Published paper

Adcock, C.J. (2010) *Asset pricing and portfolio selection based on the multivariate extended skew-Student-t distribution*, Annals of Operations Research, 176 (1), pp. 221-234
http://dx.doi.org/10.1007/s10479-009-0586-4

Asset Pricing and Portfolio Selection Based on the Multivariate Extended Skew-Student-t Distribution

by

C J Adcock

The University of Sheffield, UK

Abstract

The returns on most financial assets exhibit kurtosis and many also have probability distributions that possess skewness as well. In this paper a general multivariate model for the probability distribution of assets returns, which incorporates both kurtosis and skewness, is described. It is based on the multivariate extended skew-Student-t distribution. Salient features of the distribution are described and these are applied to the task of asset pricing. The paper shows that the market model is non-linear in general and that the sensitivity of asset returns to return on the market portfolio is not the same as the conventional beta, although this measure does arise in special cases. It is shown that the variance of asset returns is time varying and depends on the squared deviation of market portfolio return from its location parameter. The first order conditions for portfolio selection are described. Expected utility maximisers will select portfolios from an efficient surface, which is an analogue of the familiar mean-variance frontier, and which may be implemented using quadratic programming.

Key words: Capital Asset Pricing Model, Efficient frontier, Market Model, Multivariate skew-normal distribution, Multivariate Student distribution, Portfolio selection, Utility functions

Correspondence Address

C J Adcock
The Management School
University of Sheffield
9, Mappin Street
Sheffield, S1 4DT
UK

Tel: +44 (0)114 222 3402
Fax: +44 (0)114 222 3348
Email: c.j.adcock@shef.ac.uk

1. Introduction

Modern portfolio theory makes extensive use, both explicit and implicit, of the normal distribution. Two aspects of the importance of normality of asset returns are as follows. First, when asset returns follow a multivariate normal distribution, the market model, the equation that links assets returns to returns on the market portfolio, is linear and may be correctly estimated using OLS. Secondly, investors who select portfolios by maximising the expected utility of return will always be located on Markowitz' mean-variance efficient frontier. This well-known feature is a consequence of Stein's lemma, Stein (1981). In reality, it is accepted that asset returns are not normally distributed. It is now common practice to recognise the non-normality inherent in returns and to build more appropriate models. Nowadays, there is a very wide variety of models, ARCH/GARCH and neural networks to name just two, that is in common use. For applications which require models of the multivariate distribution of asset returns, there has been substantial growth in recent years of the use of copulae in finance; see for example the recent book by Cherubini *et al* (2004). Such models, however, can leave two issues unresolved. These are (a) the correct form of the market model and (b) the choice of an appropriate utility function for portfolio selection.

The form of the market model is an important issue for the following reason. It is conventionally assumed that this takes the familiar linear regression form. However, the OLS model makes strong assumptions about the form of the multivariate probability distribution of asset returns. In the absence of normality, OLS is not the optimal estimation method. Furthermore in the absence of elliptical symmetry there is no guarantee that the relationship between asset returns and market returns is linear. The choice of utility function is also an important consideration. Under some return distributions, certain choices of the functional form of the utility function are prohibited because the expected values do not exist. From a more pragmatic perspective, the question of whether there is a utility function that is "optimal" in some sense for a given return distribution is always likely to receive attention. In addition, to date anyway, the use of copulae is restricted to a small number of dimensions. This does not of course limit their importance or usefulness, but does mean that applications like large-scale portfolio selection require simplifying assumptions to be made.

The past two decades has seen the development of a large body of theory concerning multivariate probability distributions which are elliptically symmetric. The standard source reference is the well-known monograph by Fang *et al* (1990). These distributions result in market models which are linear, although of course OLS is not the optimal method for parameter estimation. Papers by Liu (1994), Landsman (2006) and Landsman and Nešlehová (2008) present extensions to Stein's lemma which indicate that, *ceteris paribus*, there is a single efficient frontier for all expected utility maximisers.

Self evidently, elliptically symmetric distributions will not deal with asset returns in the presence of skewness, which is an empirical feature of returns of some assets. The aim of this paper is to present the multivariate extended skew-Student-t distribution (henceforth MEST distribution) as a coherent and tractable model for asset returns and hence for the development of various aspects of portfolio theory. It is readily apparent that the use of a single multivariate distribution requires that compromises be made in the ability to estimate marginal distributions. The view taken in this work is that the gains offered by a coherent multivariate model offset the effects of such compromises. Furthermore, as is very well known, estimation error for associated with models used in finance is almost always substantial and it is debatable whether a more complex model will yield superior properties in practice.

