FinAnalytica Foundation Paper

STABLE ETL OPTIMAL PORTFOLIOS & EXTREME RISK MANAGEMENT

Svetlozar (Zari) Rachev[1]
R. Douglas Martin[2]
Borjana Racheva[3]
Stoyan Stoyanov[4]

Original Publication - Wilmott Magazine of Finance, November 2003
Updated 2006, 2009

FinAnalytica

[1] Chief Scientist, FinAnalytica and Chair Professor of Statistics, Econometrics and Mathematical Finance , School of Economics and Business Engineering, University of Karlsruhe
[2] Founder, FinAnalytica and Professor of Statistics, University of Washington
[3] President, Managing Director of Research and Development, FinAnalytica
[4] Chief Financial Researcher, FinAnalytica

Abstract

We introduce a practical alternative to Gaussian risk factor distributions based on Svetlozar Rachev's work on *Stable Paretian Models in Finance* (see Rachev and Mittnik, 2000) and called the Stable Distribution Framework. In contrast to normal distributions, stable distributions capture the fat tails and the asymmetries of real-world risk factor distributions. In addition, we make use of copulas, a generalization of overly restrictive linear correlation models, to account for the dependencies between risk factors during extreme events, and multivariate ARCH-type processes with stable innovations to account for joint volatility clustering. We demonstrate that the application of these techniques results in more accurate modeling of extreme risk event probabilities, and consequently delivers more accurate risk measures for both trading and risk management. Using these superior models, VaR becomes a much more accurate measure of downside risk. More importantly Stable Expected Tail Loss (SETL) can be accurately calculated and used as a more informative risk measure for both market and credit portfolios. Along with being a superior risk measure, SETL enables an elegant approach to portfolio optimization via convex optimization that can be solved using standard scalable linear programming software. We show that SETL portfolio optimization yields superior risk adjusted returns relative to Markowitz portfolios. Finally, we introduce an alternative investment performance measurement tools: the Stable Tail Adjusted Return Ratio (STARR), which is a generalization of the Sharpe ratio in the Stable Distribution Framework.

1 Extreme Asset Returns Demands New Solutions

Professor Paul Wilmott (www.wilmott.com) likes to recount the ritual by which he questions his undergraduate students on the likelihood of Black Monday 1987. Under the commonly accepted Gaussian risk factor distribution assumption, they consistently reply that there should be no such event in the entire existence of the universe and beyond!

The last two decades have witnessed a considerable increase in fat-tailed kurtosis and skewness of asset returns at all levels, individual assets, portfolios and market indices. - *Extreme* events are the corollary of the increased kurtosis. Legacy risk and portfolio management systems have done a reasonable job at managing *ordinary* financial events. However up to now, very few institutions or vendors have demonstrated the systematic ability to deal with the unusual or extreme event, the one that should *almost never* happen using conventional modeling approaches. Therefore, one can reasonably question the soundness of some of the current risk management practices and tools used in Wall Street as far as extreme risk is concerned.

The two main conventional approaches to modeling asset returns are based either on a historical or a normal (Gaussian) distribution for returns. Neither approach adequately captures unusual asset price and return behaviors. The *historical* model is bounded by the extent of the available observations and the *normal* model inherently cannot produce atypical returns. The financial industry is beleaguered with both under-optimized portfolios with often-poor ex-post risk-adjusted returns, as well as overly optimistic aggregate risk indicators (e.g. VaR) that lead to substantial unexpected losses.

The inadequacy of the normal distribution is well recognized by the risk management community. Yet up to now, no consistent and comprehensive alternative has adequately addressed unusual returns. To quote one major vendor:

> ``It has often been argued that the true distributions returns (even after standardizing by the volatility) imply a larger probability of extreme returns than that implied from the normal distribution. Although we could try to specify a distribution that fits returns better, it would be a daunting task, especially if we consider that the new distribution would have to provide a good fit across all asset classes.“ (Technical Manual, RMG, 2001)

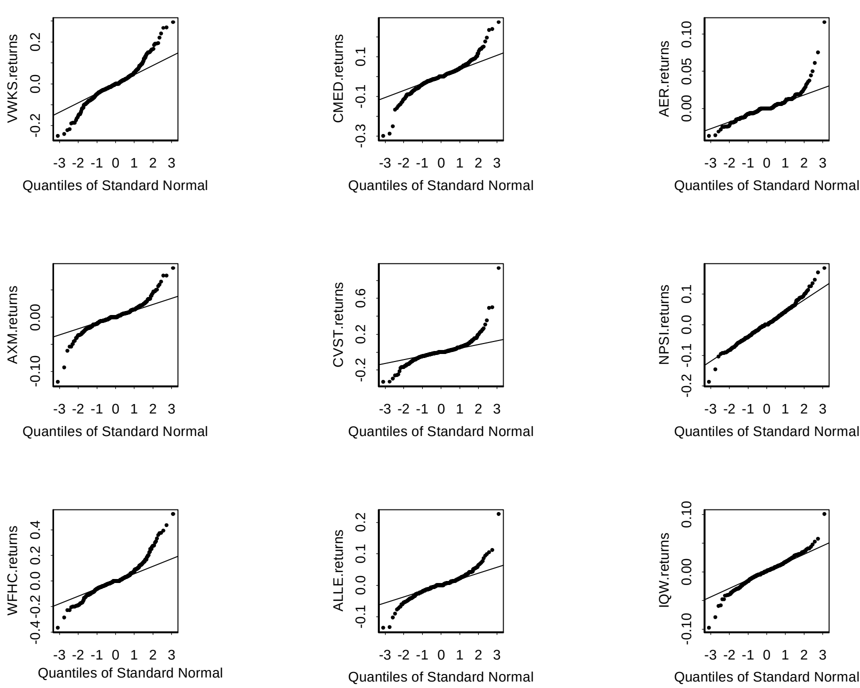
In response to the challenge, we use generalized multivariate stable (GMstable) distributions and generalized risk-factor dependencies, thereby creating a paradigm shift to consistent and uniform use of the most viable class of non-normal probability models in finance. This approach leads to distinctly improved financial risk management and portfolio optimization solutions for assets with extreme events.

2 The Stable Distribution Framework

2.1 Stable Distributions

In spite of wide-spread awareness that most risk factor distributions are heavy-tailed, to date, risk management systems have essentially relied either on historical, or on univariate and multivariate normal (or Gaussian) distributions for Monte Carlo scenario generation. Unfortunately, historical scenarios only capture conditions actually observed in the past, and in effect use empirical probabilities that are zero outside the range of the observed data, a clearly undesirable feature. On the other hand Gaussian Monte Carlo scenarios have probability densities that converge to zero too quickly (exponentially fast) to accurately model real-world risk factor distributions that generate extreme losses. When such large returns occur separately from the bulk of the data they are often called outliers.

The figure below shows quantile-quantile (qq)-plots of daily returns versus the best-fit normal distribution of nine randomly selected microcap stocks for the two-year period 2000-2001. If the returns were normally distributed, the quantile points in the qq-plots would all fall close to a straight line. Instead they all deviate significantly from a straight line (particularly in the tails), reflecting a higher probability of occurrence of extreme values than predicted by the normal distribution, and showing several outliers.



Such behavior occurs in many asset and risk factor classes, including well-known indices such as the S&P 500, and corporate bond prices. The latter are well known to have quite non-Gaussian distributions that have substantial negative skews to reflect down-grading and default events. For such returns, non-normal distribution models are required to accurately model the tail behavior and compute probabilities of extreme returns.

Various non-normal distributions have been proposed for modeling extreme events, including:

- Mixtures of two or more normal distributions
- t-distributions, hyperbolic distributions, and other scale mixtures of normal distributions
- Gamma distributions
- Extreme Value distributions,
- Stable non-Gaussian distributions (also known as Lévy-stable and Pareto-stable distributions)

Among the above, only stable distributions have attractive enough mathematical properties to be a viable alternative to normal distributions in trading, optimization and risk management systems. A major drawback of all alternative models is their lack of stability. Benoit Mandelbrot (1963) demonstrated that the stability property is highly desirable for asset returns. These advantages are particularly evident in the context of portfolio analysis and risk management.

An attractive feature of stable models, not shared by other distribution models, is that they allow generation of Gaussian-based financial theories and, thus allow construction of a coherent and general framework for financial modeling. These generalizations are possible only because of specific probabilistic properties that are unique to (Gaussian and non-Gaussian) stable laws, namely: the stability property, the central limit theorem, and the invariance principle for stable processes.

Benoit Mandelbrot (1963), then Eugene Fama (1965), provided seminal evidence that stable distributions are good models for capturing the heavy-tailed (leptokurtic) returns of securities. Many follow-on studies came to the same conclusion, and the overall stable distributions theory for finance is provided in the definitive work of Rachev and Mittnik (2000).

But in spite the convincing evidence, stable distributions have seen virtually no use in capital markets. There have been several barriers to the application of stable models, both conceptual and technical:

- Except for three special cases, described below, stable distributions have no closed form expressions for their probability densities.

- Except for normal distributions, which are a limiting case of stable distributions (with α=2 and β=0), stable distributions have infinite variance and only a mean value for $\alpha > 1$.

- Without a general expression for stable probability densities, one cannot directly implement maximum likelihood methods for fitting these densities, even in the case of a single (univariate) set of returns.

The availability of practical techniques for fitting univariate and multivariate stable distributions to asset and risk factor returns has been *the* barrier to the progress of stable distributions in finance. Only the recent development of advanced numerical methods has removed this obstacle. These patent-protected methods are at the foundation of the ***Cognity™*** risk management and portfolio optimization software system (see further comments in section 4.7).

