupon bonds, with maturities of two, three and four years, is available in the Chilean financial market on a daily frequency only since December 2001. Therefore, the time series were too short for carrying out our estimation, and we preferred the proxy just described.

All computations of this and the following sections were programmed in the S+FinMetrics module of S-Plus 6.1. Table 1 shows some descriptive statistics of the daily returns on the above series. The mean return is close to zero for all of the four series. However, they differ considerably in terms of both volatility and tail thickness, which are quantified by the interquartile range and kurtosis, respectively. In particular, the Ch$/US$ exchange rate and the proxy of a zero-coupon bond return series exhibit both very low volatility and high kurtosis when compared with the IPSA and copper return series.

[Table 1 about here]

Partial autocorrelations are mostly statistically insignificant for all return series from the thirteen lag onwards. This in turn translates into rejection of the null hypothesis of a unit root. The assumption of normally distributed returns is strongly rejected by all series, as Figure 1 suggests.

[Figure 1 about here]

The next step was to fit a GARCH (1, 1) model of the sort described in the previous section, and compute the standardized residuals for every return series according to expression (18). Specification tests are reported in Table 2. The Lagrange multiplier test (TR^2) applied to each series cannot reject the null hypothesis of “no residual ARCH”. The Ljung-Box test in turn finds no evidence to reject the null hypothesis of no autocorrelation.

[Table 2 about here]

Given that a GARCH(1,1) specification cannot be rejected in any case, we next set a threshold u for each individual series and assume that the (standardized) residuals exceeding u follow a generalized Pareto distribution (GDP). Table 3 shows GDP estimates for both tails of the innovations for each of the four return series. In each case, the number of points above the threshold u is 10 percent of the observations in each tail. (This is roughly the same percentage used by McNeil and Frey, 2000).

[Table 3 about here]

Except for the innovations of the Ch$/US$ exchange rate returns, the shape parameter ξ for losses (large positive residuals) and gains (large negative residuals) turns out to be statistically insignificant for the other three series. That means that tail distributions in those cases do not departure substantially from the Gumbel type (thin tailed distributions). This can be also measured by the ratio of the expected shortfall to VaR. For the 99-percent quantile, this ratio is 1.15 for a standard normal (see McNeil and Frey, op. cit). As we see, except for the exchange rate return series, the expected shortfall to VaR is around 1.2 for losses and gains. (Although the gains distribution of our proxy of a zero-coupon bond is slightly more heavy tailed). This number is not considerably greater than that of a N(0,1) distribution.

Panels (a) through (d) of Figure 2 shed more light on this issue. For all of the four cases, the normal distribution understates the extent of large losses and gains. The t distribution, on the other hand, overestimates large losses and gains for IPSA, copper, and the 1-year zero coupon bond, but it does only slightly for the Ch$/US$ exchange rate.

[Figure 2 about here]

3.2 Estimation of VaR in-Sample and Out of Sample

In this section, we follow an approach similar to Engle (2001)'s, but we explicitly model the behavior of tails according to the EVT approach described earlier. We used all observations except for the last two years to estimate the 99-percent VaR in-sample. The last two years of the data were used for backtesting. The results are depicted in Panels (a) through (d) of Figures 3.

[Figure 3 about here]

The expected in-sample error is 1 percent. The log-negative returns on the proxy of a zero-coupon bond (sample period: 1993-1999) and copper (sample period: 1998-2000) exceeded VaR 0.87 and 0.99 percent of the time, respectively, which is close to the expected. By contrast log-negative returns on IPSA (sample period: 1990-2000) and the Ch$/US$ exchange rate (sample period: 1988-2000) tended to underestimate losses (0.61 and 0.77 percent, respectively) more often than in the other two cases.

Estimation of the 99-percent VaR out of sample is computed without updating the parameter estimates previously obtained. Likewise, the 99-percent quantile of the innovation distribution for each series is not recalculated either. The time period for backtesting is 2001-2002 for all series except for the proxy of a zero-coupon bond, which is 2000-2001. The value at risk is exceeded by the log-negative returns on IPSA and copper 2 and 5 times, respectively (that is, 0.42% and 0.66% of the time, respectively). This suggests that our measure of risk is rather conservative, especially for IPSA. This is not as true for the Ch$/US$ exchange rate and the proxy of a zero-coupon bond, for which VaR is exceeded 1.6% and 1.43% of the time, respectively.

However, as Engle points out, in this case it not easy to asses how accurate a measure of risk VaR is out of sample. In particular, neither parameter estimates nor quantiles incorporate the new information that becomes available in the backtesting period. That is why in the next section we focus on a backtesting procedure that dynamically adjusts quantiles, and that allows us to conclude statistically which way to compute VaR might be best.

3.3 Dynamic Backtesting

In order to asses the accuracy of the EVT approach and alternative ways to compute VaR, we backtested the method on the four return series described earlier by the following procedure. Let $r_1, r_2, \ldots, r_m$ be a historical return series. The conditional quantile $\hat{r}_q^t$ is computed on t days in the set of T={n, …, m–1) using an n-day window each time. The large of n we set depended on the sample size of each returns series. For IPSA, the Ch$/US$ exchange rate, and our proxy of 1-year zero coupon bond, n took on the value of 974, 992, and 999, respectively, so that we had about the last four years of data for prediction. Given that the daily series of copper only covered the period 1998-2002, we set n=504, so that we had the last two years of data for prediction.

The constant k, which defines the number of exceedences above the threshold u as described in Section II, was set following McNeil and Frey (2000)'s approach. In particular, the authors set k so that the 90th percentile of the innovation distribution is estimated by historical simulation. For instance, for copper we set k=50.

On each day $t \in T$, we estimated a new GARCH(1,1) model and fitted a new generalized Pareto distribution to losses, which were computed from the series of standardized residuals. This procedure, as explained in Section II, is called conditional EVT. In addition, we estimated the unconditional EVT quantile, which corresponds to expression (11).