The MEST distribution described in this paper is a development of the multivariate skew-normal (MSN) distribution introduced by Azzalini and Dalla-Valle (1996), henceforth AZ&DV. The MSN distribution was first used as a model for returns in finance in Adcock and Shutes (2001), henceforth A&S, who introduced an extension of the model which is often referred to the multivariate extended skew-normal or MESN distribution.

The motivation for an extended skew-Student-t model as distinct to the skew-normal is that kurtosis is almost always present in asset returns. The ability to incorporate fat-tails as well as skewness is therefore an important practical consideration. The use of a Student model is motivated to some extent by tractability, but also by a number of papers spanning several decades that indicate that the Student distribution is a useful model for fat-tailed returns. Two examples of such work include the well-known papers by Praetz (1972) and Blattberg and Gonedes (1974) and the study of returns on European securities by Aparicio and Estrada (2001).

Development of the multivariate skew-Student-t and related distributions is an active area of research. The publication of initial papers by Branco and Dey (2001), Azzalini and Capitanio (2003) and Sahu *et al* (2003) has been followed by articles by several authors including Arellano-Valle and Genton (2005), Arrellano-Valle *et al* (2006), Arrellano-Valle and Azzalini (2006) and Azzalini and Genton (2008). An extended skew Student distribution and many of its properties are described in Arellano-Valle and Genton (2008), henceforth AV&G.

There are also several papers which present other multivariate distributions, all of which possess skewness and which are based on the Student distribution. Examples of such work are in Jones (2001, 2002), Jones and Faddy (2003) and Bauwens and Laurent (2005). There is also a univariate skewed t distribution, which was introduced by McDonald and Xu (1995). This has been generalised and applied to financial data in Theodossiou (1998), who calls the model the skewed generalized t distribution. However, no multivariate form of this model appears in the literature.

The results presented in this paper are used first to derive a market model, which has the appropriate functional form when asset returns follow the MEST distribution. The market model is non-linear. Specifically, it is shown that the MEST market model consists of two components, only the first of which is linear in the return on the market portfolio. If, in addition, the market portfolio has zero skewness, then the linear component is essentially the same as that in the conventional model based on the normal distribution. The non-linear component is a function of the squared deviation of the return of the market portfolio from its location parameter and exhibits different properties depending on the skewness of the market portfolio. Secondly, the first order conditions for portfolio selection are described. There is a single surface on which expected utility maximizing portfolios are located, regardless of individual investors' utility function. Under the MEST distribution, portfolio selection may be implemented using quadratic programming.

The structure of the paper is as follows. Section 2 contains a short summary of the multivariate extended skew-normal distribution. Section 3 describes the derivation and probability density function of the MEST distribution that is used in this paper and presents properties that are required in subsequent sections. This includes the general conditional distribution for the case when the vector of returns $\mathbf{R}$ is partitioned into two components $\mathbf{R}_1$ and $\mathbf{R}_2$ say and $\mathbf{R}_1$ is taken as given. In section 4 this is applied to the situation where $\mathbf{R}_1$ is the scalar return on a market portfolio. The resulting market model is shown to be non-linear in general. Section 5 is concerned with portfolio selection and efficient set mathematics. Section 6 concludes the paper. There are two short appendices.

In the text of this paper, the word beta refers to the conventional measure of risk; that is the covariance of asset returns with the return on the market portfolio divided by market variance. The notation β is reserved for a different definition of risk that arises with market models based on the MEST distribution. However, as is shown in the paper, there are cases where β is essentially the same as beta. Notation is that in common use.