Univariate Stable Distributions

A stable distribution for a random risk factor *X* is defined by its characteristic function:

$$F(t) = E\left(e^{itX}\right) = \int e^{itx} f_{\mu,\sigma}(x)dx,$$

where

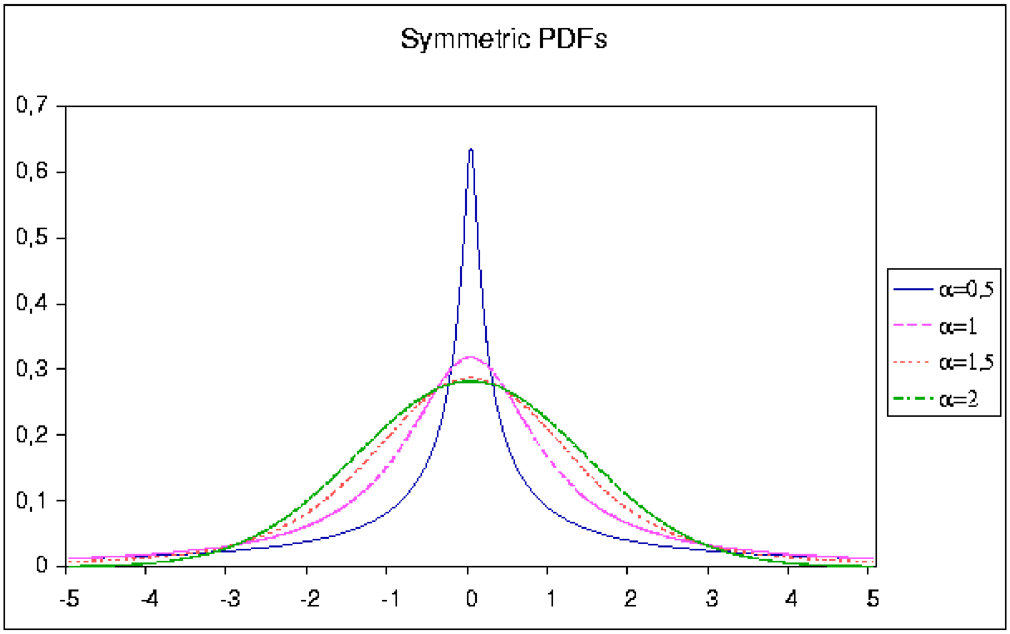
$$f_{\mu,\sigma}(x) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$$

is any probability density function in a location-scale family for *X*:

$$\log F(t) = \begin{cases} -\sigma^{\alpha} |t|^{\alpha} \left(1 - i\beta \, \mathrm{sgn}(t) \tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu t, & \alpha \neq 1 \\ -\sigma |t| \left(1 - i\beta \frac{2}{\pi} \mathrm{sgn}(t) \log|t|\right) + i\mu t, & \alpha = 1 \end{cases}$$

A stable distribution is therefore determined by the four key parameters:

1. α determines density's kurtosis with $0 < \alpha \leq 2$ (e.g. tail weight)
2. β determines density's skewness with $-1 \leq \beta \leq 1$
3. σ is a scale parameter (in the Gaussian case, α =2 and $2\sigma^2$ is the variance)
4. μ is a location parameter (μ is the mean if $1 < \alpha \leq 2$)



Stable distributions for risk factors allow for skewed distributions when $\beta \neq 0$ and fat tails relative to the Gaussian distribution when $\alpha < 2$. The graph above shows the effect of α on tail thickness of the density as well as peakedness at the origin relative to the normal distribution (collectively the "kurtosis" of the density), for the case of $\beta = 0$, $\mu = 0$, and $\sigma = 1$. As the values of α decrease the distribution exhibits fatter tails and more peakedness at the origin.

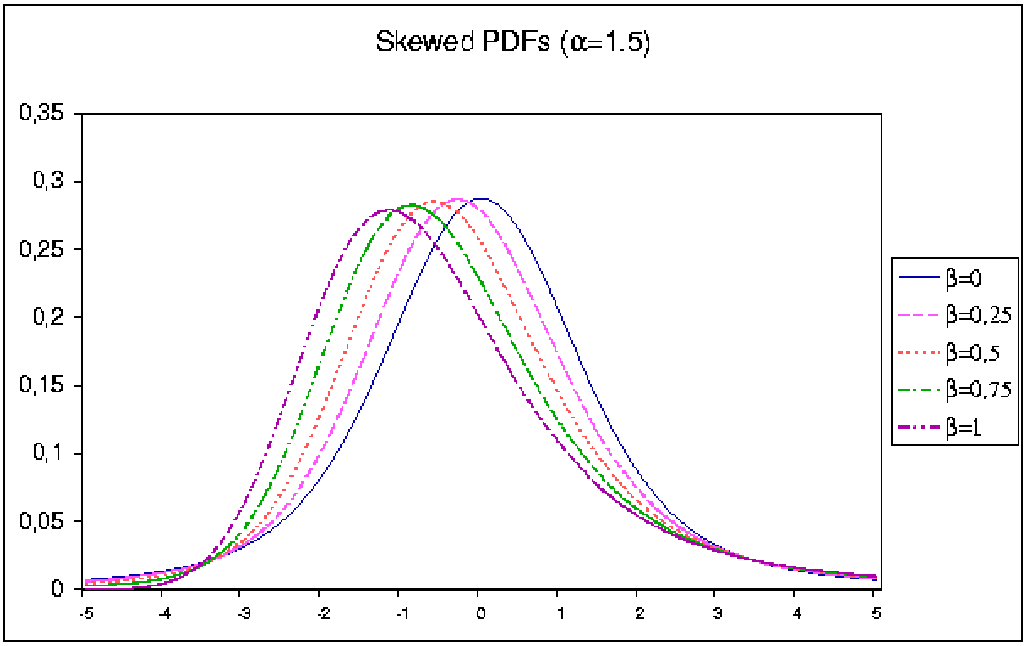
The case of $\alpha = 2$ and $\beta = 0$ and with the reparameterization in scale, $\tilde{\sigma} = \sqrt{2}\sigma$, yields the Gaussian distribution, whose density is given by:

$$f_{\mu,\tilde{\sigma}}(x) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}} e^{-\frac{(x-\mu)^2}{2\tilde{\sigma}^2}} .$$

The case $\alpha=1$ and $\beta=0$ yields the Cauchy distribution with much fatter tails than the Gaussian, and is given by:

$$f_{\mu,\sigma}(x) = \frac{1}{\pi \cdot \sigma}\left(1 + \left(\frac{x-\mu}{\sigma}\right)^2\right)^{-1}$$

The figure below illustrates the influence of β on the skewness of the density for α=1.5, μ=0 and σ=1. Increasing (decreasing) values of β result in skewness to the right (left).



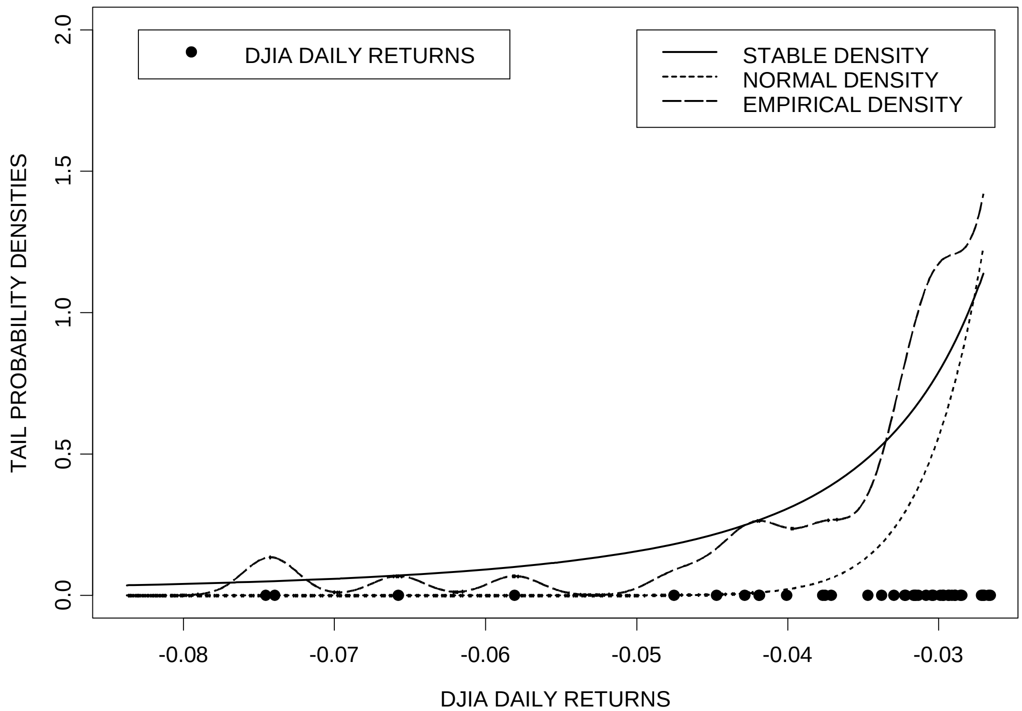
Fitting Stable and Normal Distributions: DJIA Example

Aside from the Gaussian, Cauchy, and one other special case of stable distribution for a positive random variable with $\alpha = 0.5$, there is no closed form expression for the probability density of a stable random variable.

Thus one is not able to directly estimate the parameters of a stable distribution by the method of maximum likelihood. To estimate the four parameters of the stable laws, the ***Cognity™*** system uses a special patent-pending version of the FFT (Fast Fourier Transform) approach to numerically calculate the densities with high accuracy, and then applies maximum likelihood estimation (MLE) to estimate the parameters.

The results from applying the ***Cognity™*** stable distribution modeling to the DJIA daily returns from January 1, 1990 to February 14, 2003 is displayed in the figure below. The figure shows the left-hand tail detail of the resulting stable density, along with that of a normal density fitted using the sample mean and sample standard deviation, and that of a non-parametric kernel density estimate (labeled "Empirical" in the plot legend). The parameter estimates are:

- Stable parameters $\hat{\alpha} = 1.699$, $\hat{\beta} = -.120$, $\hat{\mu} = .0002$, and $\hat{\sigma} = .006$,
- Normal density parameter estimates $\hat{\mu} = .0003$, and $\hat{\sigma} = .010$.