The conditional normal quantile 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 $z_q = \sqrt{(\upsilon - 2)/\upsilon}\, F_T^{-1}(q)$, where T follows a t-distribution with υ degrees of freedom (υ>2). On each day t, we estimated a GARCH(1,1) model with Student-t innovations and estimated a new υ and new quantiles. The value at risk was computed according to formula (20) for both the normal and t conditional cases.

If Z_t is assumed to be distributed as t with υ degrees of freedom in equation (15), the log-likelihood function of the sample becomes (see, for example, Hamilton, 1994, chapter 21)

$$L(\theta) = n\log\left(\frac{\Gamma(\upsilon+1)/2}{\pi^{1/2}\Gamma(\nu/2)}(\nu-2)^{-1/2}\right) - \frac{1}{2}\sum_{t=1}^{n}\log(\sigma_t) - \frac{(\upsilon+1)}{2}\sum_{t=1}^{n}\log\left(1+\frac{(r_t-\mu)}{\sigma_t(\nu-2)}^{2}\right) \quad (21)$$

where n is the sample size. This is a better approximation to the data generating process in case observed returns appear to come from a (symmetric) fat-tailed distribution. As discussed below, the t distribution works better in VaR estimation than the conditional normal approach.

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

$$\frac{\frac{Y}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{d} N(0,1) \quad (22)$$

where n is the sample size and Y is the number of violations, so that Y/n is the actual proportion of violations. The proportion p is the expected number of violations, under the assumption that the indicator function $I_t \equiv 1_{\{r_{t+1} > r_q^t\}} = 1_{\{Z_{t+1} > z_q\}} \sim Be(p)$. This is a one-tailed test that is asymptotically distributed as N(0,1) (see, for example, Larsen and Marx, 1986, chapter 5). If Y/n<p, we test the null hypothesis of estimating correctly the conditional quantile against the alternative that the method systematically underestimates the conditional quantile. Otherwise, we test the null against the alternative that the method systematically overestimates the conditional quantile.

Panel (a) of Table 4 presents backtesting results based on population quantiles of 95, 99 and 99.5 percent. The conditional t, normal, EVT and the unconditional EVT approaches were computed as described above. Our rule is that the null hypothesis is rejected whenever the p-value of the binomial test is less than 5 percent. Our results show that the conditional normal approach is the one that rejects the null hypothesis most often (in 7 out of 12 cases). As expected, this approach tends to work worse the higher the confidence level. For instance, for the 99.5 percent quantile this approach rejects the null for all return series except for copper.

The conditional t and the EVT approaches are the closest to the mark: the former rejects the null hypothesis only once while the latter rejects it twice. In particular, the conditional t approach beats the conditional EVT approach for the 99.5 percent quantile of our proxy of the zero-coupon bond. In turn the unconditional EVT approach also works well: it rejects the null in 3 out of 12 occasions. Conditional and unconditional EVT estimates along with log-negative returns are depicted in Panels (a) through (d) of Figure 4.

[Figure 4 about here]

We also explore the performance of empirical quantiles in computing VaR, an issue that is not addressed by McNeil and Frey. Panel (b) of Table 4 shows backtesting using empirical quantiles. This procedure is similar to those described above but, instead of parameterizing the tails of the innovation distribution, quantiles are computed from the empirical distribution of standardized residuals each time a new GARCH-model is fitted to the data. This procedure works well (the null hypothesis is rejected in 3 out of 12 cases), and it is easy to compute. In particular, unlike the conditional normal approach, it takes account of the thickness of tails. Again, the poorest fit is for the exchange rate (the null is rejected twice), and for our proxy of a zero-coupon bond return series (the null is rejected once). However, the null is never rejected for copper and IPSA.

[Table 4 about here]

Our results are similar in nature to those find by McNeil and Frey (2000), although for their returns series the conditional EVT approach never rejects the null hypothesis. At this stage, it worth noticing that most rejections of the null hypothesis occur for the exchange rate and our proxy of a zero-coupon bond return series. As Table 1 shows, both are characterized by relatively low volatility (e.g., the interquartile ranges of both series are much lower than those of the IPSA and the copper returns series), and by very high kurtosis.

This can be easily seen in the corresponding histograms of Figure 1 and Panels (a) and (b) of Table 5. For the exchange rate return series the 25 and 75 percent quantiles are −0.12 and 0.16 percent, respectively, and 99.7 percent of the returns are between −2 and 2 percent. The dispersion is even lower for our proxy of a 1-year zero: the 25 and 75 percent quantiles are −0.048 and 0.047 percent, respectively, and 98.7 percent of the returns are between −0.5 and 0.5 percent. In other words, the low dispersion of returns along with the presence of a few outliers would explain the relatively poor fit of tails in these two cases.

In summary, the conditional t and EVT approaches are the best. In addition, good alternatives to the normal approach are the unconditional EVT and the empirical quantiles approaches.

IV Conclusions

Value at Risk (VaR) is a popular measure of market risk, 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.

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. Extreme value theory (EVT) focuses on modeling the tail behavior of a loss distribution using only extreme value rather than the whole data set.

In this paper, we estimate assets volatility with GARCH-type models and compute tails distributions of GARCH innovations by EVT. This allows us to compute conditional quantiles, and compare the EVT approach to other alternatives, such as conditional normal, Student-t, and non-parametric quantiles. Our results show that EVT outdoes a GARCH model with normal innovations by far, and that it gives similar results to a GARCH model with t innovations, as long as innovations come from a relatively symmetric and fat-tailed distribution. In turn, GARCH models with non-parametric estimation of quantiles give also more accurate VaR estimates than the assumption of conditional normality. And, they have the advantage of being easy to compute.

References

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*. Springer series in statistics, Springer-Verlag London Limited.

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

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

Hamilton, J. (1994), *Time Series Analysis*. Princeton University Press. Princeton, New Jersey.

Jorion, P. (2001), *Value at Risk*. Second edition. McGraw Hill, New York.

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

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.

Reiss, R.D. and M. Thomas (2001), *Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser Verlag, second edition. Basel, Switzerland.

