Portfolio optimization for Student t and skewed t returns

Wenbo Hu
Quantitative Trader
Bell Trading
111 W Jackson Blvd, Suite 1122
Chicago, IL 60604
312-379-5343
whu@belltrading.us

Alec N. Kercheval
Associate Professor
Department of Mathematics, Florida State University
Tallahassee, FL 32306-4510
850-644-2202
kercheval@math.fsu.edu

May 13, 2007
revised Nov 14, 2008

Abstract

It is well-established that equity returns are not Normally distributed, but what should the portfolio manager do about this, and is it worth the effort? It is now feasible to employ better multivariate distribution families that capture heavy tails and skewness in the data; we argue that among the best are the Student t and skewed t distributions. These can be efficiently fitted to data, and show a much better fit to real returns than the Normal distribution. By examining efficient frontiers computed using different distributional assumptions, we show, using for illustration 5 stocks chosen from the Dow index, that the choice of distribution has a significant effect on how much available return can be captured by an optimal portfolio on the efficient frontier.

Portfolio optimization requires balancing risk and return; for this purpose one needs to employ some precise concept of "risk". Already in 1952, Markowitz used the standard deviation (StD) of portfolio return as a risk measure, and, thinking of returns as Normally distributed, described the efficient frontier of fully invested portfolios having minimum risk among those with a specified return. This concept has been extremely valuable in portfolio management because a rational portfolio manager will always choose to invest on this frontier.

The construction of an efficient frontier depends on two inputs: a choice of risk measure (such as StD, VaR, or ES, described below), and a probability distribution used to model returns.

Using StD (or equivalently, variance) as the risk measure has the drawback that it is generally insensitive to extreme events, and sometimes these are of most interest to the investor. Value at Risk (VaR) better reflects extreme events, but it does not aggregate risk in the sense of being subadditive on portfolios. This is a well-known difficulty addressed by the concept of a "coherent risk measure" in the sense of Artzner, et. al. [1999]. A popular example of a coherent risk measure is expected shortfall (ES), though VaR is still more commonly seen in practice.

Perhaps unexpectedly, the choice of risk measure has no effect on the actual efficient frontier when the underlying distribution of returns is Normal – or more generally any "elliptical" distribution. Embrechts, McNeil, and Straumann [2001] show that when returns are elliptically distributed, the minimum risk portfolio for a given return is the same whether the risk measure is standard deviation, VaR, ES, or any other positive, homogeneous, translation-invariant risk measure.

This fact suggests that the portfolio manager should pay at least as much attention to the family of probability distributions chosen to model returns as to the choice of which risk measure to use.

It is now commonly understood that the multivariate Normal distribution is a poor model of generally acknowledged "stylized facts" of equity returns:

- return distributions are fat-tailed and skewed

- volatility is time-varying and clustered

- returns are serially uncorrelated, but squared returns are serially correlated.

The aim of this paper is to focus attention on the question of what underlying family of distributions should be used for fitting and describing returns data in portfolio optimization problems. We show why portfolio managers should use heavy-tailed, rather than Normal, distributions as models for equity returns – especially the multivariate Student t and skewed t distributions. Recently other authors have also argued, with different data that these distributions are empirically superior, e.g. Keel, et. al. [2006], and Aas and Hobaek Haff [2006].

Does it really cost anything to find optimal portfolios by fitting a Normal distribution to returns data, rather than some heavier-tailed choice? The answer is yes: not only do other distributions do a better job of modeling extreme events, but using them allows the manager to capture portfolio returns that are unrecognized when using the Normal model.

We illustrate this below with a portfolio of 5 stocks using daily log-returns data to optimize the one-day forecast portfolio return at a fixed risk level. We use a GARCH filter to remove serial correlations of squared log-returns; we then fit this approximately *i.i.d.* five dimensional data using a selection of potential distributions from the Generalized Hyperbolic family, including Normal, hyperbolic (Hy), Normal inverse Gaussian (NIG), variance gamma (VG), Student t, and skewed t (defined below). We observe that the Student t and skewed t have the largest log likelihood, despite having one fewer parameter than Hy, NIG, or VG.

After discussion of coherent risk measures, value at risk, and expected shortfall, we examine the problem of portfolio optimization for these different risk measures and returns distributions, concentrating on the Student t and skewed t. We show (proposition 6) that for zero skewness, these distributions produce the same efficient frontiers no matter which risk measure or degree of freedom is chosen, so long as the same means and correlations are used. Nevertheless, our data set illustrates how much potential return is lost by a manager who estimates Normal distributions with returns that are in reality closer to Student t or skewed t distributed. The reason is that fitting the tails better, as the t distributions do, leads to better estimates of the correlations, which in turn affects the efficient frontier.

The Multivariate Generalized Hyperbolic Distributions

The family of multivariate skewed t distributions is a subfamily of the larger family of "generalized hyperbolic (GH) distributions", introduced by Barndorff-Nielsen [1977] and championed for financial applications in McNeil, Frey, and Embrechts [2005].

These distributions are usefully understood as examples of a nice class of distributions called Normal mean-variance mixture distributions, defined as follows.

Definition 1 Normal Mean-Variance Mixture. *The d-dimensional random variable $\mathbf{X}$ is said to have a multivariate Normal mean-variance mixture distribution if*

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + W\boldsymbol{\gamma} + \sqrt{W}\mathbf{Z}, \quad \textit{where} \tag{1}$$

1. *$\mathbf{Z} \sim N_k(\mathbf{0}, \Sigma)$, the k-dimensional Normal distribution with mean zero and covariance Σ (a positive semi-definite matrix),*
2. *$W \geq 0$ is a positive, scalar-valued r.v. which is independent of $\mathbf{Z}$, and*
3. *$\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are parameter vectors in $\mathbb{R}^d$.*

The mixture variable W can be interpreted as a shock which changes the volatility and mean of an underlying Normal distribution. From the definition, we can see that, conditional on W, $\mathbf{X}$ is Normal:

$$\mathbf{X} \mid W \sim N_d(\boldsymbol{\mu} + W\boldsymbol{\gamma}, W\Sigma), \tag{2}$$

and

$$E(\mathbf{X}) = \boldsymbol{\mu} + E(W)\boldsymbol{\gamma} \tag{3}$$

$$COV(\mathbf{X}) = E(W)\Sigma + var(W)\boldsymbol{\gamma}\boldsymbol{\gamma}' \tag{4}$$

the latter defined when the mixture variable W has finite variance $var(W)$.

If the mixture variable W is generalized inverse Gaussian (GIG) (see Appendix), then $\mathbf{X}$ is said to have a generalized hyperbolic distribution (GH). As described in the Appendix, the GIG distribution has three real parameters, λ, χ, ψ, and we write $W \sim N^{-}(\lambda, \chi, \psi)$ when W is GIG.

Therefore the multivariate generalized hyperbolic distribution depends on three real paramters λ, χ, ψ, two d-dimensional parameter vectors $\boldsymbol{\mu}$ (location) and $\boldsymbol{\gamma}$ (skewness) in $\mathbb{R}^d$, and a $d \times d$ positive semidefinite matrix Σ. We then write

$$\mathbf{X} \sim GH_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \boldsymbol{\gamma}, \Sigma).$$

Some Special Cases

Hyperbolic distributions (Hy):

When $\lambda = 1$, we get the multivariate generalized hyperbolic distribution whose univariate margins are one-dimensional hyperbolic distributions. (For $\lambda = (d+1)/2$, we get the d-dimensional hyperbolic distribution. However, its marginal distributions are no longer hyperbolic.)

The one dimensional hyperbolic distribution is widely used in the modeling of univariate financial data, for example in Eberlein and Keller [1995] and Farjado and Farias [2003].

Normal Inverse Gaussian distributions (NIG):

When $\lambda = -1/2$, the distribution is known as Normal inverse Gaussian (NIG). NIG is also commonly used in the modeling of univariate financial returns. Hu [2005] contains a fast estimation algorithm.

Variance Gamma distribution (VG):

When $\lambda > 0$ and $\chi = 0$, we get a limiting case known as the variance gamma distribution. For the variance gamma distribution, we can estimate all the parameters including λ; see Hu [2005].