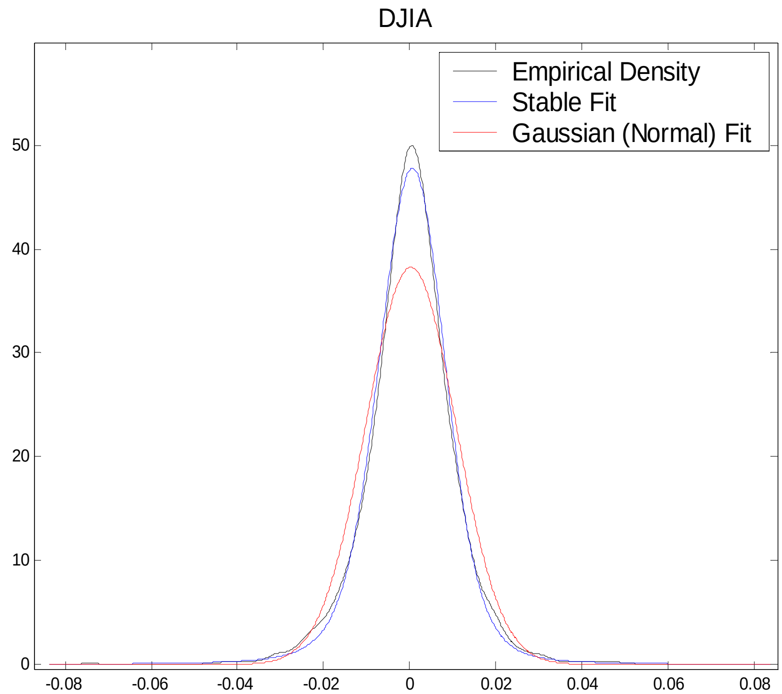
Note that the stable density tail behavior is reasonably consistent with the empirical non-parametric density estimate, indicating the existence of some extreme returns. At the same time it is clear from the figure that the tail of the normal density is much too thin, and will provide inaccurate estimates of tail probabilities for the DJIA returns. The table below shows just how bad the normal tail probabilities are for several negative returns values.

	PROBABILITY (DJIA RETURN < x)			
x	-0.04	-0.05	-0.06	-0.07
Stable Fit	0.0066	0.0043	0.0031	0.0023
Normal Fit	0.000056	0.0000007	3.68E-09	7.86E-12

A daily return smaller than -0.04 with the stable distribution occurs with probability 0.0066, or roughly seven times every four years, whereas such a return with the normal fit occurs on the order of once every four years.

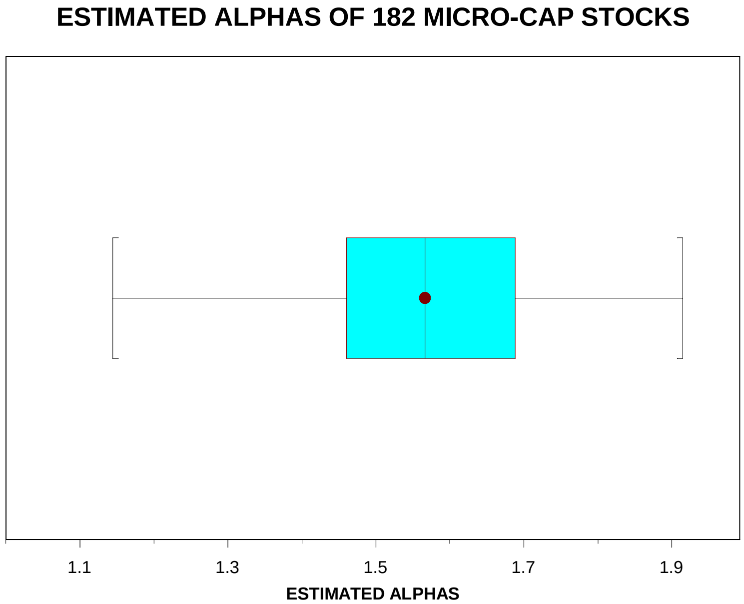
Similarly, a return smaller than -0.05 with the stable occurs about once per year and with the normal fit about once every forty years. Clearly the normal distribution fit is an exceedingly optimistic predictor of DJIA tail return values.

The figure below displays the central portion of the fitted densities as well as the tails, and shows that the normal fit is not nearly peaked enough near the origin as compared with the empirical density estimate (even though the GARCH model was applied), while the stable distribution matches the empirical estimate quite well in the center as well as in the tails.



Fitting Stable Distributions: Micro-Caps Example

Noting that micro-cap stock returns are consistently strongly non-normal (see sample of normal qq-plots at the beginning of this section), we fit stable distributions to a random sample of 182 micro-cap daily returns for the two-year period 2000 – 2001. The results are displayed in the box plot below.



The median of the estimated alphas is 1.57, and the upper and lower quartiles are 1.69 and 1.46 respectively. Somewhat surprisingly, the distribution of the estimated alphas turns out to be quite normal.

Generalized Multivariate Stable Distribution Modeling

Generalized stable distribution GMstable) modeling is based on fitting univariate stable distributions for each one dimensional set of returns or risk factors, each with its own parameter estimates α_i, β_i, μ_i, σ_i, $i=1,2,\ldots,K$, where K is the number of risk factors, along with a dependency structure.

One way to produce the cross-sectional dependency structure is through a scale mixing process (called a "subordinated" process in the mathematical finance literature) as follows. First compute a robust mean vector and covariance matrix estimate of the risk factors to get rid of the outliers, and have a good covariance matrix estimate for the central bulk of the data. Next we generate multivariate normal scenarios with this mean vector and covariance matrix. Then we multiply each of random variable component of the scenarios by a strictly positive stable random variable with index $\alpha_i/2$, $i=1,2,\ldots,K$.

The vector of stable random variable scale multipliers is usually independent of the normal scenario vectors, but it can also be dependent. See for example Rachev S. and Mittnik S. (2000), and Rachev S., Schwartz E. and Khindanova I. (2003).

Another very promising approach to building the cross-sectional dependence model is through the use of copulas, an approach that is quite attractive because it allows for modeling higher correlations during extreme market movements, thereby accurately reflecting lower portfolio diversification at such times. The next section briefly discussion copulas.

2.2 Copula Multivariate Dependence Models

Why Copulas?

Classical correlations and covariances are quite limited measures of dependence, and are only adequate in the case of multivariate Gaussian distributions. A key failure of correlations is that, for non-Gaussian distributions, zero correlation does not imply independence, a phenomenon that arises in the context of time-varying volatilities represented by ARCH and GARH models. The reason we use copulas is that we need more general models of dependence, ones which:

- Are not tied to the elliptical character of the multivariate normal distribution
- Have multivariate contours and corresponding data behavior that reflect the local variation in dependence that is related to the level of returns, in particular, those shapes that correspond to higher correlations with extreme co-movements in returns than with small to modest co-movements.

What are Copulas?

A copula may be defined as a multivariate cumulative distribution function with uniform marginal distributions:

$$C(u_1, u_2, \cdots, u_n), \qquad u_i \in [0,1] \text{ for } i = 1, 2, \cdots, n$$

where

$$C(u_i) = u_i \quad \text{for } i = 1, 2, \cdots, n\,.$$

It is known that for any multivariate cumulative distribution function:

$$F(x_1, x_2, \cdots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \cdots X_n \le x_n)$$

there exists a copula C such that

$$F(x_1, x_2, \cdots, x_n) = C(F_1(x_1), F_2(x_2), \cdots, F_n(x_n))$$

where the $F_i\left(x_i\right)$ are the marginal distributions of $F(x_1, x_2, \cdots, x_n)$, and conversely for any copula C the right-hand-side of the above equation defines a multivariate distribution function $F(x_1, x_2, \cdots, x_n)$. See for example, Bradley and Taqqu (2001) and Sklar (1996).

The main idea behind the use of copulas is that one can first specify the marginal distributions in whatever way makes sense, e.g. fitting marginal distribution models to risk factor data, and then specify a copula C to capture the multivariate dependency structure in the best suited manner.

There are many classes of copula, particularly for the special case of bivariate distributions. For more than two risk factors beside the traditional Gaussian copula, the t-copula is very tractable for implementation and provides a possibility to model dependencies of extreme events. It is defined as:

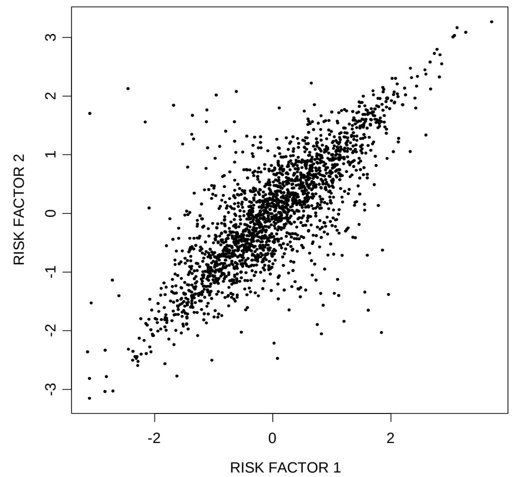
$$C_{\nu,\mathbf{c}}(u_1, u_2, \cdots, u_n) = \frac{\Gamma\left((\nu + n)/2\right)}{\Gamma\left(\nu/2\right)\sqrt{|\mathbf{c}|(\nu\pi)^n}} \int_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(u_n)} \left(1 + \frac{\mathbf{s}'\mathbf{c}^{-1}\mathbf{s}}{\nu}\right) d\mathbf{s}$$

where $\mathbf{c}$ is a correlation matrix.

A sample of 2000 bivariate simulated risk factors generated by a t-copula with 1.5 degrees of freedom and normal marginal distributions is displayed in the figure below.

The example illustrates that these two risk factors are somewhat uncorrelated for small to moderately large returns, but are highly correlated for the infrequent occurrence of very large returns. This can be seen by noting that the density contours of points in the scatter plot are somewhat eliptical near the origin, but are nowhere close to eliptical for more extreme events. This situation is in contrast to a Gaussian linear dependency relationship where the density contours are expected to be eliptical.