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

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

TABLES

Table 1 Descriptive Statistics of Daily Returns

	IPSA	**ChUS exchange rate**	**Copper**	**(Proxy) 1-year zero coupon bond**
# observations	2,742	3,240	1,257	2,209
Mean	0.09%	0.03%	–0.01%	0.00%
Median	0.04%	0.03%	–0.06%	0.00%
Std. Dev.	1.32%	0.36%	1.22%	0.14%
Interquartile range	1.42%	0.28%	1.53%	0.10%
Minimum	–7.66%	–4.82%	–5.38%	–0.89%
Maximum	8.97%	4.34%	6.00%	1.70%
Kurtosis	3.90	24.47	1.64	29.67
Skewness	0.30	–0.06	0.43	0.43
ρ_1	0.251	0.132	–0.093	–0.224
	(0.00)	(0.00)	(0.00)	(0.00)
ρ_2	–0.041	–0.012	0.001	–0.069
	(0.03)	(0.49)	(0.98)	(0.00)
ρ_3	0.003	–0.003	–0.012	–0.053
	(0.86)	(0.88)	(0.67))	(0.01)
ρ_4	0.025	–0.034	0.055	–0.055
	(0.19)	(0.05)	(0.05)	(0.01)
ρ_{13}	0.047	0.016	0.026	0.006
	(1.00)	(1.00)	(1.00)	(1.00)
ρ_{26}	0.016	0.015	0.023	–0.037
	(0.40)	(0.39)	(0.41)	(0.08)
ρ_{60}	0.021	0.010	0.014	0.014
	(0.27)	(0.55)	(0.62)	(0.51)
Jarque-Bera test	1,777.9	80,815.1	179.02	82,315.9
	(0.00)	(0.00)	(0.00)	(0.00)
Augmented Dickey-Fuller test	–14.75	–16.86	–10.54	–12.47
	(0.00)	(0.00)	(0.00)	(0.00)

Notes: IPSA stands for Price Index of Selective Stocks, and gathers the 40 most traded stocks on the Santiago Stock Exchange, Chile. The sample period for IPSA is January 1990-November 2002 (data source: Bloomberg, Central Bank of Chile).; for the Ch$/US$ exchange rate is 1988-2002 (data source: Central Bank of Chile); for copper is 1998-2002 (data source: London Metal Exchange); and, for the proxy of a 1-zero coupon bond is 1993-2001 (data source: Bloomberg). ρ_j represents the autocorrelation coefficient of order j. P-values are between parentheses.

Table 2 Specification Tests for GARCH (1,1) Models

Returns series	Lagrange multiplier test (TR^2) for serial correlation (12 lags)	Test for ARCH effects (12 df)
IPSA	17.27	16.75
	(0.14)	(0.16)
Ch$/US$ exchange rate	0.21	0.22
	(1.00)	(1.00)
Copper	7.53	6.98
	(0.82)	(0.86)
(Proxy) 1-year zero coupon bond	13.02	13.91
	(0.37)	(0.34)

Notes: P-values are between parentheses. The sample periods are the same as those described at the bottom of Table 1.

Table 3 Tails of IPSA, Ch/US Exchange Rate, Copper, and 1-year Zero Coupon Bond Innovations

(a) IPSA: 1990-2000									
Tail	u	$\hat{\xi}$	s.e	$\hat{\beta}$	s.e	Observations	q=99%		
							quantile (x_q)	sfall	sfall/ x_q
Losses	1.528	–0.023	0.074	0.572	0.064	1,408	2.44	2.98	1.22
Gains	1.748	–0.035	0.090	0.542	0.068	1,334	2.58	3.07	1.19
(b) Ch/US exchange rate: 1988-2000									
Tail	u	$\hat{\xi}$	s.e	$\hat{\beta}$	s.e	Observations	q=99%		
							quantile (x_q)	sfall	sfall/ x_q
Losses	1.310	0.300	0.088	0.516	0.060	1,555	2.34	3.52	1.50
Gains	1.403	0.402	0.108	0.475	0.061	1,685	2.51	4.04	1.61
(c) Copper: 1998-2001									
Tail	u	$\hat{\xi}$	s.e	$\hat{\beta}$	s.e	Observations	q=99%		
							quantile (x_q)	sfall	sfall/ x_q
Losses	0.955	–0.029	0.068	0.516	0.054	519	2.31	2.77	1.20
Gains	0.993	–0.036	0.082	0.679	0.079	488	2.72	3.31	1.22
(d) (proxy) 1-year zero coupon bond: 1993-2001									
Tail	u	$\hat{\xi}$	s.e	$\hat{\beta}$	s.e	Observations	q=99%		
							quantile (x_q)	sfall	sfall/ x_q
Losses	1.529	–0.012	0.096	0.613	0.086	981	2.58	3.17	1.23
Gains	1.724	0.033	0.158	0.794	0.156	742	2.89	3.78	1.31

Notes: 'sfall' stands for expected shortfall.