Skewed t Distribution:

If $\lambda = -\nu/2, \chi = \nu$ and $\psi = 0$, we obtain a limiting case which is called the skewed t distribution by Demarta and McNeil [2005], because it generalizes the Student t distribution, obtained from the skewed t by setting the skewness parameter $\boldsymbol{\gamma} = 0$. The skewed t can also be described as a

Normal mean-variance mixture distribution, where the mixture variable W is inverse gamma $Ig(\nu/2, \nu/2)$; see McNeil, Frey, and Embrechts [2005].

The one-dimensional Student t distribution is widely used in modeling univariate financial data because, in comparison to the Normal, it easily incorporates a heavy tail with a single extra parameter (the degree of freedom). The EM (expectation-maximization) algorithm, discussed below, now makes practical the use of the multivariate Student t distribution for multivariate data.

Without skewness, the Student t is elliptical, and therefore predicts, for example, that joint crashes have the same likelihood as joint booms. This partly motivates the introduction of skewness with the skewed t.

For convenience, explicit density functions of the skewed t distributions are given in the Appendix. The mean and covariance of a skewed t distributed random vector $\mathbf{X}$ are

$$E(\mathbf{X}) = \boldsymbol{\mu} + \boldsymbol{\gamma}\frac{\nu}{\nu-2} \tag{5}$$

$$COV(\mathbf{X}) = \frac{\nu}{\nu-2}\Sigma + \boldsymbol{\gamma}\boldsymbol{\gamma}'\frac{2\nu^2}{(\nu-2)^2(\nu-4)} \tag{6}$$

where the covariance matrix is defined when $\nu > 4$, and the expectation when $\nu > 2$.

Furthermore, in the limit as $\boldsymbol{\gamma} \to \mathbf{0}$, we get the joint density function of the Student t distribution:

$$f(\boldsymbol{x}) = \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}(1 + \frac{\boldsymbol{\rho}(\boldsymbol{x})}{\nu})^{-\frac{\nu+d}{2}} \tag{7}$$

with mean and covariance

$$E(\mathbf{X}) = \boldsymbol{\mu}, \quad COV(\mathbf{X}) = \frac{\nu}{\nu-2}\Sigma \tag{8}$$

The Portfolio Property

A great advantage of the generalized hyperbolic distributions with this parametrization is that they are well-behaved under linear transformation. More precisely (see McNeil, Frey, and Embrechts [2005]), if

$$\boldsymbol{X} \sim GH_d(\lambda, \chi, \psi, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$$

and $\boldsymbol{Y} = B\boldsymbol{X} + \boldsymbol{b}$ for $B \in \mathbb{R}^{k\times d}$ and $\boldsymbol{b} \in \mathbb{R}^k$, then

$$\boldsymbol{Y} \sim GH_d(\lambda, \chi, \psi, B\boldsymbol{\mu} + \boldsymbol{b}, B\Sigma B', B\boldsymbol{\gamma}) \tag{9}$$

In particular, if $\boldsymbol{X} \sim SkewedT_d(\nu, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$, we have

$$\boldsymbol{Y} \sim SkewedT_k(\nu, B\boldsymbol{\mu} + \boldsymbol{b}, B\Sigma B', B\boldsymbol{\gamma}) \tag{10}$$

Forming a linear portfolio $y = \boldsymbol{\omega}^T\boldsymbol{X}$ of the components of $\boldsymbol{X}$ amounts to choosing $B = \boldsymbol{\omega}^T = (\omega_1, \cdots, \omega_d)$ and $\boldsymbol{b} = \mathbf{0}$. In this case,

$$y \sim GH_1(\lambda, \chi, \psi, \boldsymbol{\omega}^T\boldsymbol{\mu}, \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega}, \boldsymbol{\omega}^T\boldsymbol{\gamma})$$

or, in the skewed t case,

$$y \sim SkewedT_1(\nu, \boldsymbol{\omega}^T\boldsymbol{\mu}, \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega}, \boldsymbol{\omega}^T\boldsymbol{\gamma}) \tag{11}$$

That is, all portfolios share the same degree of freedom ν.

This also shows that the marginal distributions are automatically known once we have estimated the multivariate generalized hyperbolic distributions, i.e., $X_i \sim SkewedT_1(\nu, \mu_i, \Sigma_{ii}, \gamma_i)$.

Estimation of Student t and Skewed t Distributions Using the EM Algorithm

The mean-variance representation of the multivariate skewed t distribution has the great advantage that the EM algorithm is directly applicable to the estimation problem. See McNeil, Frey, and Embrechts [2005] for a general discussion of this algorithm for estimating generalized hyperbolic distributions.

The EM (expectation-maximization) algorithm is a two-step iterative process in which (the E-step) an expected log likelihood function is calculated using current parameter values, and then (the M-step) this function is maximized to produce updated parameter values. After each E and M step, the log likelihood is increased, and the method converges to a maximum log likelihood estimate of the distribution parameters.

What helps this along is that the skewed t distribution can be represented as a conditional Normal distribution, so most of the parameters $(\Sigma, \boldsymbol{\mu}, \boldsymbol{\gamma})$ can be estimated, conditional on W, like a Gaussian distribution. See Hu [2005] for details of our implementation and comparisons with other versions.

Returns, Risk Measures and Portfolio Optimization

It will help to be more precise about what is meant by "return" in order to clarify later approximations. If P_t is the price of an asset or portfolio at time t, then the time t return over a unit time interval could mean either *arithmetic return:* $(P_t - P_{t-1})/P_{t-1}$ or *log-return*[1]*:* $\log(P_t/P_{t-1})$.

The log-return is most natural when studying time series, since the log-return over a longer time interval is the sum of the log-returns over sub-intervals. However, when studying portfolios, it is the arithmetic return that is more natural, because the arithmetic return of a portfolio is the weighted average of the arithmetic returns of the individual portfolio securities, weighted by capital value.

When returns are small, as in our study of daily stock returns, the difference between the log-return and the arithmetic return is negligible. However, for clarity we will distinguish these from now on.

Now we turn to a discussion of portfolio risk. Suppose $\boldsymbol{\omega}^T = (\omega_1, \cdots, \omega_d)$ is the capital amount invested in each risky security in a portfolio, and $\mathbf{X}^T = (X_1, \cdots, X_d)$ is the arithmetic return of each risky security. The portfolio property of arithmetic returns is that the total return of the portfolio is $\boldsymbol{\omega}^T\mathbf{X}$. Let

$$L(\boldsymbol{\omega}, \mathbf{X}) = -\sum_{i=1}^{d} \omega_i X_i = -\boldsymbol{\omega}^T\mathbf{X}$$

denote the loss of this portfolio over a fixed time interval Δ and F_L its distribution function. (The time interval Δ is usually one, ten, or 30 days for equity portfolio management.)

From the portfolio property of GH distributions, if $\boldsymbol{X}$ has distribution $N_d(\boldsymbol{\mu}, \Sigma)$ (Normal), $t_d(\nu, \boldsymbol{\mu}, \Sigma)$ (Student t), or $SkewedT_d(\nu, \boldsymbol{\mu}, \Sigma, \boldsymbol{\gamma})$ (skewed t), then the loss $L(\boldsymbol{\omega}, \boldsymbol{X})$ has distribution

$$L \sim N_1(-\boldsymbol{\omega}^T\boldsymbol{\mu}, \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega}) \tag{12}$$

$$L \sim t_1(\nu, -\boldsymbol{\omega}^T\boldsymbol{\mu}, \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega}) \tag{13}$$

[1]As usual, "log" denotes the natural logarithm.

or

$$L \sim SkewedT_1(\nu, -\boldsymbol{\omega}^T\boldsymbol{\mu}, \boldsymbol{\omega}^T\Sigma\boldsymbol{\omega}, -\boldsymbol{\omega}^T\boldsymbol{\gamma}) \tag{14}$$

respectively.

Whatever the model distribution of the loss random variable L, we independently need to choose a risk measure that associates L with some numerical measure of risk.