T-COPULA WITH 1.5 DOF AND NORMAL MARGINALS



2.3 Volatility Clustering Models and Stable VaR

It is well known that asset returns and risk factors returns exhibit volatility clustering, and that even after adjusting for such clustering the returns will still be non-normal and contain extreme values. There may also be some serial dependency effects to account for. In order to adequately model these collective behaviors we recommend using ARIMA models with an ARCH/GARCH "time-varying" volatility input, where the latter has non-normal stable innovations. This approach is more flexible and accurate than the commonly used simple exponentially weighted moving average (EWMA) volatility model, and provides accurate time-varying estimates of VaR and expected tail loss (ETL) risk measures. See section 3 for discussion of ETL vs. VaR that emphasizes the advantages of ETL. However, we stress that those who must use VaR to satisfy regulatory requirements will get much more accurate results with stable VaR than with normal VaR, as the following example vividly shows.

Consider the following portfolio of Brady bonds:

- Brazil C 04/14
- Brazil EIB 04/06
- Venezuela DCB Floater 12/07
- Samsung KRW Ord Shares
- Thai Farmers Bank THB

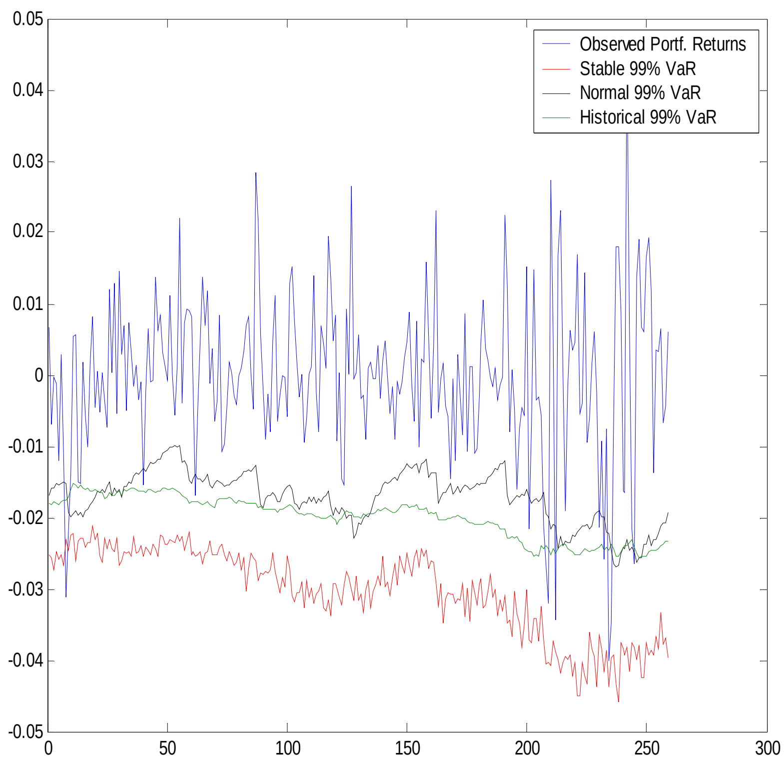
We have run normal, historical and stable 99% (1% tail probability) VaR calculations for one-year of daily data from January 9, 2001 to January 9, 2002. We used a moving window with 250 historical observations for the normal VaR model, 500 for the historical VaR model and 700 for the stable VaR model. For each of these cases we used a GARCH(1,1) model for volatility clustering of the risk factors, with stable innovations.

We back-tested these VaR calculations by using the VaR values as one-step ahead predictors, and got the results shown in the figure below.

The figure shows: the returns of the Brady bond portfolio (top curve); the normal+EWMA (a la RiskMetrics) VaR (curve with jumpy behavior, just below the returns); the historical VaR (the smoother curve mostly below but sometimes crossing the normal+EWMA VaR); the stable+GARCH VaR (the bottom curve). The results with regard to exceedances of the 99% VaR, and keeping in mind Basel II guidelines, may be summarized as follows:

- Normal 99% VaR produced 12 exceedances (red zone)
- Historical 99% VaR produced 9 exceedances (on upper edge of yellow zone)
- Stable 99% VaR produced 1 exceedence and nearly two (well in the green zone)

Clearly stable (+GARCH) 99% VaR produces much better results with regard to Basel II compliance. This comes at the price of higher initial capital reserves, but results in a much safer level of capital reserves and a very clean bill of health with regard to compliance. Note that organizations in the red zone will have to increase their capital reserves by 33%, which at some times for some portfolios will result in larger capital reserves than when using the stable VaR, this in addition to being viewed as having inadequate risk measures relative to the organization using stable VaR.



3 ETL is the Next Generation Risk Measure

3.1 Why Not Value-at-Risk (VaR)?

There is no doubt that VaR's popularity is in large part due to its simplicity and its ease of calculation for 1 to 5% confidence levels. However, there is a price to be paid for the simplicity of VaR in the form of several limitations:

- VaR does not give any indication of the risk beyond the quantile, and so provides very weak information on downside risk.
- VaR portfolio optimization is a non-convex, non-smooth problem with multiple local minima that can result in portfolio composition discontinuities. Furthermore it requires complex calculation techniques such as integer programming.
- VaR is not sub-additive; i.e. the VaR of the aggregated portfolio can be larger than the sum of the VaR's of the sub-portfolios.
- Historical VaR limits the range of the scenarios to data values that have actually been observed, while normal Monte Carlo tends to seriously underestimate the probability of extreme returns. In either case, the probability functions beyond the sample range are either zero or excessively close to zero.

Yamai and Yoshiba (2002) note in their concluding remarks: "The widespread use of VaR for risk management could lead to market instability. Basak and Shapiro (2001) show that when investors use VaR for their risk management, their optimizing behavior may result in market positions that are subject to extreme loss because VaR provides misleading information regarding the distribution's tail."

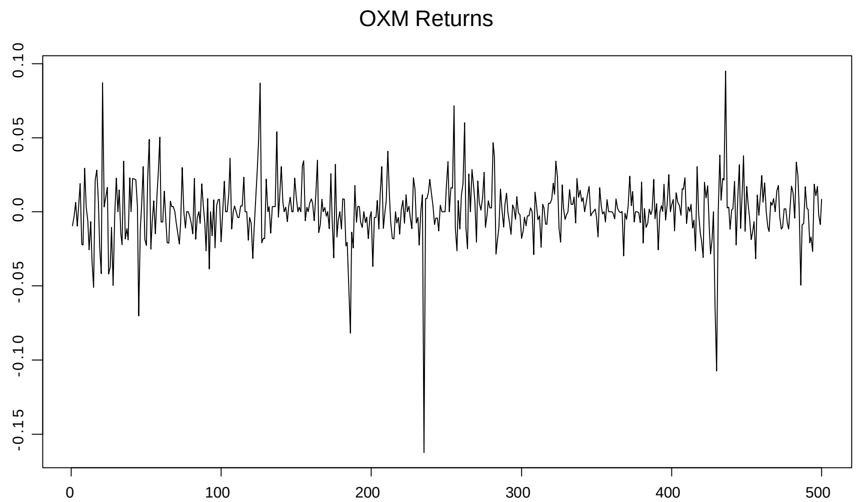
3.2 ETL and Stable versus Normal Distributions

Expected Tail Loss (ETL) is simply the average (or expected value) loss conditioned on the loss being larger than VaR. ETL is also known as *Conditional Value-at-Risk* (CVaR), or *Expected Shortfall* (ES). As such ETL is intuitively much more informative than VaR. We note however that ETL offers little benefit to investors who use a normal distribution to calculate VaR at the usual 99% confidence limit (1% tail probability). The reason is that the resulting VaR and ETL values differ by very little, specifically:

- For CI = 1%, VaR = 2.336 and ETL = 2.667

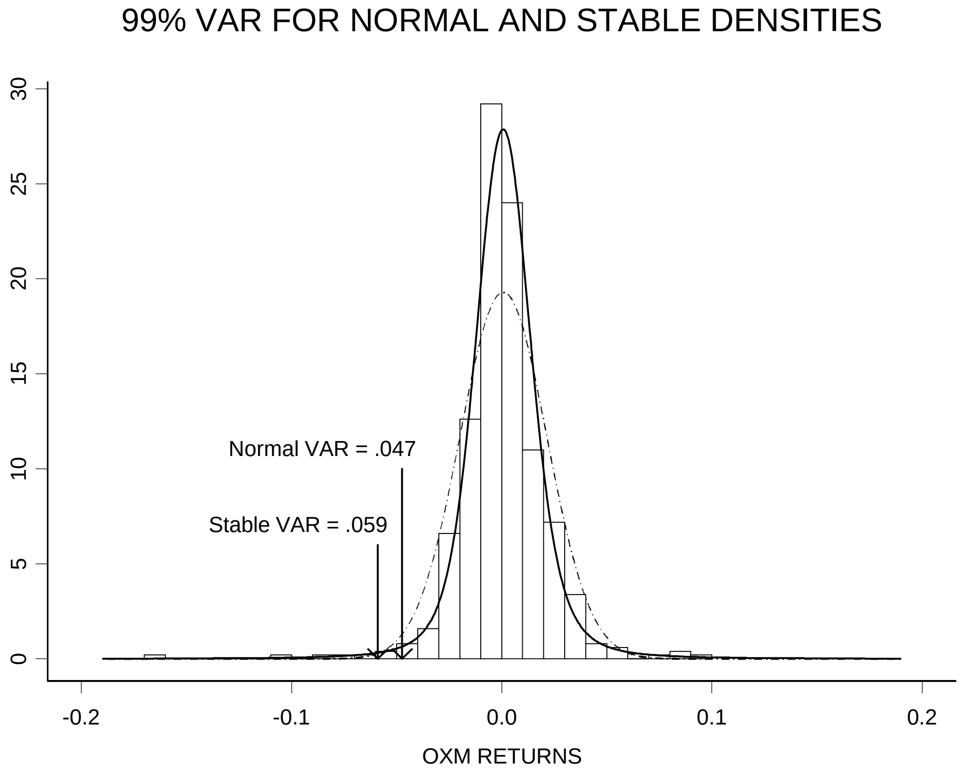
ETL really comes into its own when coupled with stable distribution models that capture leptokurtic tails ("fat tails"). In this case ETL and VaR values will be quite different ,with the resulting ETL often being much larger than the VaR.