Table 4 Backtesting Results

(a) Population Quantiles

	95% Cond. t	95% Cond. normal	95% Cond. EVT	95% Unc. EVT	99% Cond. t	99% Cond. normal	99% Cond. EVT	99% Unc. EVT	99.5% Cond. t	99.5% Cond. normal	99.5% Cond. EVT	99.5% Unc. EVT
IPSA												
% error	4.82%	4.59%	5.45%	5.00%	0.99%	1.23%	1.14%	1.05%	0.45%	0.86%	0.50%	0.59%
expected	5.00%	5.00%	5.00%	5.00%	1.00%	1.00%	1.00%	1.00%	0.50%	0.50%	0.50%	0.50%
binomial test	–0.39	–0.88	0.98	0.00	–0.04	1.07	0.64	0.21	–0.33	2.42	0.00	0.60
p-value	0.35	0.19	0.16	0.50	0.52	0.14	0.26	0.42	0.63	0.01	0.50	0.27
rejection of null	0	0	0	0	0	0	0	0	0	1	0	0
Exchange rate												
% error	4.88%	3.39%	6.19%	6.96%	0.91%	1.35%	1.17%	1.31%	0.36%	0.80%	0.62%	0.69%
expected	5.00%	5.00%	5.00%	5.00%	1.00%	1.00%	1.00%	1.00%	0.50%	0.50%	0.50%	0.50%
binomial test	–0.28	–3.88	2.87	4.71	–0.47	1.83	0.87	1.64	–1.01	2.24	0.89	1.43
p-value	0.39	0.00	0.00	0.00	0.32	0.03	0.19	0.05	0.16	0.01	0.19	0.08
rejection of null	0	1	1	1	0	1	0	0	0	1	0	0
Copper												
% error	3.58%	3.58%	5.04%	4.38%	0.40%	0.80%	0.80%	0.53%	0.27%	0.40%	0.27%	0.27%
expected	5.00%	5.00%	5.00%	5.00%	1.00%	1.00%	1.00%	1.00%	0.50%	0.50%	0.50%	0.50%
binomial test	–1.79	–1.79	0.05	–0.79	–1.66	–0.56	–0.56	–1.30	–0.91	–0.40	–0.91	–0.91
p-value	0.04	0.04	0.48	0.22	0.05	0.29	0.29	0.10	0.18	0.35	0.18	0.18
rejection of null	1	1	0	0	0	0	0	0	0	0	0	0
Zero												
% error	5.68%	4.87%	5.28%	5.76%	1.06%	2.19%	1.30%	1.62%	0.57%	1.62%	0.89%	0.97%
expected	5.00%	5.00%	5.00%	5.00%	1.00%	1.00%	1.00%	1.00%	0.50%	0.50%	0.50%	0.50%
binomial test	1.10	–0.21	0.44	1.23	0.19	4.20	1.05	2.20	0.34	5.59	1.95	2.36
p-value	0.14	0.42	0.33	0.11	0.42	0.00	0.15	0.01	0.37	0.00	0.03	0.01
rejection of null	0	0	0	0	0	1	0	1	0	1	1	1
rejection of null by quantile	1	2	1	1	0	2	0	1	0	3	1	1

Table 4 Continued

(b) Empirical Quantiles

Quantile	**95%**	**99%**	**99.5%**
IPSA			
% error	5.5%	1.3%	0.6%
expected	5.0%	1.0%	0.5%
binomial test	1.08	1.29	0.91
p-value	0.14	0.10	0.18
rejection of null	0	0	0
Exchange rate			
% error	6.2%	1.4%	0.7%
expected	5.0%	1.0%	0.5%
binomial test	2.87	2.22	1.43
p-value	0.00	0.01	0.08
rejection of null	1	1	0
Copper			
% error	5.2%	0.8%	0.5%
expected	5.0%	1.0%	0.5%
binomial test	0.49	–0.56	0.08
p-value	0.31	0.29	0.47
rejection of null	0	0	0
Zero			
% error	5.2%	1.4%	1.1%
expected	5.0%	1.0%	0.5%
binomial test	0.31	1.34	2.76
p-value	0.38	0.09	0.00
rejection of null	0	0	1
rejection of null by quantile	1	1	1

Notes: (1) Parameter estimates are obtained by the method of maximum likelihood, as described earlier. (2) The value of "1", under the category of "rejection null", indicates that the p-value of the binomial test is less than 5 percent and, hence, the null hypothesis is rejected; and, 0 otherwise.

Table 5 Tabulation of Returns on the Ch$/US$ Exchange rate and the (Proxy) of a Zero-coupon Bond

(a) Ch$/US$ exchange rate

Return value	Count	Percent	Cumulative Count	Cumulative Percent
[–0.06, –0.04)	1	0.03	1	0.03
[–0.04, –0.02)	2	0.05	3	0.08
[–0.02, 0)	1654	44.26	1657	44.34
[0, 0.02)	2071	55.42	3728	99.76
[0.02, 0.04)	8	0.21	3736	99.97
[0.04, 0.06)	1	0.03	3737	100.00
Total	3737	100.00	3737	100.00

(b) (Proxy) of zero-coupon bond

Return value	Count	Percent	Cumulative Count	Cumulative Percent
[–0.01, –0.005)	12	0.52	12	0.52
[–0.005, 0)	925	40.17	937	40.69
[0, 0.005)	1351	58.66	2288	99.35
[0.005, 0.01)	11	0.48	2299	99.83
[0.01, 0.015)	2	0.09	2301	99.91
[0.015, 0.02)	2	0.09	2303	100.00
Total	2303	100.00	2303	100.00