2. The Multivariate Extended Skew-Normal Distribution

There are several ways of deriving multivariate skew-normal distributions. With financial applications in mind, the following method from A&S is used. Let $\mathbf{Y}$ be an n-vector with the full rank multivariate normal distribution $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Let X be a scalar random variable which is independent of $\mathbf{Y}$ and which has a normal distribution with mean τ, variance 1 and is left truncated at zero. The vector of variables $\mathbf{R} = \mathbf{Y} + \boldsymbol{\lambda} X$ has a multivariate extended skew-normal, henceforth MESN, distribution. This may be interpreted as follows. If $\mathbf{R}$ denotes returns on risky financial assets over a single period, the definition above means that each return R_i has two components: Y_i which is normally distributed with mean μ_i and a skewness shock $\lambda_i X$. The shock is generated by the single variable X. The sensitivity of security i to the shock is λ_i, which may take any real value including zero. The skewness shock may be interpreted as a departure from market efficiency in the sense of Fama (1970).

Using the notation above, the distribution of the vector of returns $\mathbf{R}$ has probability density function

$$f(\mathbf{r}) = \varphi_n\left(\mathbf{r};\boldsymbol{\mu} + \boldsymbol{\lambda}\tau, \boldsymbol{\Sigma} + \boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathrm{T}}\right)\Phi\left(\frac{\tau + (\mathbf{r}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}}{\sqrt{1+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}}}\right)\Big/\Phi(\tau), \qquad (1.)$$

where $\varphi_n(\mathbf{x};\boldsymbol{\varsigma},\boldsymbol{\Omega})$ denotes the probability density function of a multivariate normal distribution with mean vector $\boldsymbol{\varsigma}$ and covariance matrix $\boldsymbol{\Omega}$ evaluated at $\mathbf{x}$ and

$\Phi(\tau)$ denotes the standard normal distribution function evaluated at τ. When τ takes the value zero, this is equivalent to the corresponding result in AZ&DV with a change of notation. The explicit use of non-zero values of τ facilitates some subsequent manipulations of the density. It also creates a richer family of probability distributions. The distribution at (1.) is denoted $\mathbf{R} \sim \mathrm{MESN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \tau)$. Moments and related properties of this distribution, expressed in the above notation, are described in A&S. AZ&DV show that when the vector $\mathbf{R}$ is partitioned into two components, $\mathbf{R}_1$ and $\mathbf{R}_2$, containing n_1 and $n_2 = n - n_1$ elements respectively, the conditional distribution of $\mathbf{R}_2$ given that $\mathbf{R}_1 = \mathbf{r}_1$ is multivariate skew-normal of the same family. In the usual way, partition as $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\lambda}$ as

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, V = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda}_1 \\ \boldsymbol{\lambda}_2 \end{bmatrix}.$$

The conditional distribution of $\mathbf{R}_2$ given that $\mathbf{R}_1 = \mathbf{r}_1$ is $\mathrm{MESN}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1} \boldsymbol{\lambda}_{2|1}, \tau_{2|1})$ where

$$\boldsymbol{\mu}_{2|1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{r}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{2|1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12},$$

$$\boldsymbol{\lambda}_{2|1} = \frac{\boldsymbol{\lambda}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}{\sqrt{1 + \boldsymbol{\lambda}_1^{\mathrm{T}}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}}, \tau_{2|1} = \frac{\tau + \boldsymbol{\lambda}_1^{\mathrm{T}}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{r}_1 - \boldsymbol{\mu}_1)}{\sqrt{1 + \boldsymbol{\lambda}_1^{\mathrm{T}}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}}.$$

A&S apply this to the situation where $\mathbf{R}_2$ represents individual asset returns and $\mathbf{R}_1$, which is a scalar variable, is the return on the market portfolio, R_p say. It will be apparent from the form of the parameters in the above equations that the market model, the expected value of the vector of asset returns conditional on the return on the market portfolio, is non-linear in r_p, the given value of the market return. This model may also be applied to general regression. In this case, the conditioning vector $\mathbf{R}_1$ represents selected explanatory variables. The vector $\mathbf{R}_1$ may have a multivariate normal distribution, in which case $\boldsymbol{\lambda}_1 = \mathbf{0}$. In this case, the regression model that links the dependent variables $\mathbf{R}_2$ to the independent variables $\mathbf{r}_1$ is linear in $\mathbf{r}_1$.

In addition to finance applications using the multivariate skew-normal distribution reported in the book editted by Marc Genton, Genton (2004), and for example Adcock(2005), there is a comprehensive study in Harvey *et al* (2004). This uses a version of the distribution that allows for more than one truncated variable: in the notation used above X is now a vector and $\boldsymbol{\lambda}$ is a matrix. Harvey *et al* make the important point that skewness in asset returns is not always evident from a univariate histogram and present a number of bivariate plots to illustrate the point. A recent paper which also uses the distribution is by Meucci (2006).