As in the graph below, consider the time series of daily returns for the stock OXM from January 2000 to December 2001. Observe the occurrences of extreme values.



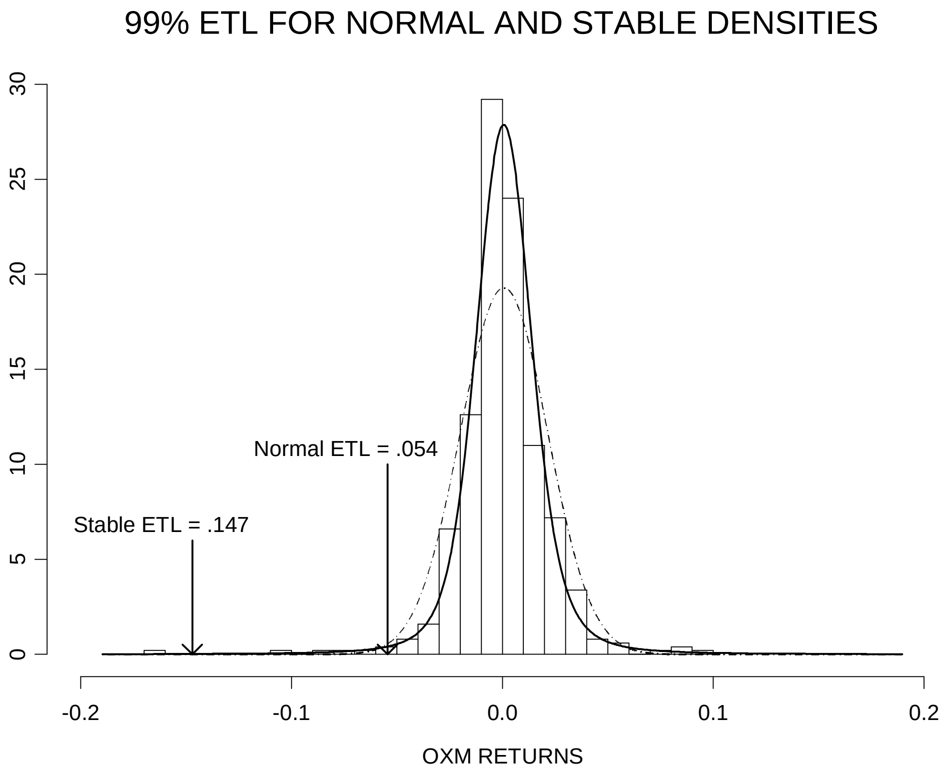
While this series also displays obvious volatility clustering that deserves to be modeled as described in section 3.3, we shall ignore this aspect for the moment. Rather, here we provide a compelling example of the difference between ETL and VaR based on a well-fitting stable distribution, as compared with a poor fitting normal distribution.

The figure below shows a histogram of the OXM returns with a normal density fitted using the sample mean and sample standard deviation, and a stable density fitted using



maximum-likelihood estimates of the stable distribution parameters. The stable density is shown by the solid line and the normal density is shown by the dashed line. The former is obviously a better fit than the latter, when using the histogram of the data values as a reference. The estimated stable tail thickness index is $\hat{\alpha} = 1.62$. The 1% VaR values for the normal and stable fitted densities are .047 and .059 respectively, a ratio of 1.26 which reflects the heavier-tailed nature of the stable fit.

The figure below displays the same histogram and fitted densities with 1% ETL values instead of the 1% VaR values. The 1% ETL values for the normal and stable fitted densities are .054 and .174 respectively, a ratio of a little over three-to-one. This larger ratio is due to the stable density's heavy tail contribution to ETL relative to the normal density fit.



Unlike VaR, ETL has a number of attractive properties:

- ETL gives an informed view of losses beyond VaR.
- ETL is a convex, smooth function of portfolio weights, and is therefore attractive to optimize portfolios (see Uryasev & Rockafellar, 2000). This point is vividly illustrated in the subsection below on ETL and Portfolio Optimization.
- ETL is sub-additive and satisfies a set of intuitively appealing coherent risk measure properties (see Artzner et. al., 1999).
- ETL is a form of expected loss (i.e. a conditional expected loss) and is a very convenient form for use in scenario-based portfolio optimization. It is also quite a natural risk-adjustment to expected return (see STARR, or Stable Tail Adjusted Return Ratio).

The limitations of current normal risk factor models and the absence of regulator blessing have held back the widespread use of ETL, in spite of its highly attractive properties. However, we expect ETL to be a widely accepted risk measure as portfolio and risk managers become more familiar with its attractive properties.

For portfolio optimization, we recommend the use of Stable distribution ETL (SETL), and limiting the use of historical, normal or stable VaR to required regulatory reporting purposes only. Finally, organizations should consider the advantages of Stable ETL for risk assessment purposes and non-regulatory reporting purposes.

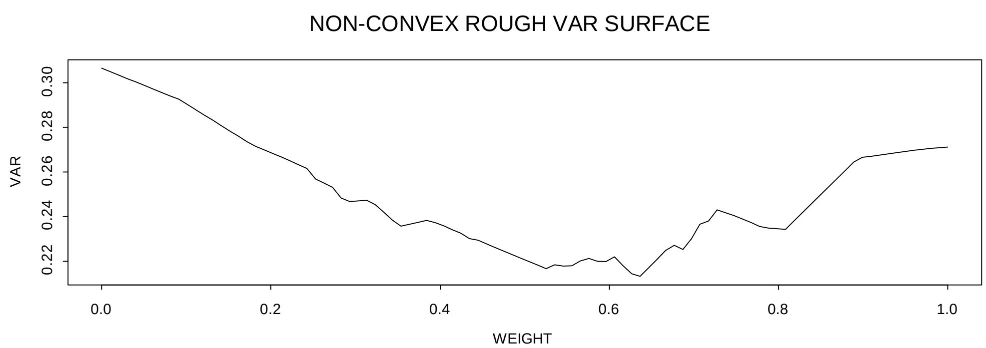
The following quotation is relevant: “Expected Tail Loss gives an indication of extreme losses, should they occur. Although it has not become a standard in the financial industry, expected tail loss is likely to play a major role, as it currently does in the insurance industry” (Embrechts et. al. 1997).

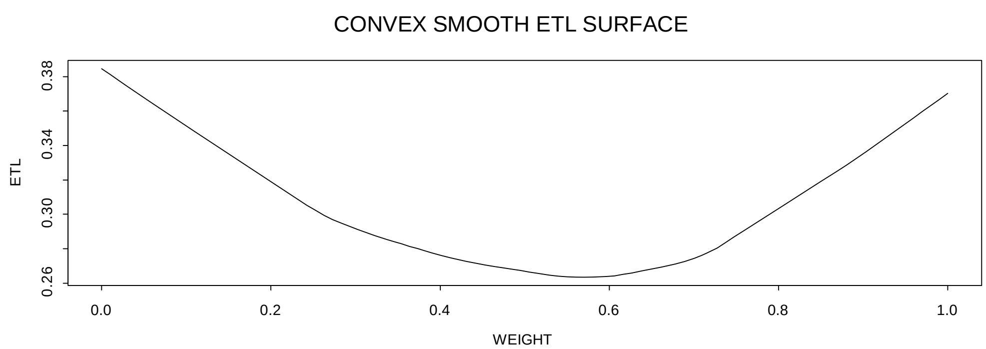
3.3 Portfolio Optimization and ETL versus VaR

To the surprise of many, portfolio optimization with ETL turns out to be a smooth, convex problem with a unique solution (Rockafellar and Uryasev, 2000). These properties are in sharp contrast to the non-convex, rough VaR optimization problem.

The contrast between VAR and ETL portfolio optimization surfaces is illustrated with the following pair of figures for a simple two-asset portfolio. The horizontal axes show one of the portfolio weights (from 0 to 100%) and the vertical axes display portfolio VAR and ETL respectively. The data consist of 200 simulated uncorrelated returns.

The VAR objective function is quite rough with respect to varying the portfolio weight(s), while that of the ETL objective function is smooth and convex. One can see that optimizing with ETL is a much more tractable problem than optimizing with VaR.





Rockafellar and Uryasev (2000), show that the ETL optimal portfolio weight vector can be obtained based on historical (or scenario) returns data by minimizing a relatively simple convex function (Rockafellar and Uryasev used the term CVaR whereas we use the, less confusing, synonym ETL). Assuming p assets with single period returns $\mathbf{r}_i = (r_{i1}, r_{i2}, \cdots, r_{ip})$, for period i, and a portfolio weight vector $\mathbf{w} = (w_1, w_2, \ldots, w_p)$, the function to be minimized is

$$F(\mathbf{w}, \gamma) = \gamma + \frac{1}{\varepsilon \cdot n} \sum_{i=1}^{n} \left[\mathbf{w}'\mathbf{r}_i - \gamma \right]^+ .$$

where $[x]^+$ denotes the positive part of x. This function is to be minimized jointly with respect to $\mathbf{w}$ and γ, where ε is the tail probability for which the expected tail loss is computed. Typically $\varepsilon = .05$ or $.01$, but larger values may be useful, as we discuss in section 4.6. The authors further show that this optimization problem can be cast as a LP (linear programming) problem, solvable using any high-quality LP software.