FIGURES

Figure 1 Histograms of Daily Returns

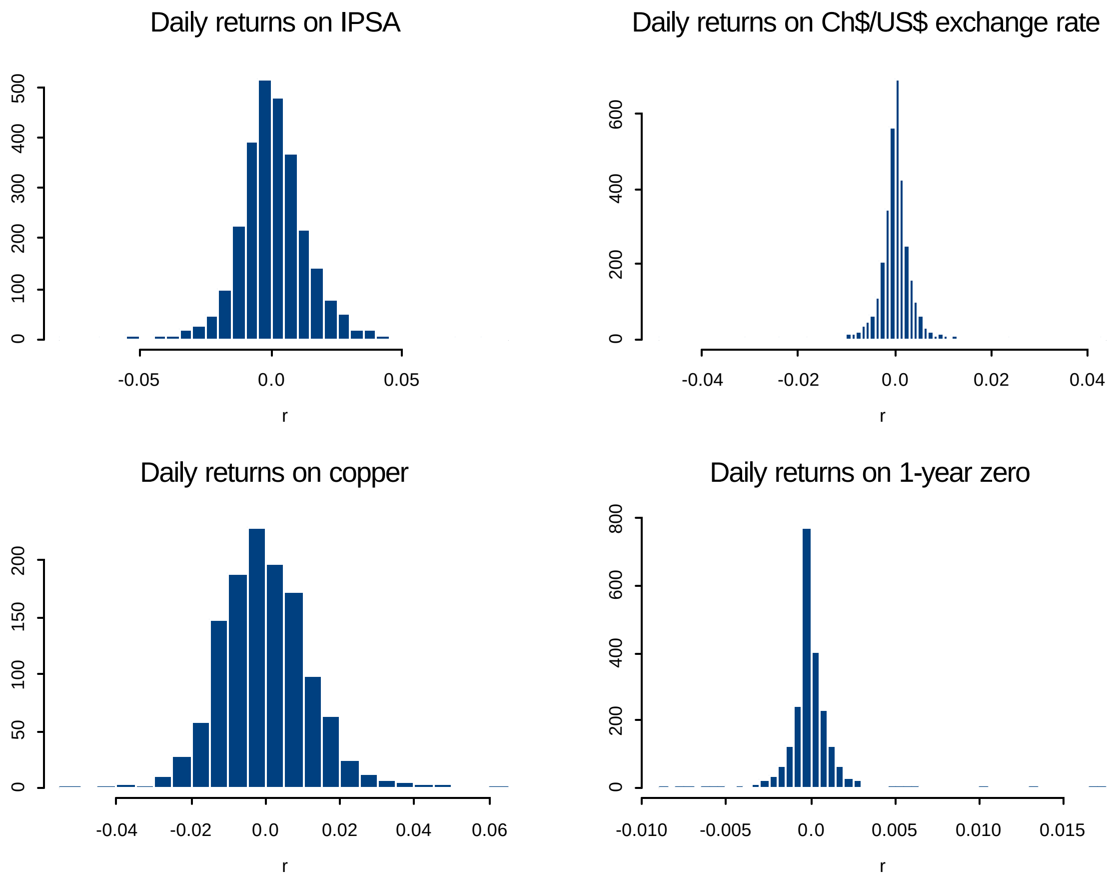
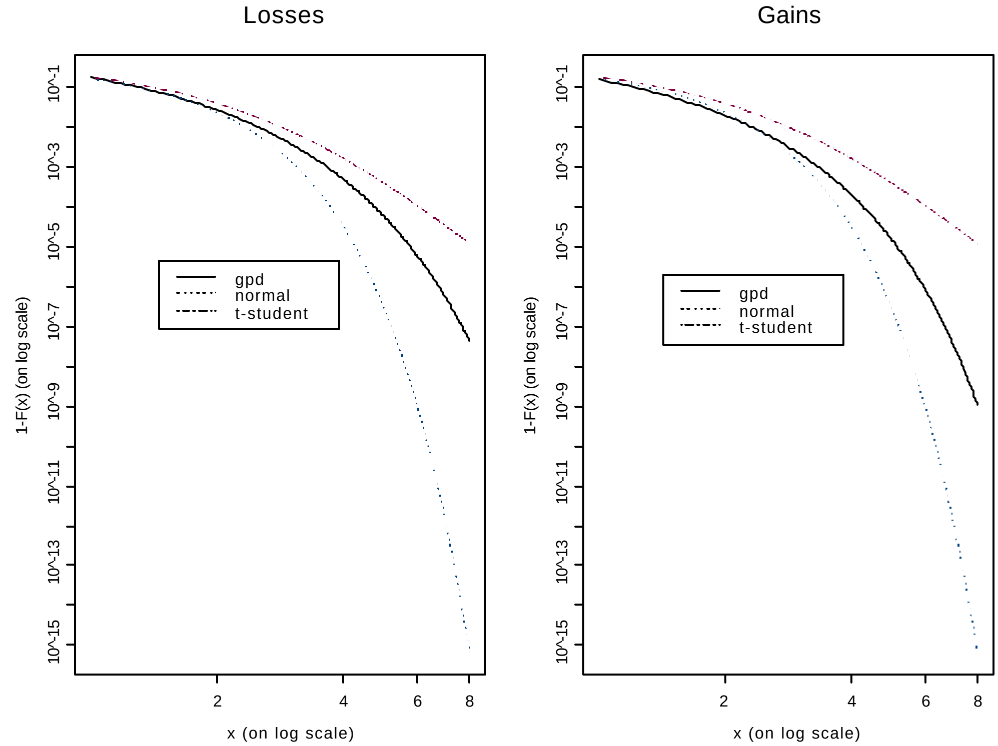
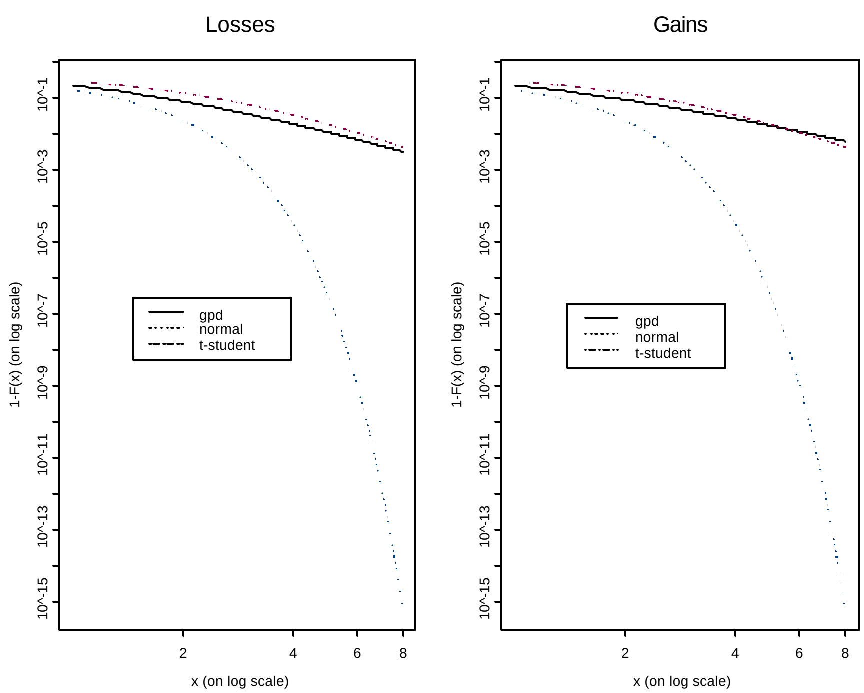