3. The Multivariate Extended Skew-Student-t Distribution

The MEST distribution may be derived using the standard form of multivariate Student distribution (see for example Johnson & Kotz, 1972, page 162 *et sec* or Bernardo and Smith, 1994, page 435 and 441). The vector $\mathbf{R}$ is defined as above, except that the vector

$$\begin{bmatrix} \mathbf{Y}-\boldsymbol{\mu} \\ X-\tau \end{bmatrix},$$

now has a centred multivariate Student distribution with υ degrees of freedom and dispersion matrix equal to

$$\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix},$$

with the variable X being left-truncated at zero. The joint distribution of $\mathbf{Y}$ and X has probability density function proportional to

$$\left[1+\{(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})+(x-\tau)^2\}/\upsilon\right]^{-\frac{1}{2}(\upsilon+n+1)}\Big/T_\upsilon(\tau)\,,$$

for $x>0$ where $T_\upsilon(\tau)$ denotes the distribution function for the Student-t distribution with υ degrees of freedom evaluated at τ. Standard transformation of the variables gives the joint distribution of $\mathbf{R}$ and X. Writing this joint probability density in the conditional form $f(x\,|\,\mathbf{r})f(\mathbf{r})$ and integrating with respect to x gives the MEST distribution, with probability density function

$$f(\boldsymbol{r})=t_n\left(\mathbf{r};\boldsymbol{\mu}+\boldsymbol{\lambda}\tau,\boldsymbol{\Sigma}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T,\upsilon\right)\frac{T_{\upsilon+n}\left\{\sqrt{\frac{\upsilon+n}{\upsilon}}\,\frac{\tau+(\mathbf{r}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}}{\sqrt{\{1+Q(\mathbf{r},\boldsymbol{\mu},\boldsymbol{\lambda},\tau,\boldsymbol{\Sigma})/\upsilon\}\left(1+\boldsymbol{\lambda}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}\right)}}\right\}}{T_\upsilon(\tau)}. \tag{2.}$$

The quadratic form $Q(\mathbf{r},\boldsymbol{\mu},\boldsymbol{\lambda},\tau,\boldsymbol{\Sigma})$ is defined as

$$Q(\mathbf{r},\boldsymbol{\mu},\boldsymbol{\lambda},\tau,\boldsymbol{\Sigma})=\left(\mathbf{r}-\boldsymbol{\mu}-\boldsymbol{\lambda}\tau\right)^T\left(\boldsymbol{\Sigma}+\boldsymbol{\lambda}\boldsymbol{\lambda}^T\right)^{-1}\left(\mathbf{r}-\boldsymbol{\mu}-\boldsymbol{\lambda}\tau\right). \tag{3.}$$

The function $t_n\left(\mathbf{x};\varsigma,\boldsymbol{\Omega},\upsilon\right)$ denotes the probability density function of an n-variate Student distribution with υ degrees of freedom, location parameter vector ς and dispersion matrix $\boldsymbol{\Omega}$ evaluated at $\mathbf{x}$. As $\upsilon\to\infty$, the MEST density (2.) tends to the multivariate skew-normal at (1.). The notation $\mathbf{R}\sim MEST\left(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\lambda},\tau,\upsilon\right)$ is used to describe the distribution at (2.).

The distribution derived in Sahu *et al* (2003) uses more than one skewness variable. An extended MEST distribution with more that one truncated variable is also reported in AV&G and Adcock (2008). It is a development of similar extensions to the multivariate skew normal distribution described in Gonzàlez-Farías et al (2004) and Arellano-Valle and Azzalini (2006). As AV&G rightly remark these *"are ... generalizations that are of interest from a theoretical point of view"*. As the papers by Horrace (2005) and Adcock (2007) make clear, the computational difficulties alone are substantial.

The rest of this section presents properties of the MEST distribution, which are used in Section 4. A number of other properties of the distribution which are not required are omitted.