Definition 2 Value at Risk *Given a confidence level α between 0 and 1 (such as 99% or 95%), the VaR at confidence level α is the smallest value l such that the probability that the loss L exceeds l is no larger than $(1-\alpha)$. In other words,*

$$VaR_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}$$

For the Normal and Student t distributions, the following explicit VaR formulas are easy to verify. When the loss L is Normally distributed with mean μ and variance σ^2, then

$$VaR_\alpha = \mu + \sigma\Phi^{-1}(\alpha) \tag{15}$$

where Φ denotes the standard Normal distribution function. When the loss L is Student t distributed, $L \sim t_1(\nu, \mu, \sigma^2)$, then

$$VaR_\alpha = \mu + \sigma t_\nu^{-1}(\alpha) \tag{16}$$

where t_ν denotes the distribution function of the Student t with degree of freedom ν.

It's helpful to consider more generally some desirable properties for a risk measure.

Definition 3 Coherent Risk Measure *(Artzner et. al. [1999]). A real valued function ρ of a random variable is a coherent risk measure if it satisfies the following properties,*

1. **Subadditivity**. *For any two random variables X and Y, $\rho(X+Y) \leq \rho(X) + \rho(Y)$.*

2. **Monotonicity**. *For any two random variables* $X \geq Y$ *,* $\rho(X) \geq \rho(Y)$*.*

3. **Positive homogeneity**. *For* $\lambda \geq 0$*,* $\rho(\lambda X) = \lambda\rho(X)$*.*

4. **Translation invariance**. *For any* $a \in \mathbb{R}$*,* $\rho(a + X) = a + \rho(X)$*.*

In the language above, StD is not a coherent risk measure; VaR is a coherent measure if the underlying distribution is elliptical, but not generally. Expected shortfall (ES), also called Conditional Value at Risk (see Rockafellar and Uryasev [2002]), is always coherent.

Definition 4 Expected Shortfall (ES**).** *For a continuous loss distribution with* $\int_{\mathbb{R}} |l| dF_L(l) < \infty$*, the* ES_α *at confidence level* $\alpha \in (0, 1)$ *for loss L of a security or a portfolio is defined to be*

$$ES_\alpha = E(L|L \geq VaR_\alpha) = \frac{\int_{VaR_\alpha}^{\infty} l dF_L(l)}{1 - \alpha} \tag{17}$$

$$= \frac{\int I_{\{-(\boldsymbol{\omega}^T \mathbf{x}) \geq VaR_\alpha\}}[-(\boldsymbol{\omega}^T \mathbf{x})] f(\mathbf{x}) d\mathbf{x}}{1 - \alpha} \tag{18}$$

ES can also be computed explicitly for some loss distributions: if L is Normally distributed $N(\mu, \sigma^2)$, then

$$ES_\alpha = \mu + \sigma \frac{\psi(\Phi^{-1}(\alpha))}{1 - \alpha} \tag{19}$$

where ψ is the density of standard Normal distribution. If L is Student t distributed $t(\nu, \mu, \sigma^2)$, then

$$ES_\alpha = \mu + \sigma \frac{f_\nu(t_\nu^{-1}(\alpha))}{1 - \alpha} \left(\frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu - 1} \right) \tag{20}$$

where f_ν is the density function of the Student t with degree of freedom ν.

For skewed t, there is no closed formula for VaR or ES – but see below.

Next we need the concept of an *elliptical distribution*. Briefly, an elliptical distribution is an affine transform of a spherical distribution; a spherical distribution is one which is invariant under rotations and reflections (that is, spherically symmetric). Explicit definitions are available from many sources, e.g. Bradley and Taqqu [2002]. The Normal and Student t distributions are elliptical; the skewed t is not when $\boldsymbol{\gamma} \neq \mathbf{0}$.

Proposition 5 Efficient Frontier for Elliptical Distributions. *(Embrechts, McNeil, and Straumann [2001]). Suppose* $\mathbf{X}$ *is elliptically distributed and all univariate marginals have finite variance. For any* $r \in \mathbb{R}$*, let*

$$\mathcal{Q} = \{Z = \sum_{i=1}^{d} \omega_i X_i | \omega_i \in \mathbb{R}, \sum_{i=1}^{d} \omega_i = 1, E(Z) = r\}$$

be the set of all fully invested portfolio returns with expectation r*. Then for any positively homogeneous, translation invariant risk measure* ρ*,*

$$argmin_{Z \in \mathcal{Q}} \rho(Z) = argmin_{Z \in \mathcal{Q}} \sigma_Z^2.$$

This proposition means that if we assume that the underlying distribution is elliptical, then the Markowitz minimum variance portfolio, for a given return, will be the same as the optimized portfolio obtained by minimizing any other translation invariant and positively homogeneous risk measure, such as VaR or ES. That is, the portfolio allocation does not depend on the choice of risk measure (or confidence level), but only on the choice of distribution.

The skewed t distribution is not elliptical if $\boldsymbol{\gamma} \neq \mathbf{0}$. In this case we see in practice that the efficient portfolios do depend on the choice of confidence level, and on the whether we use VaR, ES, or StD. The practitioner might view this as a disadvantage of using distributions with skewness – she will have to decide whether the data show enough skewness to justify the need to confront these extra choices.

A practical disadvantage of skewed distributions is that we do not have a closed form formula for VaR or ES. Instead, we turn to Monte Carlo simulation to minimize ES at confidence level α by sampling the multivariate distribution of returns. (The Monte Carlo method also can be applied to the elliptical distributions mentioned above, yielding results that are identical to the naked eye.)

More specifically, from (18), we can rewrite the definition of expected shortfall as follows,

$$ES_\alpha = VaR_\alpha + \frac{\int [-(\boldsymbol{\omega}^T \mathbf{x}) - VaR_\alpha]^+ f(\mathbf{x}) d\mathbf{x}}{1 - \alpha},$$

where $[x]^+ := max(x, 0)$.

We get a new objective function by replacing VaR by p,

$$F_\alpha(\boldsymbol{\omega}, p) = p + \frac{\int[-(\boldsymbol{\omega}^T\mathbf{x}) - p]^+ f(\mathbf{x})d\mathbf{x}}{1-\alpha}. \tag{21}$$

Rockafellar and Uryasev [2002] showed that ES can be minimized by minimizing this convex function with respect to $\boldsymbol{\omega}$ and p. If the minimum is attained at $(\boldsymbol{\omega}^*, p^*)$, then $\boldsymbol{\omega}^*$ is the optimized portfolio composition (and $p*$, if unique for that $\boldsymbol{\omega}^*$, is the corresponding portfolio's VaR at confidence level α). In general the minimizing portfolio need not be unique, though this does not matter to our efficient frontier analysis below. (Empirically, we always observed a unique minimizer.)

Below, we sample the multivariate density by Monte Carlo simulation to estimate $F_\alpha(\boldsymbol{\omega}, p)$ by

$$\hat{F}_\alpha(\boldsymbol{\omega}, p) = p + \frac{\sum_{k=1}^{n}[-(\boldsymbol{\omega}^T\mathbf{x}_k) - p]^+}{n(1-\alpha)}, \tag{22}$$

where $\mathbf{x}_k$ is the k-th sample from some distribution and n is the number of samples.

Data Sets and estimation

For illustration, we consider in this study portfolios composed of the following 5 stocks: WALT DISNEY, EXXON MOBIL, PFIZER, ALTRIA GROUP and INTEL, and use adjusted daily closing prices for the period 7/1/2002 to 08/04/2005. The prices are converted to daily log-returns, and from Exhibit 1 we can see that squared log-returns series show some evidence of serial correlation.

For each stock, we use a $GARCH(1,1)$ model with Gaussian innovations for each stock to remove the observed serial return dependence. That is, we fit parameters α_0, α_1 and β_1 in the following $GARCH(1,1)$ model of the log-return series X_t:

$$X_t = \sigma_t Z_t \text{ where } Z_t \sim N(0,1) \quad i.i.d., \tag{23}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{24}$$

We then think of Z_t as a "filtered return" which we hope is $i.i.d.$

From Exhibit 2, we can see that the squared filtered log-returns series show no evidence of serial correlation; Exhibit 3 shows that heteroscedasticity clearly exists in the returns of the five stocks.

After obtaining the approximately $i.i.d.$ filtered data, we can estimate the multivariate density. Note that in the $GARCH$ fitting we assume filtered marginal returns are Gaussian in order to arrive at best fit *pseudo-maximum-likelihood* $GARCH$ parameters (see Bradley and Taqqu [2002]), but the multivariate distribution that best fits the filtered data need not be *a posteriori* Gaussian.