Figure 2 Tail Behavior of Innovations

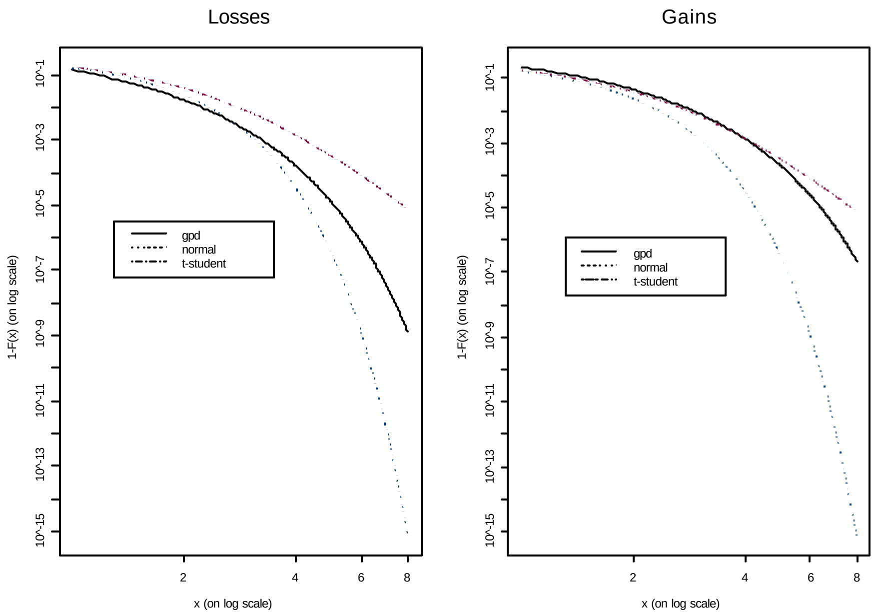
(a) IPSA



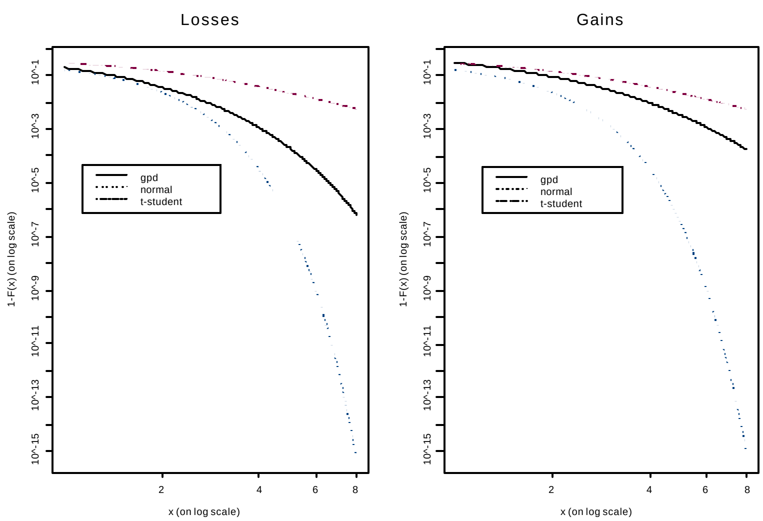
(b) Ch$/US$ Exchange Rate



(c) Copper



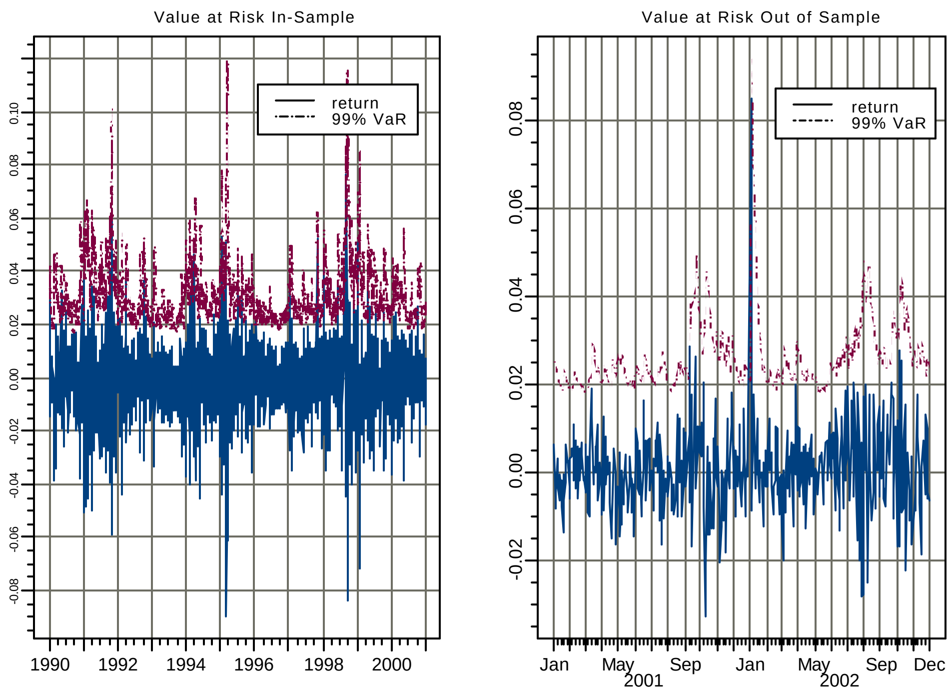
(d) (Proxy) 1-year zero coupon bond innovations



Notes: the degrees of freedom of the t distribution have been determined in each case by fitting a GARCH(1,1) model to the data assuming t innovations. They are the following: 8.9 for IPSA, 3.6 for the CH$/US$ exchange rate, 9.6 for copper, and 3.4 for the proxy of a 1-year zero coupon bond.