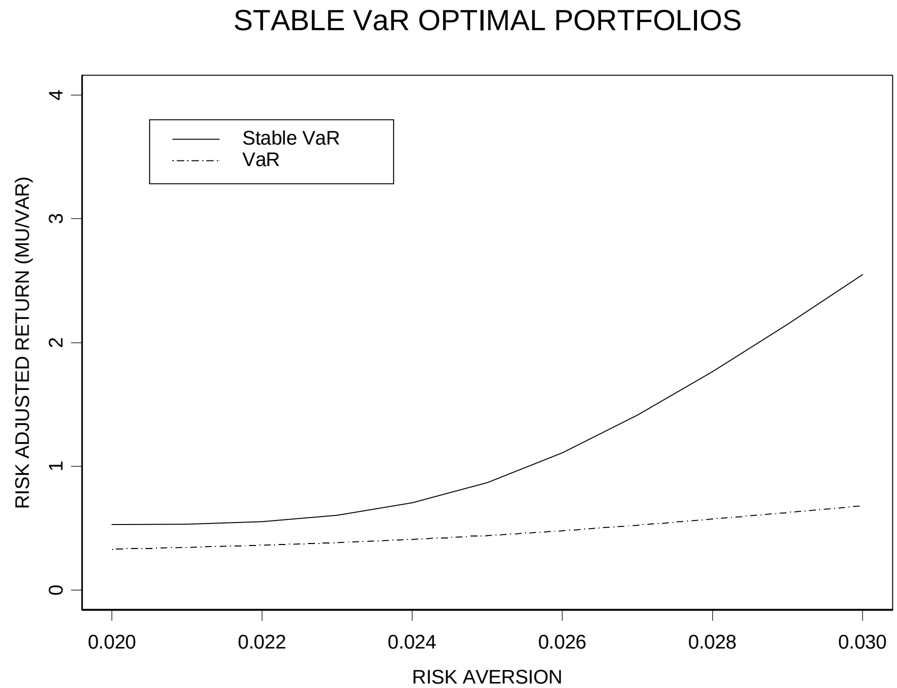
Cognity™ combines this approach with fitting GMstable distribution models for scenario generation. The stable scenarios provide accurate and well-behaved estimates of ETL for the optimization problem.

3.4 Stable ETL Leads to Higher Risk Adjusted Returns

ETL portfolio optimization based on GMstable distribution modeling, which we refer to as SETL portfolios, can lead to significant improvements in risk adjusted return as compared to the conventional Markowitz mean-variance portfolio optimization.

The figures below are supplied to illustrate the claim that stable ETL optimal portfolios produce consistently better risk-adjusted returns. These figures show the risk adjusted return MU/VaR (mean return divided by VaR) and MU/ETL (mean return divided by ETL) for 1% VaR optimal portfolios and ETL optimal portfolios, and using a multi-period fixed-mix optimization in all cases.

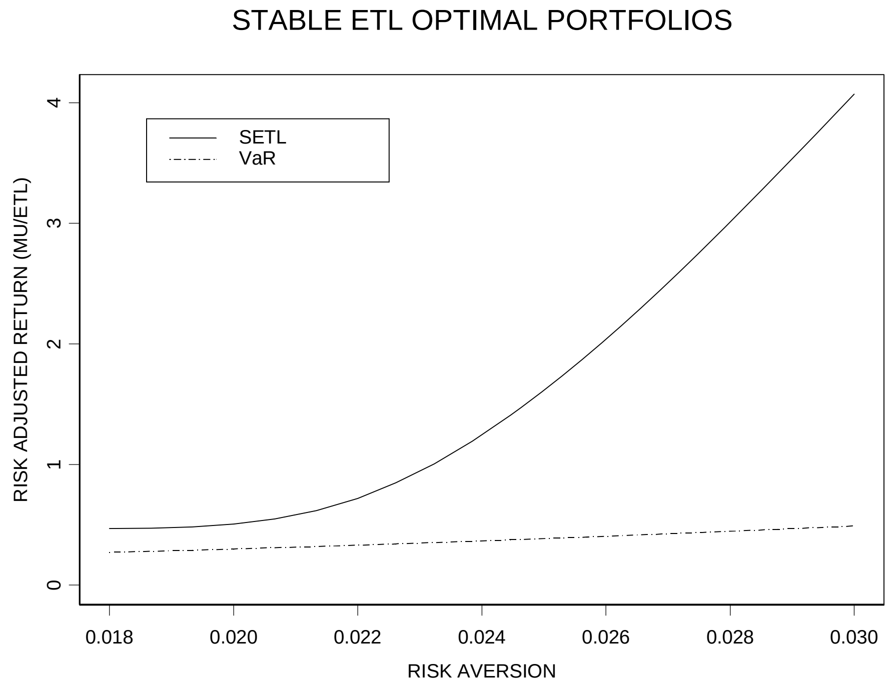
In this simple example, the portfolio to be optimized consists of two assets, cash and the S&P 500. The example is based on monthly data from February 1965 to December 1999. Since we assume full investment, the VaR optimal portfolio depends only on a single portfolio weight and the optimal weight(s) is found by a simple grid search on the interval 0 to 1. The use of a grid search technique, overcomes the problems with non-convex and non-smooth VaR optimization. In this example the optimizer is maximizing $MU - c \cdot VAR$ and $MU - c \cdot ETL$, where c is the risk aversion (parameter), and with VaR or ETL as the penalty function.



The first plot shows that even using the VaR optimal portfolio, one gets a significant relative gain in risk-adjusted return using stable scenarios when compared to normal scenarios, and with the relative gain increasing with increasing risk aversion. The reason for the latter behavior is that with stable distributions the optimization pays more attention to the S&P returns distribution tails, and allocates less investment to the S&P under stable distributions than under normal distributions as risk aversion increases.

The figure below for the risk-adjusted return for the ETL optimal portfolio has the same vertical axis range as the previous plot for the VaR optimal portfolio. The figure below shows that the use of ETL results in much greater gain under the stable distribution relative to the normal than in the case of the VaR optimal portfolio.

At every level of risk aversion, the investment in the S&P 500 is even less in the ETL optimal portfolio than in the case of the VaR optimal portfolio. This behavior is to be expected because the ETL approach pays attention to the losses beyond VaR (the expected value of the extreme loss), and which in the stable case are much greater than in the normal case.



4 The Stable ETL Paradigm

4.1 The Stable ETL Framework

Our risk management and portfolio optimization framework uses multi-dimensional asset and risk factor returns models based on GMstable distributions, and stresses the use of Stable ETL (SETL) as the risk measure of choice. These stable distribution models incorporate generalized dependence structure with copulas, and include time varying volatilities based on GARCH models with stable innovations. Henceforth we use the term *GMstable distribution* to include the generalized dependence structure and volatility clustering model aspects of the model. Collectively, these modeling foundations form the basis of a new and powerful overall basis for investment decisions that we call the *SETL Framework*.

Currently the SETL framework has the following basic components:

- SETL scenario engines
- SETL factor models
- SETL integrated market risk and credit risk
- SETL optimal portfolios and efficient frontiers
- SETL derivative pricing

Going forward, additional classes of SETL investment decision models will be developed, such as SETL betas and SETL asset liability models. The rich structure of these models will encompass the heavy-tailed distributions of the asset returns, stochastic trends, heteroscedasticity, short-and long-range dependence, and more. We use the term "SETL model" to describe any such model in order to keep in mind the importance of the stable tail-thickness parameter α and skewness parameter β, along with volatility clustering and general dependence models, in financial investment decisions.

It is essential to keep in mind the following SETL fundamental principles concerning risk factors:

> P1) Asset and risk factor returns have stable distributions where each asset or risk factor typically has a different stable tail-index α_i and skewness parameter β_i.
>
> P2) Asset and risk factor returns are associated through models that describe the dependence between the individual factors more accurately than classical correlations. Often these will be copula models.
>
> P3) Asset and risk factor modeling typically includes a SETL econometric model in the form of multivariate ARIMA-GARCH processes with residuals driven by fractional stable innovations. The SETL econometric model captures clustering and long-range dependence of the volatility.

4.2 Stable ETL Optimal Portfolios

A SETL optimal portfolio is one that minimizes portfolio expected tail loss subject to a constraint of achieving expected portfolio returns at least as large as an investor defined level, along with other typical constraints on weights, where both quantities are evaluated in the SETL framework. Alternatively, a SETL optimal portfolio solves the dual problem of maximizing portfolio expected return subject to a constraint that portfolio expected tail loss is not greater than an investor defined level, where again both quantities are evaluated in the SETL framework. In order to define the above ETL precisely we use the following quantities:

R_p :	the random return of portfolio *p*
SER_p :	the stable distribution expected return of portfolio *p*
$L_p = -R_p + SER_p$:	the loss of portfolio *p* relative to its expected return
ε :	a tail probability of the SETL distribution L_p
$SVaR_p(\varepsilon)$:	the stable distribution Value-at-Risk for portfolio *p*

The latter is defined by the equation

$$\Pr[L_p > SVaR_p(\varepsilon)] = \varepsilon$$

where the probability is calculated in the SETL framework, that is $SVaR_p(\varepsilon)$ is the ε-quantile of the stable distribution of L_p. In the value-at-risk literature $(1-\varepsilon)\times 100\%$ is called the confidence level. Here we prefer to use the simpler, unambiguous term *tail probability*. Now we define SETL of a portfolio *p* as

$$SETL_p(\varepsilon) = E[L_p \mid L_p > SVaR_p(\varepsilon)]$$

where the conditional expectation is also computed in the SETL framework. We use the "S" in SER_p, $SVaR_p(\varepsilon)$ and $SETL_p(\varepsilon)$ as a reminder that stable distributions are a key aspect of the framework (but not the only aspect!).

Proponents of normal distribution VaR typically use tail probabilities of .01 or .05. When using $SETL_p(\varepsilon)$ risk managers may wish to use other tail probabilities such as .1, .15, .20, .25, or .5. We note that use of different tail probabilities is similar in spirit to using different utility functions.

The following assumptions are in force for the SETL investor:

A1) The universe of assets is Q (the set of mandate admissible portfolios)
A2) The investor may borrow or deposit at the risk-free rate r_f without restriction
A3) The portfolio is optimized under a set of asset allocation constraints λ
A4) The investor seeks an expected return of at least μ

To simplify the notation we shall let A3 be implicit in the following discussion. At times we shall also suppress the ε when its value is taken as fixed and understood.