By examining QQ-plots versus Normal for those five stocks, or otherwise, it is easily verified that a Normal distribution is not a good fit in the tails (see Hu [2005] for some illustrative plots). Therefore we consider several distributions in the generalized hyperbolic family to model the multivariate density.

Exhibit 4 shows the maximized log likelihood for fitting the filtered log-returns to various distributions. It shows again that all the generalized hyperbolic distributions we examine have higher log likelihood than the Normal distribution, and the skewed t has the highest log likelihood, with the Student t close behind. (QQ plots suggest this is due to better fit in the tails.)

Since models with more parameters can have higher log likelihood simply due to overfitting of data, we also consider the Akaike information criterion (AIC), which adds to the log likelihood function a penalty proportional to the number of model parameters in order to create a measure of the goodness of fit of a model. Exhibit 4 shows that the models with the optimal AIC are still the Student t and the skewed t, with the Student t now coming out slightly ahead. (The related measure known as the Bayesian information criterion (BIC) leads to similar outcomes.)

Efficient Frontier Analysis

We now study the possible efficient frontiers, for this 5-stock universe, as we vary the risk measure (StD, 99% VaR, 99% ES) and the modeling distribution (Normal, Student t, skewed t).

Suppose we are standing at August 4, 2005, the last date in our data set, and the holding period is one day. 750 sample data are used in the estimation. The one day ahead forecasted $GARCH$ volatilities for all the stocks are denoted $\boldsymbol{\sigma} = (\sigma_1, \cdots, \sigma_5)^T$ at that date. The weight constraint condition is written as

$$\sum_{i=1}^{5} \omega_i = 1, \tag{25}$$

where we assume the initial capital is 1 and ω_i is the capital invested in risky stock i. We suppose short sales are allowed.

Suppose that the estimated filtered expected log-return of stock i is $\hat{\mu}_i$. The de-filtered forecasted log-return is then $\mu_i = \sigma_i \hat{\mu}_i$; let

$$\boldsymbol{\mu} = (\mu_1, \cdots, \mu_5)^T. \tag{26}$$

We have used log-returns in estimating the multivariate distribution with time-series data, but now this raises the question of how we compute portfolio return. To compute the expected portfolio log-return corresponding to a portfolio $\boldsymbol{\omega}$, we should first convert the individual log returns to arithmetic returns, weight them with $\boldsymbol{\omega}$ to obtain the portfolio arithmetic return, and finally convert back to log return. However, for our daily data with a 1-day horizon, the difference between log returns and arithmetic returns is negligible[2], so we use $\boldsymbol{\omega}^T \boldsymbol{\mu}$ as a close approximation to the expected portfolio log-return. (This approximation means that our use of Proposition 5 below for elliptical distributions is only approximately true, though in practice the approximation error is not insignificant.)

We now set the expected portfolio log return equal to a constant c,

$$\boldsymbol{\omega}^T \boldsymbol{\mu} = c, \tag{27}$$

and find the efficient frontier by minimizing StD, VaR, or ES subject to the constraints (25) and (27).

[2]The typical error in portfolio log-returns due to neglecting the conversion between log-return and arithmetic return is one part in 10^5 or less for our 1 day horizon, and therefore is commonly ignored. For much longer horizons, the extra conversion step may be needed.

Normal Frontier

Under the modeling assumption that returns are Normal, we estimate the mean and covariance of a multivariate Normal with our filtered data, and then compute the Normal efficient frontier.

Exhibit 5 shows the filtered expected log returns and the $GARCH$ volatility forecast on Aug 4, 2005, as well as the best-fit correlation matrix for the filtered returns of the five stocks.[3]

We numerically minimize the portfolio variance, 99% VaR, and 99% ES to compute the Normal frontiers. Exhibit 6 shows the portfolio compositions and the corresponding StD, 99% VaR and 99% ES. Because the Normal distribution is elliptical, we expect and see that these three risk measures all yield the same portfolio composition for a given return. Exhibit 7 shows two efficient frontiers plotted against 99% ES – where the objective function is either variance (StD) or 99% ES. The two frontiers are the same because the optimal portfolios are the same. Note also that changing the confidence level of the objective function will leave the picture unchanged as long as we plot the same variable on the horizontal axis.

Student t Frontier

Exhibit 8 shows the expected log return and GARCH volatility forecast for the filtered data when fitting a Student t distribution; also shown is the correlation matrix for the five stocks.[4] The estimated degree of freedom is $\nu = 5.87$.

Since the Student t distribution is elliptical, we again expect the same portfolio compositions on the Student t frontier, whether we minimize StD, 99% VaR, or 99% ES. This is confirmed in Exhibit 9.

From this table, we can also see that the portfolio compositions are different than those of the Normal frontier. Exhibit 10 displays the StD and

[3]The expected return $\hat{\boldsymbol{\mu}}$ and covariance matrix $\hat{\Sigma}$ are first estimated with filtered returns. We then restore the de-filtered expected return $\boldsymbol{\mu}$ and covariance matrix Σ by $\mu_i = \hat{\mu}_i \sigma_i$ and $\Sigma = A\hat{\Sigma}A$, where A=Diag($\boldsymbol{\sigma}$).

[4]The expected return $\hat{\boldsymbol{\mu}}$ and dispersion matrix $\hat{\Sigma}$ are first estimated using filtered returns. We then restore the de-filtered expected return $\boldsymbol{\mu}$ and dispersion matrix Σ by $\mu_i = \hat{\mu}_i \sigma_i$ and $\Sigma = A\hat{\Sigma}A$, where A=Diag($\boldsymbol{\sigma}$).

ES (coincident) frontiers on 99% ES - return axes, with the Normal frontier included for reference.

Normal vs. Student t frontiers

From the portfolio property and our explicit formulas (15), (16), (19), (20) for VaR and ES, we have

$$VaR_\alpha = \boldsymbol{\omega}^T \boldsymbol{\mu} + c_1 \boldsymbol{\omega}^T \Sigma \boldsymbol{\omega} \tag{28}$$

where c_1 is a constant depending only on α for the Normal distribution, and a different constant depending only on α and ν for the Student t distribution. Similarly,

$$ES_\alpha = \boldsymbol{\omega}^T \boldsymbol{\mu} + c_2 \boldsymbol{\omega}^T \Sigma \boldsymbol{\omega} \tag{29}$$

for c_2 depending only on α and ν.

Since $\boldsymbol{\omega}^T \boldsymbol{\mu}$ is held fixed when minimizing risk for the efficient frontier, all three risk measures StD, VaR, ES will therefore produce the same efficient portfolios for both the Normal and the Student t distributions, provided that we use the same $\boldsymbol{\mu}$ and Σ. Since $\boldsymbol{\mu}$ is the mean of the Student t distribution and, from equation (4), Σ is a scalar multiple of the covariance matrix, we can summarize this as

Proposition 6 Invariance of efficient portfolios. *If the vector of asset returns is multivariate Normal or Student t distributed, with correlation matrix C and mean $\boldsymbol{\mu}$, then the portfolios on the efficient frontier depend on C and $\boldsymbol{\mu}$, but do not depend on the degree of freedom ν or on whether the risk measure is chosen to be StD, VaR, or ES.*

We stress that fitting a Student t distribution generally yields a different correlation matrix than one obtained by fitting a Normal distribution. This proposition means that the difference between the Student t and Normal frontiers in Exhibit 14 is due solely to the different means and correlations that arise in estimating the best-fit Normal or Student t distributions.

The cost of using a Normal model in a Student t world

For a fixed level of expected return, the corresponding fully invested risk-minimizing portfolio depends on which distribution is used to model returns, because of differing estimated means and correlations.

As a first example, suppose Adam is a traditional Markowitz mean-variance manager, using StD as his risk measure. He estimates a multi-variable Normal distribution to his filtered returns, in effect assuming the Normal distribution is a good model for realized returns. Adam now believes his efficient frontier is as shown by the dashed line in Exhibit 11.

However, as we have shown above, the Student t distribution is in fact a better fit to the data. If we suppose that the "true" distribution is the estimated Student t distribution, the actual efficient frontier is shown by the solid line in Exhibit 11. The circles indicate the efficient portfolios that Adam computes under his incorrect Normal assumption, where we are plotting the "true" expected log return and standard deviation based on the Student t distribution. (Note that these portfolios do not lie on the Normal frontier because Adam's computation of risk using his Normal distribution gives him the wrong answer.)