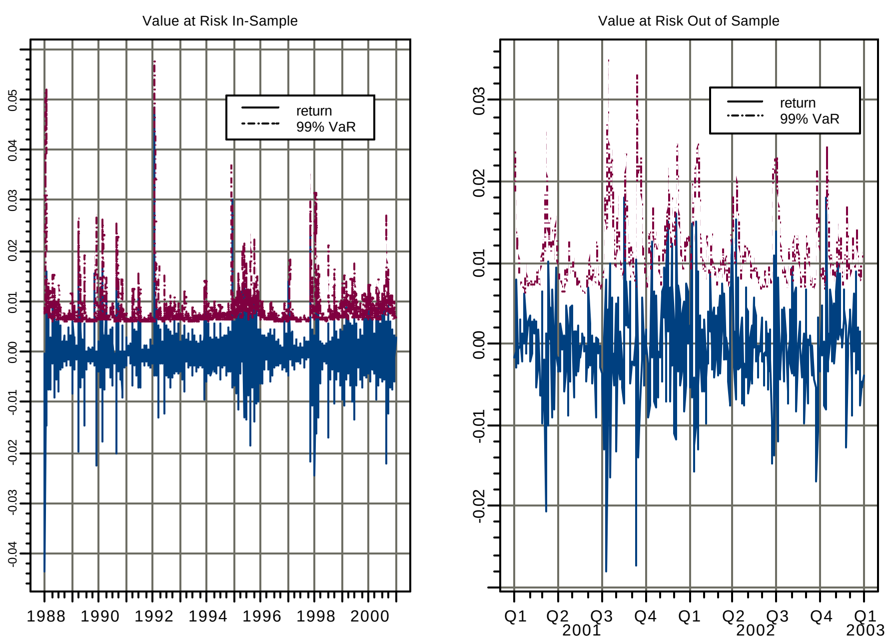
Figure 3 Conditional 99% VaR and Log-Negative Returns In-Sample and Out of Sample

(a) IPSA



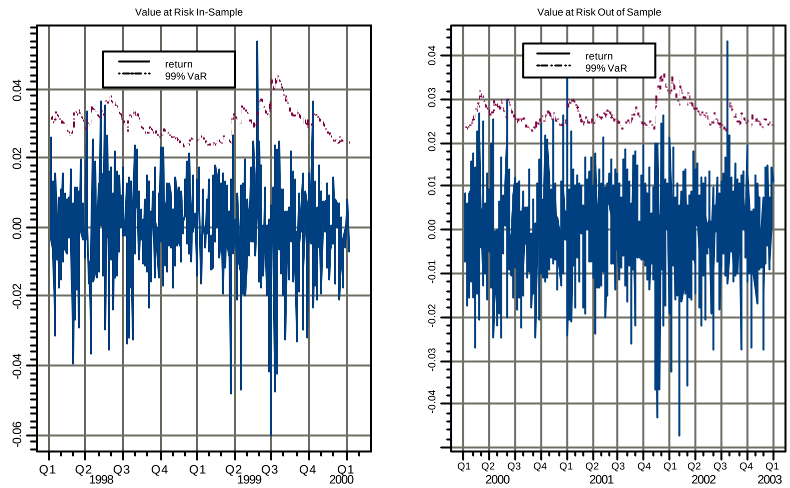
In sample (1990-2000) percent error: 0.77; Out of sample: (2001-2002) percent error=0.42.

(b) Ch$/US$ Exchange Rate:



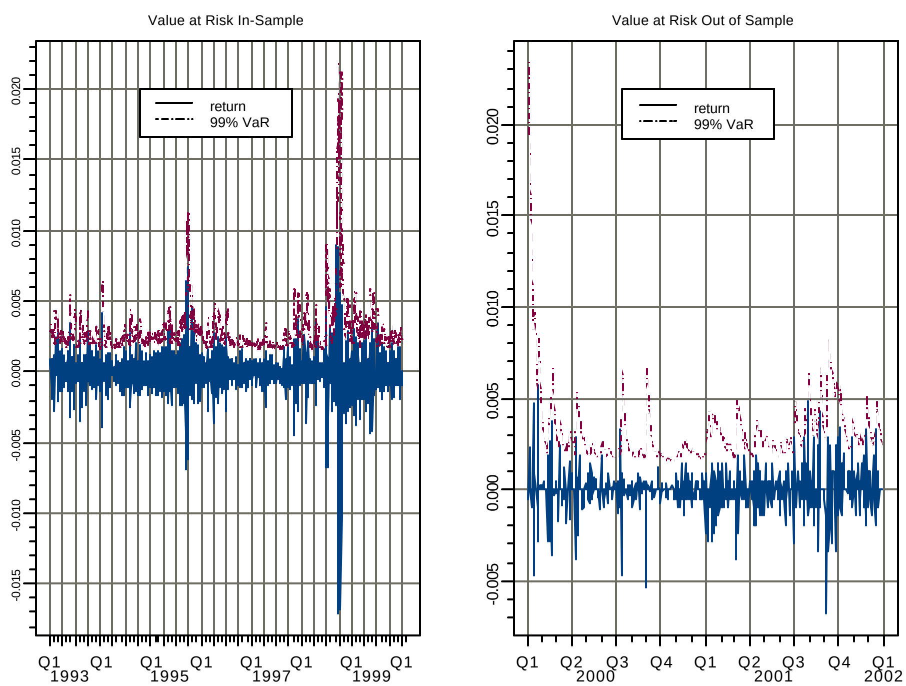
In-sample (1988-2000), percent error: 0.62; Out-of-sample (2001-2002), percent error: 1.61