3.1 Linear Transformations of Multivariate Extended Skew-Student-t Variables

When $\mathbf{R} \sim MEST(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \tau, \upsilon)$, the linear transformation $\mathbf{MR}+\boldsymbol{\eta}$, where $\mathbf{M}$ is a $m \times n$ matrix of rank less that or equal to n is distributed as $\mathrm{MEST}(\mathbf{M}\boldsymbol{\mu}+\boldsymbol{\eta}, \mathbf{M\Sigma M}^{\mathrm{T}}, \mathbf{M}\boldsymbol{\lambda}, \tau, \upsilon)$. The proof of this is omitted. Allowing for changes in notation, this result is essentially the same as proposition 5 of AV&G.

3.2 A New Elliptically Symmetric Distribution

When the $\boldsymbol{\lambda}$ equals the zero vector, the probability density function is

$$f(r) = t_n(\mathbf{r}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \upsilon) \frac{T_{\upsilon+n}\left\{\sqrt{\frac{\upsilon+n}{\upsilon}} \frac{\tau}{\sqrt{\{1+Q(\mathbf{r}, \boldsymbol{\mu}, \mathbf{0}, \tau, \boldsymbol{\Sigma})/\upsilon\}}}\right\}}{T_{\upsilon}(\tau)}.$$

The effect of $\tau > 0 (< 0)$ is to increase (decrease) the peakedness of the density around $\mathbf{R} = \mathbf{0}$ when compared with the standard multivariate Student density. This distribution is also reported in AV&G, who note that certain values of τ will produce lighter tails than the normal distribution. From the perspective of finance, the potential interest in this distribution lies in the fact that it can be more peaked than the Student distribution. It may therefore have the ability to describe the returns on assets which are thinly traded and for which the frequency of zero returns is higher than that predicted by Student's t. The following result is an interesting by-product of this distribution.

<u>Lemma 1</u>

Let $\mathbf{Y}$ be an n-vector which has a multivariate Student distribution with location parameter vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and υ degrees of freedom. Let the quadratic form $Q(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \tau, \boldsymbol{\Sigma})$ be as defined at equation (3.). The following result holds

$$E\left[T_{\upsilon+n}\left\{\sqrt{\frac{\upsilon+n}{\upsilon}} \frac{\tau}{\sqrt{1+Q(\boldsymbol{r}, \boldsymbol{\mu}, \boldsymbol{0}, 0, \boldsymbol{\Sigma})/\upsilon}}\right\}\right] = T_{\upsilon}(\tau).$$

3.3 Expectation and Variance Covariance Matrix of R

Direct evaluation of the higher unconditional moments of the distribution of $\mathbf{R}$, when they exist, is not straightforward and is beyond the scope of this paper. However, it is straightforward to show that for $\upsilon > 1$ and $\upsilon > 2$, respectively

$$E(X) = \tau + \{\upsilon t_\upsilon(\tau)(1+\tau^2/\upsilon)\}/\{(\upsilon-1)T_\upsilon(\tau)\} = \tau + \xi_\upsilon(\tau),$$

$$E\{(X-\tau)^2\} = -\tau\xi_\upsilon(\tau) + \{\upsilon T_{\upsilon-2}(\tau\sqrt{(\upsilon-2)/\upsilon})\}/\{(\upsilon-2)T_\upsilon(\tau)\} = \eta_\upsilon(\tau),$$

where $t_\upsilon(\tau)$ denotes the density function for the Student-t distribution with υ degrees of freedom evaluated at τ. The unconditional expected value of $\mathbf{R}$ is $E(\mathbf{R}) = \boldsymbol{\mu} + \boldsymbol{\lambda}\tau + \boldsymbol{\lambda}\xi_\upsilon(\tau)$. The conditional covariance matrix of $\mathbf{R}$ given $X = x$ is $var(\mathbf{R} \mid \mathrm{x}) = \upsilon\{1+(x-\tau)^2/\upsilon\}\boldsymbol{\Sigma}/(\upsilon - 1)$. For $\upsilon > 2$, the unconditional covariance matrix of $\mathbf{R}$ is

$$var(\mathbf{R}) = \upsilon\{1+\eta_\upsilon(\tau)/\upsilon\}\boldsymbol{\Sigma}/(\upsilon - 1) + \{\eta_\upsilon(\tau) - \xi_\upsilon(\tau)^2\}\boldsymbol{\lambda}\boldsymbol{\lambda}^T.$$

When all elements of $\boldsymbol{\lambda}$ are equal to zero, the covariance matrix is a function of τ, even though this variable does not then affect expected returns. Since the above result holds for all values of $\boldsymbol{\Sigma}$ and $\boldsymbol{\lambda}$ the following

$$\{1+\eta_\upsilon(\tau)/\upsilon\} \geq 0,\ \{\eta_\upsilon(\tau) - \xi_\upsilon(\tau)^2\} \geq 0\ ,$$

hold for all values of τ. As $\upsilon \to \infty$, the covariance matrix of $\mathbf{R}$ tends to the corresponding expression for the MESN distribution in Section 2.