As expected, all of Adam's portfolios lie below the true frontier. The distance between the circles and the solid curve in Exhibit 11 illustrates amount of available return Adam fails to capture because his chosen portfolios do not lie on the real efficient frontier. For moderate levels of risk he could have increased his portfolio expected return by 20 or 30 percent if he had chosen porfolios on the true efficient frontier.

Suppose now that Betty is another manager who uses 99% ES as her risk measure because of its coherence properties, but for convenience she still assumes filtered returns are Normal. Her Normal efficient frontier is plotted as the dashed curve in Exhibit 12. If filtered returns are in fact Student t distributed, then the true efficient frontier is the solid curve in the same figure. Here, the Normal frontier is actually inaccessible. As can be seen from the tables and from the plotted circles, for a fixed return, the risk-minimizing portfolio Betty chooses is actually not the true ES minimizing portfolio but is inside the accessible region. Betty is investing sub-optimally due to her choice of the Normal as modeling distribution.

Exhibit 12 also illustrates that the minimum variance portfolios are identical to the minimum ES portfolios – as expected since the Student t is an elliptical distribution.

A similar discussion, illustrated by plotting frontiers against 99% VaR in Exhibit 13, shows the results for a third hypothetical manager, Carol, a 99% VaR minimizer who assumes returns are Normal.

Skewed t Frontier

Exhibit 4 showed that a fitted skewed t distribution has a slightly higher log likelihood than the Student t because of a small amount of skewness, shown in Exhibit 14, along with the fitted correlation matrix for filtered returns. The estimated degree of freedom is 5.93.

We use Monte Carlo simulation[5] to find the skewed t frontier by minimizing expected shortfall. Exhibit 15 shows the 99% level portfolio compositions and the corresponding 99% ES, and Exhibit 16 shows the 95% level portfolio compositions and the corresponding 95% ES. Since the skewed t distribution is not elliptical, the 99% level and 95% level produce slightly different portfolios.

In Exhibit 17 we show a comparison of the two efficient frontiers, one for each of the two distributions, against 99% ES. We also include 95% ES to illustrate that the confidence level now matters. The skewed t and Student t frontiers are very close for small returns. When returns are large, the two curves diverge. Note that the estimated $\boldsymbol{\mu}$ and Σ are similar for the Student t and skewed t distributions, so the divergence is attributable to the skewness parameter in the skewed t distribution, which affects correlations according to equation (6). Here again, if we suppose that the true distribution of returns

[5] We use the filtered returns series to estimate a skewed t distribution and then use the mean-variance mixture definition to sample from the multivariate skewed t distribution to get the 1,000,000 samples $\hat{X}_{1000000\times 5}$. Specifically, in Matlab, we generate 1,000,000 multivariate Normal distributed random variables with mean $\mathbf{0}$ and covariance $\hat{\Sigma}$, which is estimated using filtered returns series, then we generate 1,000,000 $InverseGamma(\nu/2, \nu/2)$ distributed random variables, finally, we get 1,000,000 multivariate skewed t distributed random variables by using the mean-variance mixture definition. The restored samples are $X = \hat{X}A$, where $A = Diag(\boldsymbol{\sigma})$. The restored means are $\mu_i = (\hat{\mu}_i + \frac{\nu}{\nu-2}\hat{\gamma}_i)\sigma_i$ where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\gamma}}$ are location and skewness parameters respectively estimated using filtered data.

is skewed t, the manager who assumes skewness is zero arrives at the wrong efficient portfolios for large returns. Comparison of Exhibits 20 and 13 show that skewness has a noticeable effect on both the magnitude of the minimum ES for large returns, and on the portfolio composition itself.

Conclusion

Distributions matter. When fitting a Normal distribution to non-Normal data, it is not surprising that we might see inaccurate estimates of means and correlations. This is confirmed with our daily equity price data.

The result is that the composition of optimized portfolios can be quite sensitive to the kind of modeling distribution chosen. This is the case even though our data set shows very few extreme events, so the tails of the distribution are not being directly observed in any detail.

The Student t distribution forms a better fit (in the sense of log likelihood) to our equity data than does the Normal or several other common families of Generalized Hyperbolic distributions. The skewed t is slightly better. When passing from Normal to Student t, the estimated filtered means and dispersion matrices (Σ) change substantially, leading to a noticeable effect on the efficient frontiers. Introducing skewness with the skewed t distribution does not change the estimated correlations or means much, but the skewness still affects the efficient portfolios for larger values of expected return. There is some evidence of skewness in our data, but the increased log likelihood obtained by introducing a skewness parameter is small.

Estimation of the Student t and skewed t distributions can be accomplished with the EM algorithm. In the case of the skewed t, we lack explicit formulas for VaR or ES, so that we must use Monte Carlo simulation to compute risk in that case.

Since non-elliptical distributions are in many ways less convenient, managers may choose for simplicity to opt for the Student t over the skewed t distribution, especially when the estimated skewness may be small. The Student t distribution is easy enough, compared to the Normal, that we recommend managers graduate at least to that family. They can find a much better fit to the data at the cost of only one extra parameter (ν), and, because we believe real returns are fat-tailed, capture much more of the true available portfolio return at a given true level of risk.

Acknowledgement: We thank Fred Huffer for statistical advice and the referees for helpful suggestions that improved this paper. Remaining errors are our own.

Appendix

Distribution Formulas

Definition 7 Generalized Inverse Gaussian distribution(GIG). *The random variable X is said to have a generalized inverse Gaussian(GIG) distribution if its probability density function is*

$$h(x;\lambda,\chi,\psi)=\frac{\chi^{-\lambda}(\sqrt{\chi\psi})^{\lambda}}{2K_{\lambda}(\sqrt{\chi\psi})}x^{\lambda-1}exp\left(-\frac{1}{2}(\chi x^{-1}+\psi x)\right),x>0, \tag{30}$$

where K_λ is a modified Bessel function of the third kind with index λ,

$$K_{\lambda}(x)=\frac{1}{2}\int_{0}^{\infty}y^{\lambda-1}e^{-\frac{x}{2}(y+y^{-1})}dy,\quad x>0 \tag{31}$$

and the parameters satisfy

$$\begin{cases}\chi>0,\psi\geq 0 & if\lambda<0\\ \chi>0,\psi>0 & if\lambda=0\\ \chi\geq 0,\psi>0 & if\lambda>0\end{cases}$$

In short, we write $X\sim N^{-}(\lambda,\chi,\psi)$ if X is GIG distributed.

Generalized Hyperbolic Distributions. If the mixing variable $W\sim N^{-}(\lambda,\chi,\psi)$, then the density of the resulting generalized hyperbolic distribution is

$$f(\boldsymbol{x})=c\frac{K_{\lambda-\frac{d}{2}}\left(\sqrt{(\chi+(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))\,(\psi+\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}\right)e^{(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}\boldsymbol{\gamma}}}{\left(\sqrt{(\chi+(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))\,(\psi+\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}\right)^{\frac{d}{2}-\lambda}},$$

(32)

where the normalizing constant is

$$c = \frac{(\sqrt{\chi\psi})^{-\lambda}\psi^{\lambda}(\psi + \boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})^{\frac{d}{2}-\lambda}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}K_{\lambda}(\sqrt{\chi\psi})},$$

and $|\cdot|$ denotes the determinant.

Skewed t Distribution. Let $\mathbf{X}$ be skewed t distributed, and define

$$\rho(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu}). \tag{33}$$

Then the joint density function of $\boldsymbol{X}$ is given by

$$f(\boldsymbol{x}) = c\frac{K_{\frac{\nu+d}{2}}\left(\sqrt{(\nu + \rho(\boldsymbol{x}))\,(\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}\right)e^{(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}\boldsymbol{\gamma}}}{\left(\sqrt{(\nu + \rho(\boldsymbol{x}))\,(\boldsymbol{\gamma}'\Sigma^{-1}\boldsymbol{\gamma})}\right)^{-\frac{\nu+d}{2}}(1 + \frac{\rho(\boldsymbol{x})}{\nu})^{\frac{\nu+d}{2}}}, \tag{34}$$

where the normalizing constant is

$$c = \frac{2^{1-\frac{\nu+d}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}.$$

References

Aas, K. and Hobaek Haff, I.H. "The generalized hyperbolic skew student's t-distribution." *Journal of Financial Econometrics*, Vol. 4, No. 2 (2006).