(c) Copper



In sample (1998-2000) percent error: 0.99; Out-of-sample (2001-2002): 0.66%

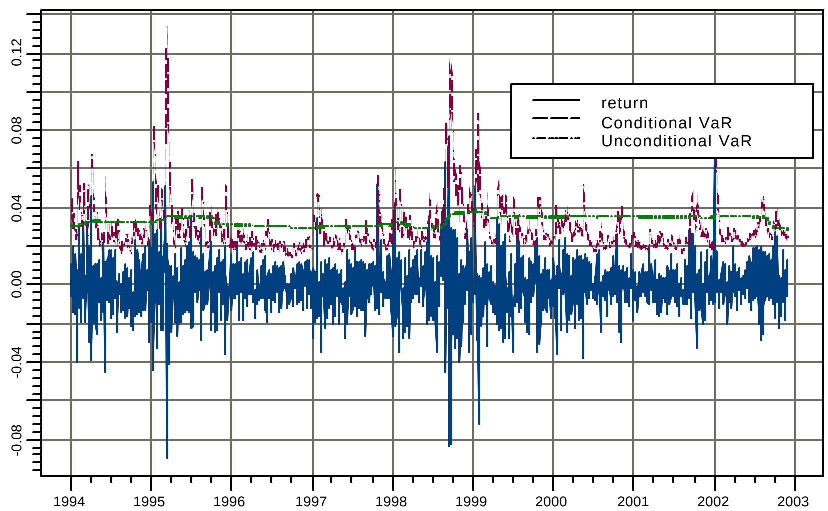
(d) (Proxy) 1-year zero-coupon bond



In sample (1993-1999) percent error: 0.88; Out-of-sample (2000-2001) percent error: 1.4

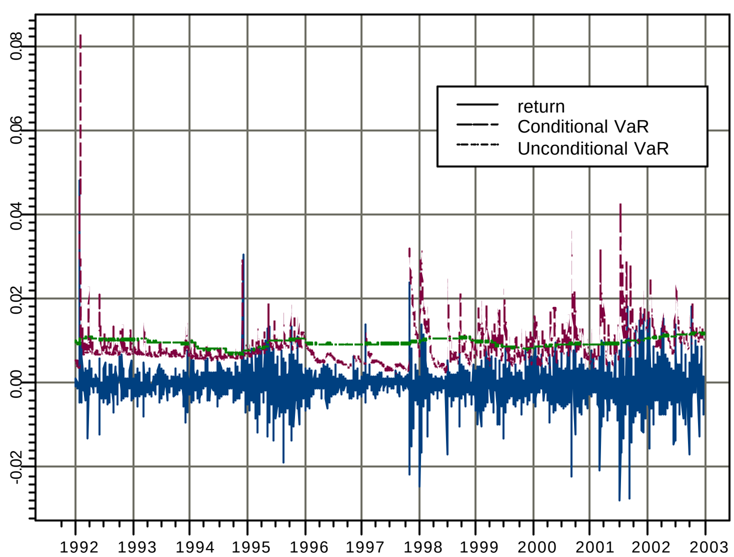
Figure 4 Backtesting: Conditional and Unconditional 99% VaR according to EVT approach

(a) IPSA: 1994-2002



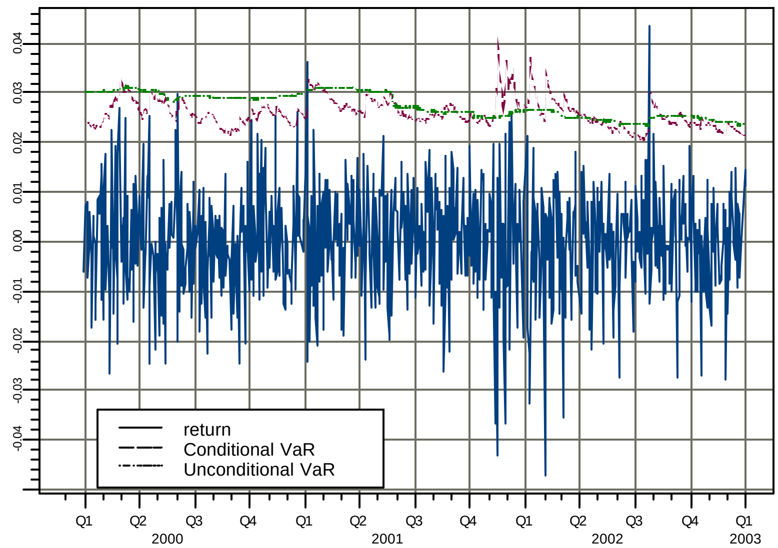
Conditional EVT percent error: 1.14; Unconditional EVT percent error: 1.05

(b) Ch$/US$ Exchange Rate: 1992-2002



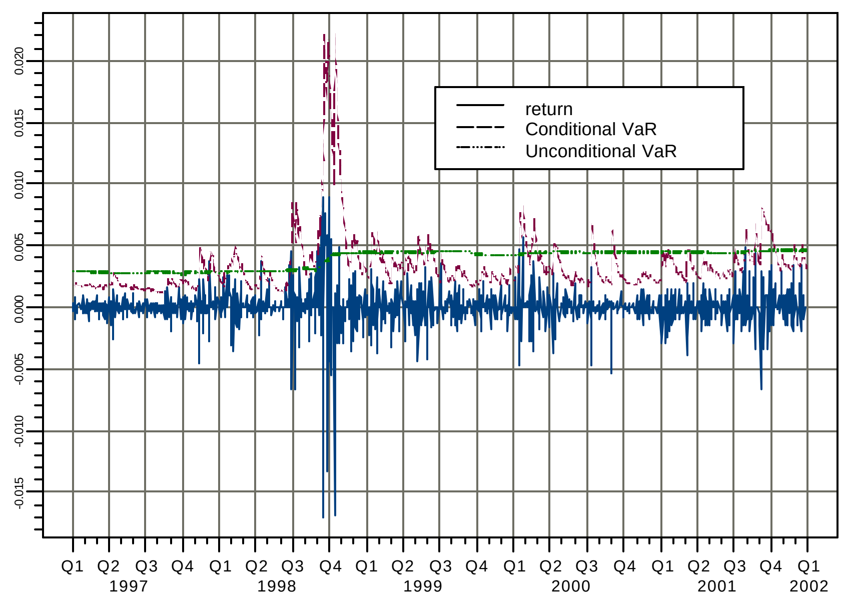
Conditional EVT percent error: 1.17; Unconditional EVT percent error: 1.31

(c) Copper Sample period 2000-2002:



Conditional EVT percent error: 0.8; Unconditional EVT percent error: 0.53

(d) (Proxy) 1-year zero coupon bond: 1997-2002



Conditional EVT percent error: 1.3; Unconditional EVT percent error: 1.62