3.4 Conditional Distributions

Let $\mathbf{R}$ be partitioned into two components, $\mathbf{R}_1$ and $\mathbf{R}_2$, containing n_1 and $n_2 = n - n_1$ elements respectively and the parameters $\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}$ partitioned correspondingly. The conditional distribution of $\mathbf{R}_2$ given that $\mathbf{R}_1 = \mathbf{r}_1$ is also MEST. The derivation is omitted, but the method is the same as that briefly described in Section 2 or in AZ&DV. The conditional distribution of $\mathbf{R}_2$ given $\mathbf{R}_1 = \mathbf{r}_1$ is $\mathrm{MEST}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}, \boldsymbol{\lambda}_{2|1}, \tau_{2|1}, \upsilon + n_1)$ where

$$\boldsymbol{\mu}_{2|1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{r}_1 - \boldsymbol{\mu}_1),\ \boldsymbol{\Sigma}_{2|1} = \frac{\upsilon}{(\upsilon + n_1)}(1 + Q_1/\upsilon)\left(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\right),$$

$$\boldsymbol{\lambda}_{2|1} = \frac{\boldsymbol{\lambda}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}{\sqrt{1 + \boldsymbol{\lambda}_1^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}}\sqrt{\upsilon(1 + Q_1/\upsilon)/(\upsilon + n_1)},$$

$$\tau_{2|1} = \sqrt{\frac{(\upsilon + n_1)}{\upsilon}}\frac{\tau + \boldsymbol{\lambda}_1^T\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{r}_1 - \boldsymbol{\mu}_1)}{\sqrt{(1 + Q_1/\upsilon)(1 + \boldsymbol{\lambda}_1^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1)}},$$

with $Q_1 = Q(\mathbf{r}_1, \boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \tau, \boldsymbol{\Sigma}_{11})$. The expected value of $\mathbf{R}_2$ given that $\mathbf{R}_1 = \mathbf{r}_1$ is a non-linear function $\mathbf{r}_1$. The details of three cases of the conditional distribution: (i) $\boldsymbol{\lambda}_1 = \mathbf{0}$; (ii) at least one element of $\boldsymbol{\lambda}_1$ is non-zero and (iii) $\boldsymbol{\lambda}_{2|1} = \mathbf{0}$ are described in Appendix A. The vector $\boldsymbol{\lambda}_{2|1}$ is referred to as conditional skewness.

4. The Market Model

When $\mathbf{R}$ represents the returns on a set of assets, the return on the market portfolio or market proxy is given by the linear transformation $R_p = \boldsymbol{w}^T\mathbf{R}$, where the vector $\boldsymbol{w}$ represents the weights of each asset. The market model, in which $\mathbf{R}_1$ is the now scalar variable R_p may be established using the results above. From the definition of the distribution, it follows that R_p has a (univariate) extended skew-Student-t distribution $R_p \sim EST\left(\boldsymbol{w}^T\boldsymbol{\mu}, \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}, \boldsymbol{w}^T\boldsymbol{\lambda}, \tau, \upsilon\right)$, equivalently $\mathrm{EST}\left(\mu_p, \sigma_p^2, \lambda_p, \tau, \upsilon\right)$. The distribution of a vector of asset returns $\mathbf{R}_2$ given that $R_p = r_p$ follows the MEST distribution with $\upsilon + 1$ degrees of freedom and parameters

$$\boldsymbol{\mu}_c = \boldsymbol{\mu}_2 + \boldsymbol{\beta}(r_p - \mu_p),\ \boldsymbol{\Sigma}_c = \frac{\upsilon}{(\upsilon + 1)}(1 + \mathrm{Q}_c/\upsilon)\left(\boldsymbol{\Sigma}_{22} - \boldsymbol{\beta}\boldsymbol{\beta}^T/\sigma_p^2\right),$$