The SETL investor's optimal portfolio is

$$\omega_\alpha(\mu \mid \varepsilon) = \arg\min_{q\in Q} SETL_q(\varepsilon)$$

subject to

$$SER_q \geq \mu .$$

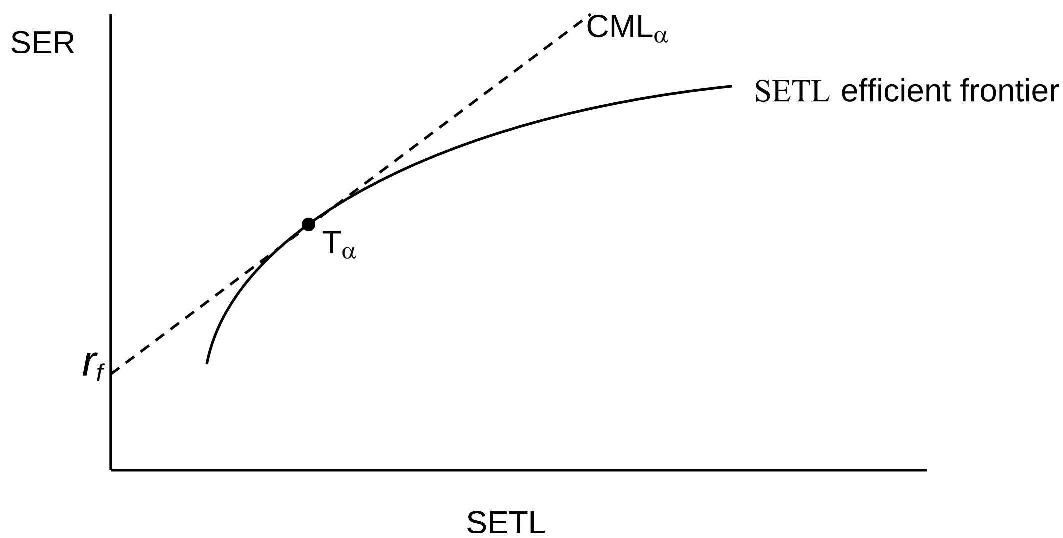
Here we use ω_α to mean either the resulting portfolio weights or the label for the portfolio itself, depending upon the context. The subscript α to remind us that we are using a GMstable distribution modeling approach (which entails different stable distribution parameters for each asset and risk factor). In other words the SETL optimum portfolio ω_α minimizes the expected tail loss among all portfolios with mean return at least μ , for fixed tail probability ε and asset allocation constraints λ. Alternatively, the SETL optimum portfolio ω_α solves the dual problem

$$\omega_\alpha(\eta \mid \varepsilon) = \arg\max_{q\in Q} SER_q$$

subject to

$$SETL_q(\varepsilon) \leq \eta \,.$$

The SETL efficient frontier is given by $\omega_\alpha(\mu \mid \varepsilon)$ as a function of μ for fixed ε, as indicated in the figure below. If the portfolio includes cash account with risk free rate



r_f, then the SETL efficient frontier will be the SETL capital market line (CML_α) that connects the risk-free rate on the vertical axis with the SETL tangency portfolio (T_α), as indicated in the figure.

We now have a SETL separation principal analogous to the classical separation principal: The tangency portfolio T_α can be computed without reference to the risk-return preferences of any investor. Then an investor chooses a portfolio along the SETL capital market line CML_α according to his/her risk-return preference.

Keep in mind that in practice when a finite sample of returns one ends up with a SETL efficient frontier, tangency portfolio and capital market line that are estimates of true values for these quantities.

4.3 Markowitz Portfolios are Sub-Optimal

While the SETL investor has optimal portfolios described above, the Markowitz investor is not aware of the SETL framework and constructs a mean-variance optimal portfolio. We assume that the Markowitz investor operates under the same assumptions A1-A4 as the SETL investor. Let ER_q be the expected return and σ_q the standard deviation of the returns of a portfolio *q*. The Markowitz investor's optimal portfolio is

$$\omega_2(\mu) = \min_{q \in Q} \sigma_q$$

subject to

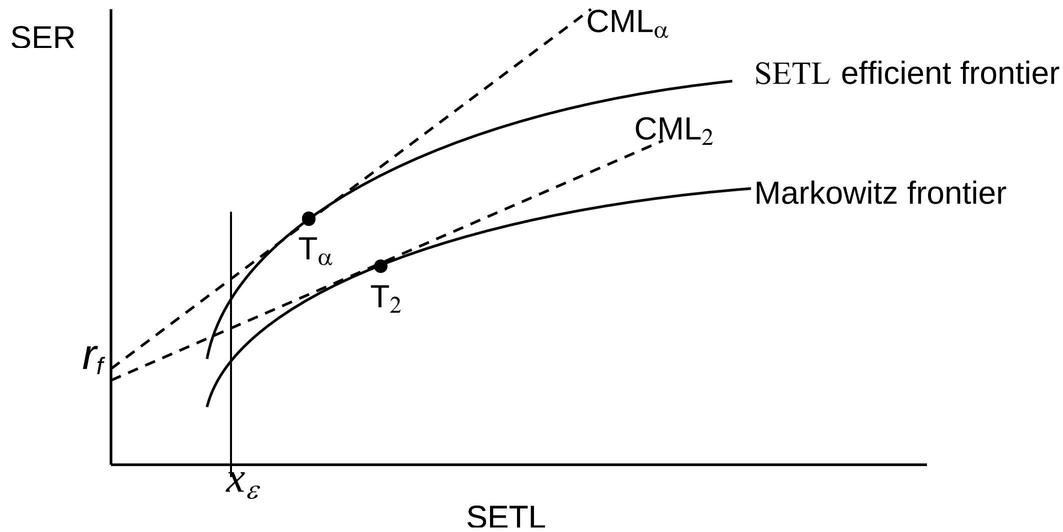
$$ER_q \geq \mu$$

along with the other constraints λ.

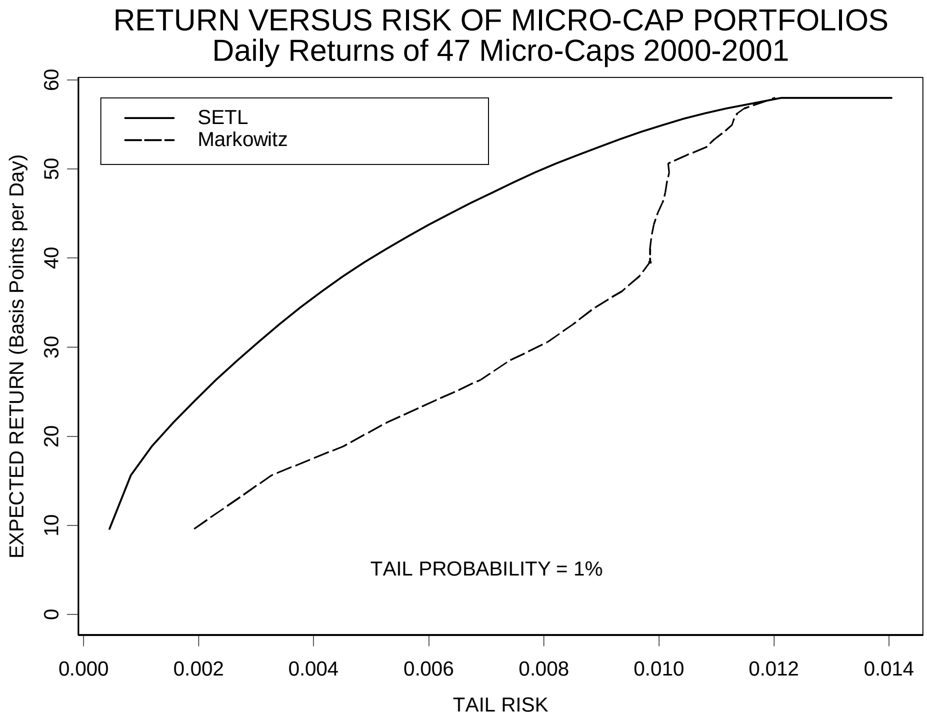
The Markowitz optimal portfolio can also be constructed by solving the obvious dual optimization problem.

The subscript 2 is used in ω_2 as a reminder that $\alpha = 2$ you have the limiting Gaussian distribution member of the stable distribution family, and in that case the Markowitz portfolio is optimal. Alternatively you can think of the subscript 2 as a reminder that the Markowitz optimal portfolio is a second-order optimal portfolio, i.e., an optimal portfolio based on only first and second moments.

The Markowitz investor ends up with a different portfolio, i.e., a different set of portfolio weights with different risk versus return characteristics, than the SETL investor. It is important to note that the performance of the Markowitz portfolio, like that of the SETL portfolio, is evaluated under a GMstable distributional model. If in fact the distribution of the returns were exactly multivariate normal (which they never are) then the SETL investor and the Markowitz investor would end up with one and the same optimal portfolio. However, when the returns are non-Gaussian SETL returns, the Markowitz portfolio is sub-optimal. This is because the SETL investor constructs his/her optimal portfolio using the correct distribution model, while the Markowitz investor does not. Thus the Markowitz investors frontier lies below and to the right of the SETL efficient frontier, as shown in the figure below, along with the Markowitz tangency portfolio T_2 and Markowitz capital market line CML_2.