Artzner, P., F. Delbaen, J. Eber, and D. Heath. "Coherent Measures of Risk." *Mathematical Finance*, Vol. 9 (1999), pp. 203–228.

Barndorff-Nielsen, O. E. "Exponentially decreasing distributions for the logarithm of the particle size." *Proceedings of the Royal Society, Series A: Mathematical and Physical Sciences*, Vol. 350 (1977), pp. 401–419.

Bradley, B. and M. Taqqu. "Financial Risk and Heavy Tails." In S. Rachev, ed., *Handbook of Heavy-tailed distributions in Finance*, North-Holland, 2003.

Demarta, S. and A. McNeil. "The t copula and related copulas." *International Statistical Review*, Vol. 73 (2005), pp. 111–129.

Eberlein, E. and U. Keller, "Hyperbolic distributions in finance." *Bernoulli*, Vol. 1 (1995), pp. 281–299.

Embrechts, P., A. McNeil, and D. Straumann. "Correlation and Dependence in Risk Management: Properties and Pitfalls." In M. Dempster and H. Moffat, eds., *Risk Management: Value at Risk and Beyond,* Cambridge: Cambridge Univ. Press, 2002.

Farjado, J. and A. Farias. "Generalized hyperbolic distributions and Brazilian data." *Ibmec Business School, Finance Lab Working Paper* FLWP-2003-10 (2003).

Hu, W. "Calibration of multivariate generalized hyperbolic distributions using the EM algorithm, with applications in risk management, portfolio optimization, and portfolio credit risk." Florida State University PhD Dissertation, 2005.

Keel, S., F. Herzog, and H. Geering. "Optimal portfolios with skewed and heavy-tailed distributions." *Proceedings of the Third IASTED International Conference on Financial Engineering and Applications,* (2006), pp. 42–48.

McNeil, A., R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton: Princeton Univ. Press, 2005.

Rockafellar, R. and S. Uryasev. "Conditional value-at-risk for general loss distributions." *Journal of Banking and Finance*, Vol. 26, (2002), pp. 1443–1471.

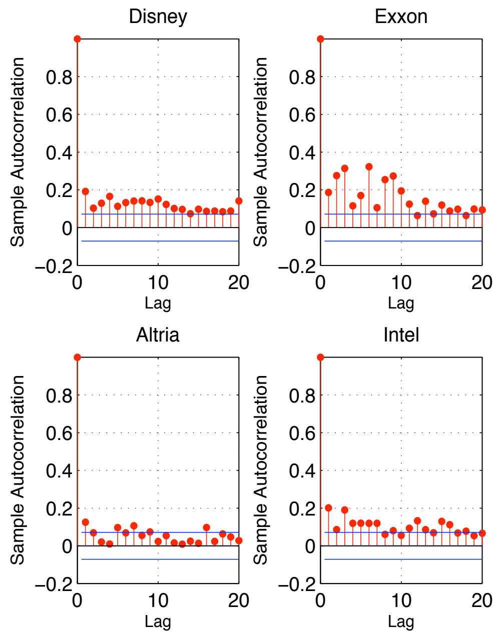

Exhibit 1: Correlograms of squared log return series for 5 stocks

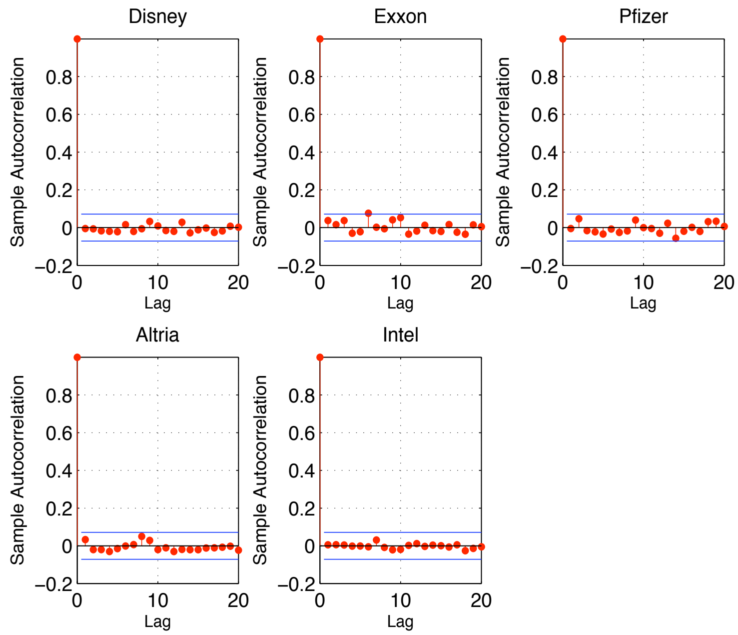

Exhibit 2: Correlograms of squared filtered log return series for 5 stocks

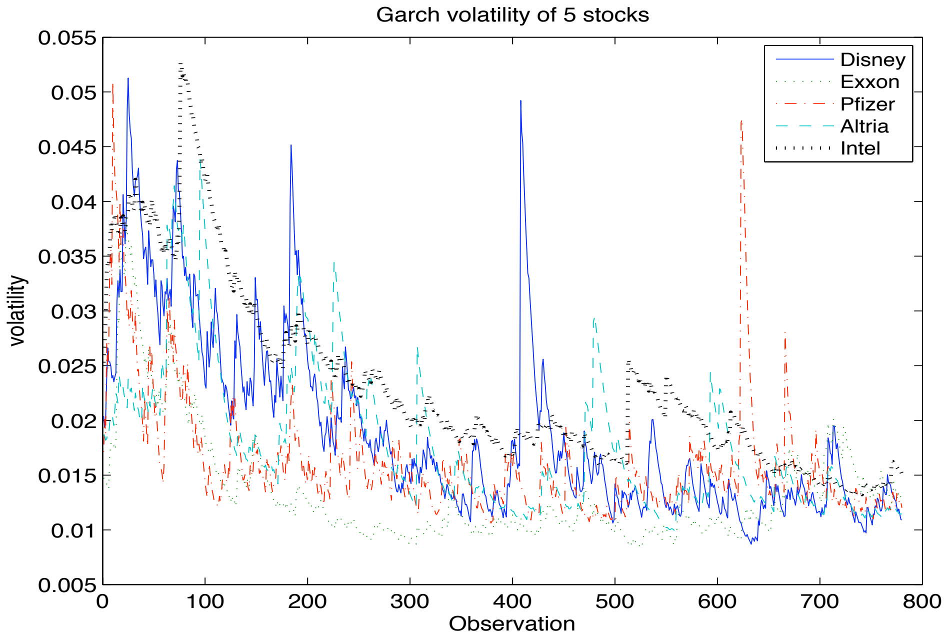

Exhibit 3: $GARCH$ volatility of log return series for five stocks over time (in days)

Model	Normal	Student t	Skewed t	VG	Hyperbolic	NIG
LL	-5095.0	-4877.8	-4873.9	-4901.7	-4891.5	-4884.2
AIC	10230	9798	9800	9858	9838	9822

Exhibit 4: Log likelihood (LL) and Akaike information criterion (AIC) of estimated multivariate densities for six distribution families (larger values of LL are better, smaller values of AIC are better)

Stock	Disney	Exxon	Pfizer	Altria	Intel
Expected filtered return	0.040	0.073	-0.015	0.039	0.027
$GARCH$ volatility	0.0107	0.0128	0.0130	0.0113	0.0156
Correlations: Disney	1				
Exxon	0.367	1			
Pfizer	0.337	0.359	1		
Altria	0.189	0.197	0.215	1	
Intel	0.420	0.303	0.297	0.168	1

Exhibit 5: Expected filtered log returns, one day ahead forecasted $GARCH$ volatility, and correlation matrix obtained by fitting a Normal distribution, on 08/04/2005