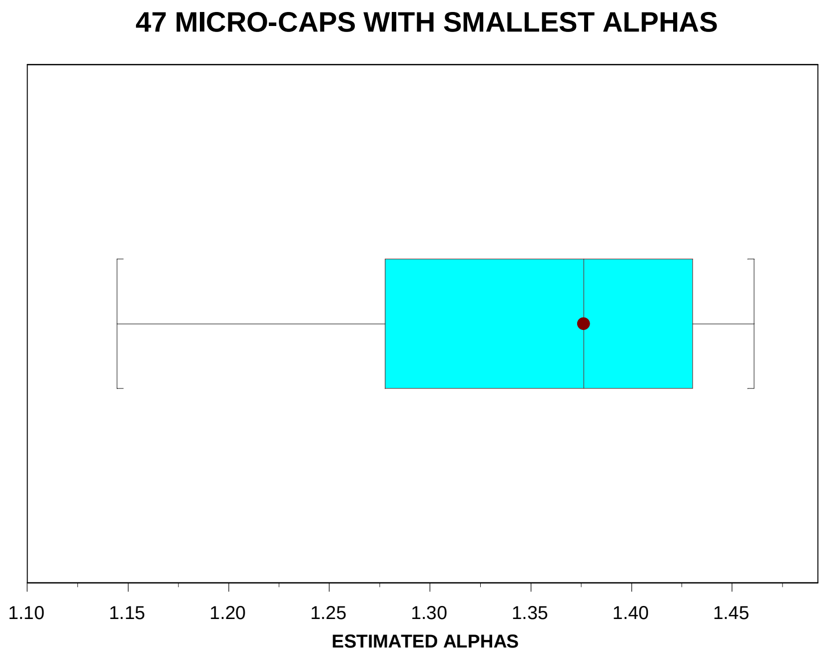


As an example of the performance improvement achievable with the SETL optimal portfolio approach, we computed the SETL efficient frontier and the Markowitz frontier for a portfolio of 47 micro-cap stocks with the smallest alphas from the random selection of 182 micro-caps in section 2.1. The results are displayed in the figure below. The results are based on 3,000 scenarios from the fitted GMstable distribution model based on two years of daily data during years 2000 and 2001. We note that, as is generally the case, each of the 47 stock returns has its own estimate stable tail index $\hat{\alpha}_i$, $i = 1, 2, \ldots, 47$.



Here we have plotted values of $TailRisk = \varepsilon \cdot SETL(\varepsilon)$, for $\varepsilon = .01$, as a natural decision theoretic risk measure, rather than $SETL(\varepsilon)$ itself. We note that over a considerable range of tail risk the SETL efficient frontier dominates the Markowitz frontier by 14 – 20 bp's daily!

We note that the 47 micro-caps with the smallest alphas used for this example have quite heavy tails as indicated by the box plot of their estimated alphas shown below.



Here the median of the estimated alphas is 1.38, while the upper and lower quartiles are 1.43 and 1.28 respectively. Evidently there is a fair amount of information in the non-Gaussian tails of such micro-caps that can be exploited by the SETL approach.

4.4 From Sharpe to STARR- and R- Performance Measures

The *Sharpe Ratio* for a given portfolio *p* is defined as follows:

$$SR_p = \frac{ER_p - r_f}{\sigma_p} \tag{2}$$

where ER_p is the portfolio expected return, σ_p is the portfolio return standard deviation as a measure of portfolio risk, and r_f is the risk-free rate. While the Sharpe ratio is the single most widely used portfolio performance measure, it has several disadvantages due to its use of the standard deviation as risk measure:

- σ_p is a symmetric measure that does not focus on downside risk
- σ_p is not a coherent measure of risk (see Artzner et. al., 1999)
- σ_p has an infinite value for non-Gaussian stable distributions.

Stable Tail Adjusted Return Ratio

As an alternative performance measure that does not suffer these disadvantages, we propose the *Stable Tail Adjusted Return Ratio* (*STARR*) defined as:

$$STARR_p(\varepsilon) = \frac{SER_p - r_f}{SETL_p(\varepsilon)}. \tag{3}$$

Referring to the first figure in section 4.3, one sees that a SETL optimal portfolio produces the maximum *STARR* under a SETL distribution model, and that this maximum *STARR* is just the slope of the SETL capital market line CML_α. On the other hand the maximum *STARR* of a Markowitz portfolio is equal to the slope of the Markowitz capital market line CML_2. The latter is always dominated by CML_α, and is equal to CML_α only in the case where the returns distribution is multivariate normal in which case $\alpha = 2$ for all asset and risk factor returns. Referring to the second figure of section 4.3, one sees that for relatively high risk-free rate of 5 bps per day, the STARR for the SETL portfolio dominates that of the Markowitz portfolio. Furthermore this dominance appears quite likely to persist if the efficient frontiers were calculated for lower risk and return positions and smaller risk-free rates were used.

We conclude that the risk adjusted return of the SETL optimal portfolio ω_α is generally superior to the risk adjusted return of the Markowitz mean variance optimal portfolio ω_2. The SETL framework results in improved investment performance.

Rachev Ratio (R-Ratio)

The R-ratio is the ratio between the expected excess tail-return at a given confidence level and the expected excess tail loss at another confidence level:

$$\rho(r) = \frac{ETL_{\gamma_1}(x'(r_f - r))}{ETL_{\gamma_2}(x'(r - r_f))}$$

Here the levels γ_1 and γ_2 are in [0,1], *x* is the vector of asset allocations and r-r_f is the vector of asset excess returns. Recall that if *r* is the portfolio return, and *L* = -*r* is the portfolio loss, we define the expected tail loss as $ETL_{\alpha\%}(r) = E(L / L > VaR_{\alpha\%})$, where $P(L > VaR_{\alpha\%}) = \alpha$, and α is in (0,1). The R-Ratio is a generalization of the STARR. Choosing appropriate levels γ_1 and γ_2 in optimizing the R-Ratio the investor can seek the best risk/return profile of her portfolio. For example, an investor with portfolio allocation maximizing the R-Ratio with $\gamma_1 = \gamma_2 = 0.01$ is seeking exceptionally high returns and protection against high losses.

4.5 The Choice of Tail Probability

We mentioned earlier that when using $SETL_p(\varepsilon)$ rather than $VaR_p(\varepsilon)$, risk managers and portfolio optimizers may wish to use other values of ε than the conventional VaR values of .01 or .05, for example values such as .1, .15, .2, .25 and .5 may be of interest. The choice of a particular ε amounts to a choice of particular risk measure in the *SETL* family of measures, and such a choice is equivalent to the choice of a utility function. The tail probability parameter ε is at the asset manager's disposal to choose according to his/her asset management and risk control objectives.

Note that choosing a tail probability ε is not the same as choosing a risk aversion parameter. Maximizing

$$SER_p - c \cdot SETL_p(\varepsilon)$$

for various choices of risk aversion parameter c for a fixed value of ε merely corresponds to choosing different points along the SETL efficient frontier. On the other hand changing ε results in different shapes and locations of the SETL efficient frontier, and corresponding different SETL excess profits relative to a Markowitz portfolio.

It is intuitively clear that increasing ε will decrease the degree to which a SETL optimal portfolio depends on extreme tail losses. In the limit of $\varepsilon = .5$, which may well be of interest to some managers since it uses the average loss below zero of L_p as its penalty function, small to moderate losses are mixed in with extreme losses in determining the optimal portfolio. There is some concern that some of the excess profit advantage relative to Markowitz portfolios will be given up as ε increases. Our studies to date indicate, not surprisingly, that this effect is most noticeable for portfolios with smaller stable tail index values.

It will be interesting to see going forward what values of ε will be used by fund managers of various types and styles.

A generalization of the SETL efficient frontier is the R-efficient frontier, obtained by replacing the stable portfolio expected return SER_p in $SER_p - c \cdot SETL_p(\varepsilon)$ by the excess tail return , the numerator in the R- ratio. R-efficient frontier allows for fine tuning of the tradeoff between high excess means returns and protection against large loss.

4.6 The Cognity Implementation of the SETL Framework

The SETL framework described in this paper has been implemented in the ***Cognity™*** Risk Management and Portfolio Construction platform. This product contains solutions for holdings-based market risk and returns-based multi-manager portfolio risk management and optimization, all with integrated factor models, scenario generation and stress testing. ***Cognity™*** is implemented in a modern Java based server architecture to support both desktop and Web delivery. For further details see www.finanalytica.com.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the extensive help provided by Stephen Elston and Frederic Siboulet in the preparation of this paper.

New York	London	Sofia
100 Park Avenue, 16th Floor	1 Liverpool Street	21 Srebarna Str., Floor 5
New York, NY 10017 USA	London EC2M 7QD UK	1407 Sofia, Bulgaria
Tel: +1 (212) 880 - 2670	Tel: +44 (0) 20 7956 2778	Tel: +359 2 962 4645

www.finanalytica.com

REFERENCES

Bradley, B. O. and Taqqu, M. S. (2003). "Financial Risk and Heavy Tails", in *Handbook of Heavy Tailed Distributions in Finance*, edited by S. T. Rachev, Elsevier/North-Holland.

Fama, E. (1963). "Mandelbrot and the Stable Paretian Hypothesis", *Journal of Business*, 36, 420-429.

Mandelbrot, B. B. (1963). "The Variation in Certain Speculative Prices", *Journal of Business*, 36, 394-419.

Rachev, Svetlozar, and Mittnik, Stefan (2000). *Stable Paretian Models in Finance*. J. Wiley.

Rachev, Svetlozar, Menn, Christian and Fabozzi, Frank J. *Fat Tailed and Skewed Asset Return Distributions: Implications for Risk*, Wiley-Finance, 2005

Racheva-Iotova , B., Stoyanov, S. and Rachev S. (2003) , Stable Non-Gaussian Credit Risk Model; The Cognity Approach, in "Credit Risk (Measurement, Evaluations and Management", edited by G.Bol, G.Nakhaheizadeh, S.Rachev, T.Rieder, K-H.Vollmer , Physica-Verlag Series: Contributions to Economics, Springer Verlag, Heidelberg,NY, 179-198.

Rockafellar, R. T. and Uryasev, S. (2000). "Optimization of Conditional Value-at-Risk", *Journal of Risk*, 3, 21-41.

Sklar, A. (1996). "Random Variables, Distribution Functions, and Copulas – a Personal Look Backward and Forward", in Ruschendorff et. al. (Eds.) *Distributions with Fixed Marginals and Related Topics*, Institute of Mathematical Sciences, Hayward , CA.

Stoyanov S. , Racheva-Iotova, B. (2004) "Univariate Stable Laws in the Field of Finance-Parameter Estimation, Journal of Concrete and Applied Mathematics,2,369-396.