Return	StD	$99\%VaR$	99% ES	Disney	Exxon	Pfizer	Altria	Intel
0	0.0096	0.0223	0.0256	0.319	-0.206	0.528	0.320	0.040
0.0002	0.0084	0.0194	0.0222	0.318	-0.038	0.344	0.333	0.043
0.0004	0.0079	0.0180	0.0207	0.318	0.131	0.161	0.345	0.045
0.0006	0.0082	0.0186	0.0214	0.317	0.300	-0.023	0.358	0.048
0.0008	0.0093	0.0209	0.0241	0.317	0.468	-0.206	0.371	0.050
0.001	0.0109	0.0244	0.0281	0.316	0.637	-0.390	0.384	0.052
0.0012	0.0129	0.0287	0.0331	0.316	0.806	-0.573	0.397	0.055
0.0014	0.0150	0.0335	0.0386	0.316	0.974	-0.757	0.409	0.057
0.0016	0.0173	0.0386	0.0444	0.315	1.143	-0.940	0.422	0.060
0.0018	0.0196	0.0438	0.0505	0.315	1.312	-1.124	0.435	0.062
0.002	0.0220	0.0492	0.0567	0.314	1.480	-1.307	0.448	0.065

Exhibit 6: Portfolio composition and corresponding standard deviation, $99\%VaR$ and $99\%ES$ for the Normal frontier

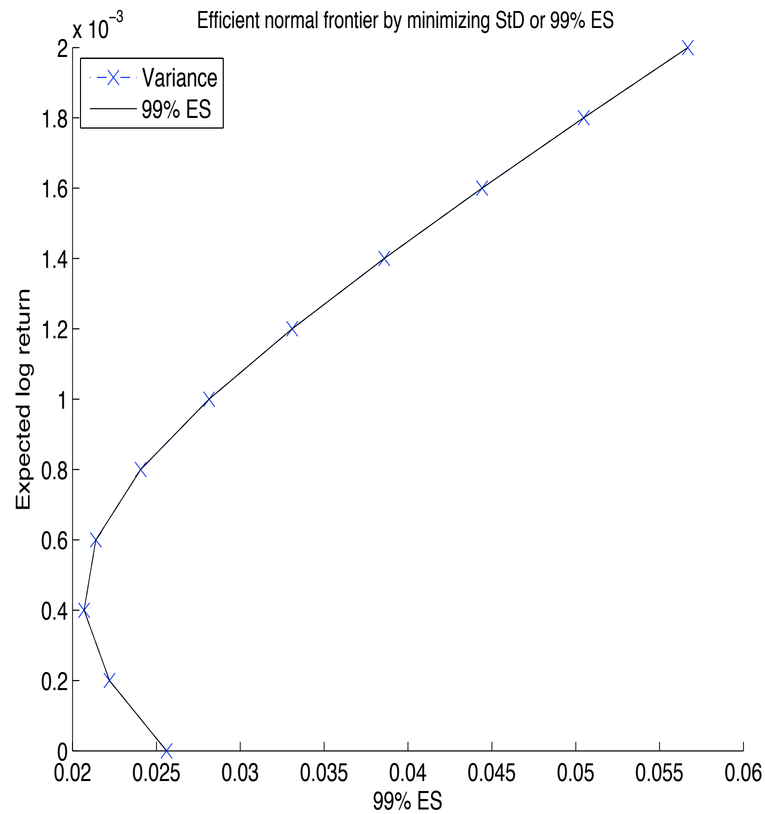

Exhibit 7: The Markowitz efficient frontier and the 99% ES frontier for the Normal distribution. Note that both risk measures give the same results.

Stock	Disney	Exxon	Pfizer	Altria	Intel
Expected filtered return	0.015	0.077	-0.018	0.069	0.030
GARCH volatility	0.0107	0.0128	0.0130	0.0113	0.0156
Correlations: Disney	1				
Exxon	0.363	1			
Pfizer	0.378	0.373	1		
Altria	0.265	0.271	0.259	1	
Intel	0.460	0.324	0.349	0.225	1

Exhibit 8: Expected filtered log returns and correlation matrix for filtered returns obtained by fitting a multivariate Student t distribution

Return	StD	99% VaR	99% ES	Disney	Exxon	Pfizer	Altria	Intel
0	0.0095	0.0245	0.0316	0.494	-0.153	0.447	0.247	-0.035
0.0002	0.0086	0.0218	0.0281	0.410	-0.048	0.315	0.336	-0.014
0.0004	0.0080	0.0203	0.0262	0.326	0.057	0.184	0.425	0.008
0.0006	0.0081	0.0201	0.0261	0.242	0.162	0.052	0.515	0.030
0.0008	0.0086	0.0214	0.0278	0.158	0.267	-0.080	0.604	0.051
0.001	0.0097	0.0238	0.0310	0.074	0.371	-0.211	0.693	0.073
0.0012	0.0110	0.0271	0.0352	-0.010	0.476	-0.343	0.782	0.094
0.0014	0.0126	0.0309	0.0402	-0.094	0.581	-0.474	0.871	0.116
0.0016	0.0143	0.0352	0.0457	-0.178	0.686	-0.606	0.961	0.138
0.0018	0.0161	0.0396	0.0515	-0.262	0.791	-0.737	1.050	0.159
0.002	0.0180	0.0443	0.0576	-0.347	0.895	-0.869	1.139	0.181

Exhibit 9: Portfolio composition and corresponding standard deviation, 99% VaR and 99% ES for the Student t frontier

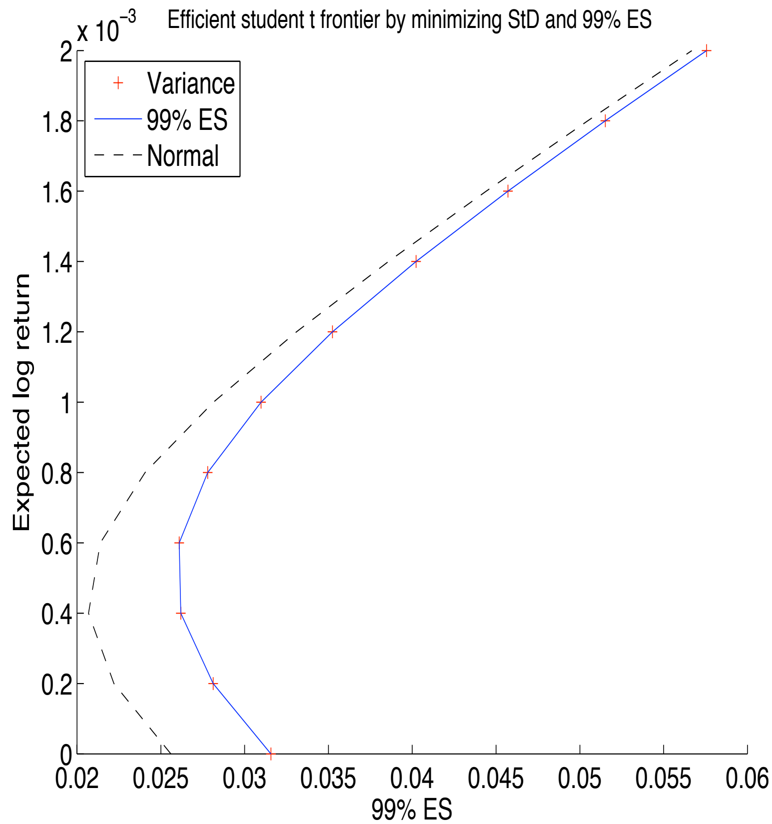

Exhibit 10: Student t frontier and Normal frontier versus 99% ES. The Student t frontier is unchanged if we minimize variance instead of ES

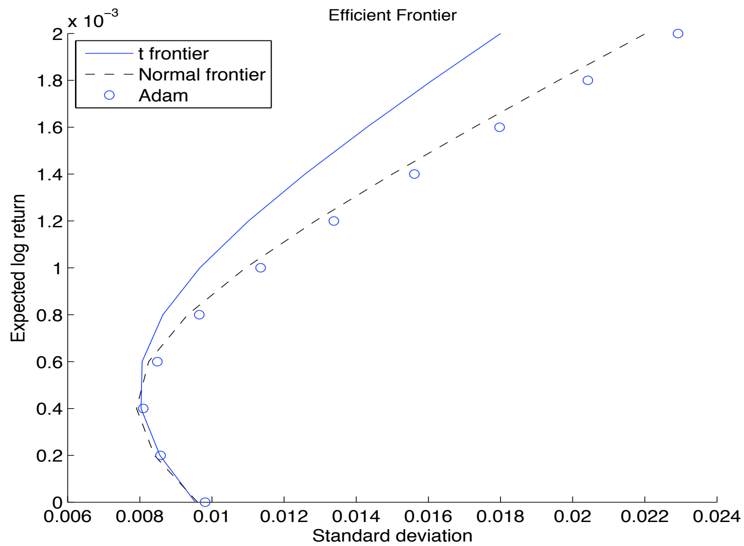

Exhibit 11: The efficient Student t and Normal frontiers vs StD, along with Adam's portfolio optimized under the assumption of Normality

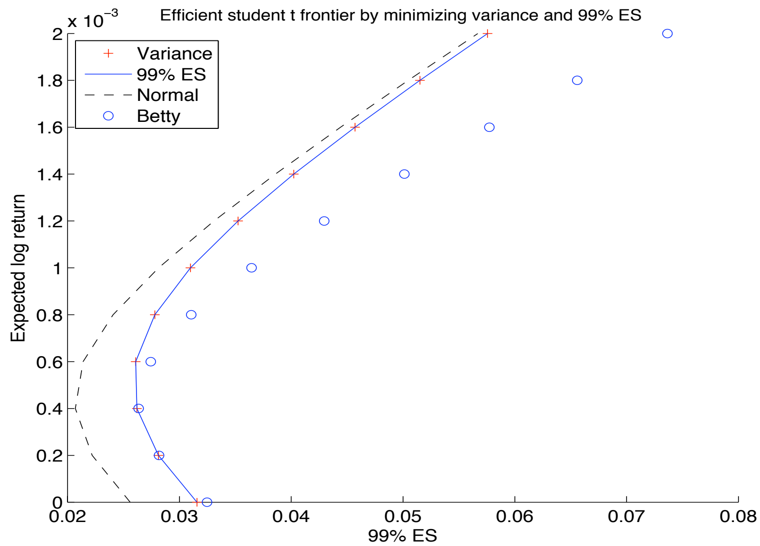

Exhibit 12: The efficient Student t and Normal frontiers vs 99% ES, along with Betty's portfolio optimized under the assumption of Normality

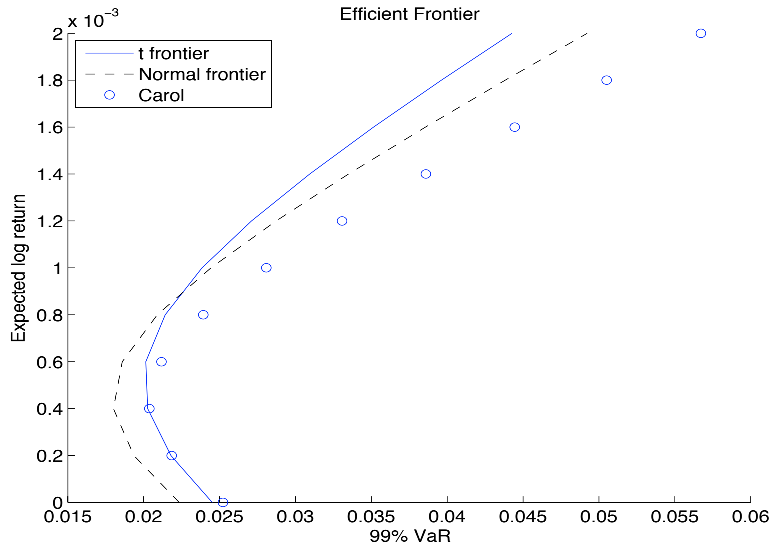

Exhibit 13: The efficient Student t and Normal frontiers vs 99% VaR, along with Carol's portfolio optimized under the assumption of Normality

Stock	Disney	Exxon	Pfizer	Altria	Intel
location parameters	-0.071	0.089	-0.030	0.161	0.042
skewness parameters	0.073	-0.010	0.010	-0.079	-0.010
$GARCH$ volatility	0.0107	0.0128	0.0130	0.0113	0.0156
Correlations: Disney	1				
Exxon	0.269	1			
Pfizer	0.267	0.275	1		
Altria	0.164	0.171	0.157	1	
Intel	0.333	0.244	0.251	0.139	1

Exhibit 14: Expected log returns, skewness parameters, GARCH volatilities, and correlations for filtered returns obtained by fitting a skewed t distribution

Return	99% ES	Disney	Exxon	Pfizer	Altria	Intel
0	0.0320	0.393	-0.219	0.515	0.337	-0.026
0.0002	0.0280	0.383	-0.058	0.328	0.363	-0.015
0.0004	0.0263	0.374	0.101	0.139	0.389	-0.003
0.0006	0.0274	0.381	0.259	-0.051	0.406	0.006
0.0008	0.0312	0.399	0.415	-0.244	0.416	0.013
0.001	0.0367	0.428	0.573	-0.436	0.415	0.021
0.0012	0.0433	0.456	0.733	-0.626	0.413	0.024
0.0014	0.0506	0.485	0.892	-0.817	0.409	0.031
0.0016	0.0583	0.514	1.052	-1.007	0.406	0.036
0.0018	0.0662	0.549	1.209	-1.200	0.404	0.038
0.002	0.0744	0.587	1.365	-1.394	0.399	0.042

Exhibit 15: Portfolio composition and corresponding 99%ES for the skewed t frontier

Return	95% ES	Disney	Exxon	Pfizer	Altria	Intel
0	0.0215	0.354	-0.222	0.515	0.367	-0.013
0.0002	0.0187	0.348	-0.065	0.325	0.393	-0.002
0.0004	0.0175	0.349	0.094	0.136	0.415	0.006
0.0006	0.0182	0.356	0.253	-0.054	0.430	0.014
0.0008	0.0206	0.369	0.412	-0.245	0.442	0.023
0.001	0.0242	0.386	0.570	-0.435	0.449	0.029
0.0012	0.0285	0.407	0.730	-0.625	0.453	0.036
0.0014	0.0333	0.426	0.889	-0.816	0.459	0.042
0.0016	0.0383	0.447	1.047	-1.007	0.466	0.047
0.0018	0.0436	0.470	1.206	-1.197	0.468	0.053
0.002	0.0489	0.492	1.366	-1.387	0.472	0.057

Exhibit 16: Portfolio composition and corresponding 95%ES for the skewed t frontier

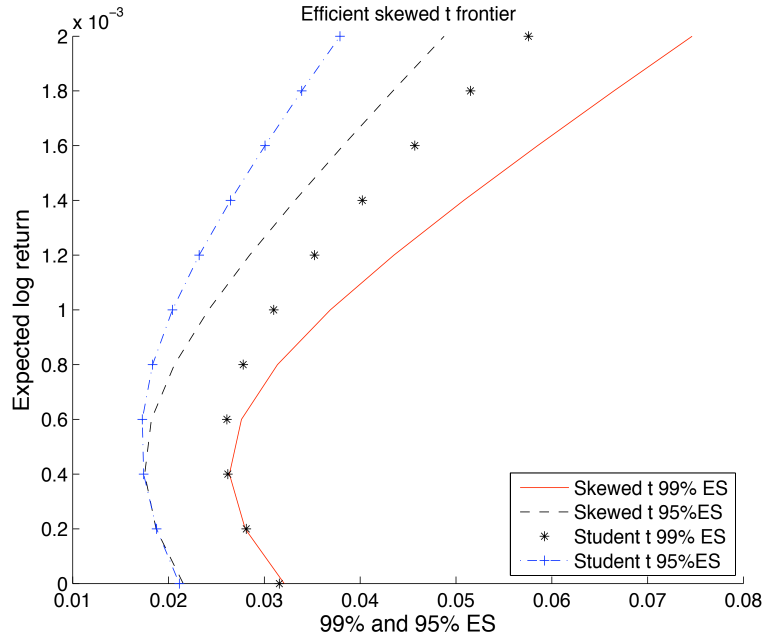

Exhibit 17: Skewed t efficient frontier at 99% ES or 95% ES versus Student t frontier