$$\boldsymbol{\lambda}_c = \frac{\boldsymbol{\lambda}_2 - \lambda_p\boldsymbol{\beta}}{\sqrt{1 + \lambda_p^2/\sigma_p^2}}\sqrt{\upsilon(1 + \mathrm{Q}_c/\upsilon)/(\upsilon + 1)},\ \tau_c = \sqrt{\frac{(\upsilon + 1)}{\upsilon}}\frac{\tau + (r_p - \mu_p)\lambda_p/\sigma_p^2}{\sqrt{\left(1 + \mathrm{Q}_c/\upsilon\right)\left(1 + \lambda_p^2/\sigma_p^2\right)}}.$$

The vector $\boldsymbol{\beta}$ (which, as stated in the introduction, is not equivalent to the conventional beta) and the quadratic form Q_c are, respectively

$$\boldsymbol{\beta} = \boldsymbol{\Sigma}_{22}\boldsymbol{w}/\sigma_p^2,\ \mathrm{Q_c} = (r_p - \mu_p - \lambda_p\tau)^2/(\sigma_p^2 + \lambda_p^2).$$

The conditional mean vector of $\mathbf{R}_2$ given r_p is therefore

$$E(\mathbf{R}_2 \mid r_p) = \boldsymbol{\mu}_2 + \boldsymbol{\lambda}_2\tau + \boldsymbol{\delta}(r_p - \mu_p - \lambda_p\tau) + \boldsymbol{\lambda}_c\xi_{\upsilon+1}(\tau_c),$$

where $\boldsymbol{\delta}$ is defined as $\boldsymbol{\delta} = \left(\boldsymbol{\Sigma}_{22} + \boldsymbol{\lambda}_2 \boldsymbol{\lambda}_2^T\right)\boldsymbol{w} / \left(\sigma_p^2 + \lambda_p^2\right)$.

This is a generalisation of the result derived for the MESN distribution in A&S. There is a difference in the last term, which reflects the underlying Student distribution. The definitions in Section 3 mean that $\boldsymbol{\delta}$ is not the vector of covariances with the market portfolio divided by market variance and so is not equal to the conventional beta. Nonetheless, $\boldsymbol{\delta}$ measures the linear sensitivity of asset returns to market return. There are three forms of the market model to consider, which depend on the skewness of the market portfolio as follows: (i) $\lambda_p = 0$, (ii) $\lambda_p \neq 0$ and (iii) $\boldsymbol{\lambda}_c = 0$. First, note that the conditional variance of asset returns is time varying and through the quadratic form defined at (3.) depends on the squared deviations of the return on the market portfolio from its location parameter.

Case 1: $\lambda_p = 0$

Since $\lambda_p = \boldsymbol{\lambda}_2^T \boldsymbol{w}$, the case $\lambda_p = 0$ implies that $\boldsymbol{\delta} = \boldsymbol{\beta}$ and that the vector $\boldsymbol{\beta}$ has the usual interpretation as long as $\upsilon > 1$. The conditional expected value is

$$E\left(\mathbf{R}_2 \mid r_p\right) = E\left(\mathbf{R}_2\right) + \boldsymbol{\beta}\left\{r_p - E\left(R_p\right)\right\} + \boldsymbol{\lambda}_2 \Delta_p,$$

where

$$\Delta_p = \sqrt{\upsilon\left(1 + Q_1/\upsilon\right)/\left(\upsilon + 1\right)}\,\xi_{\upsilon+1}\left(\Theta\right) - \xi_\upsilon\left(\tau\right),\ \Theta = \tau / \sqrt{\upsilon\left(1 + Q_1/\upsilon\right)/\left(\upsilon + 1\right)},$$

and $E\left(\Delta_p\right) = 0$.

Case 2: $\lambda_p \neq 0$

When $\lambda_p \neq 0$, the market model takes the form:

$$E\left(\mathbf{R}_2 \mid r_p\right) = E\left(\mathbf{R}_2\right) + \boldsymbol{\delta}\left\{r_p - E\left(R_p\right)\right\} + \tilde{\boldsymbol{\lambda}}_2 \tilde{\Delta}_p,$$

where

$$\tilde{\Delta}_p = \sqrt{\upsilon\left(1 + Q_1/\upsilon\right)/\left(\upsilon + 1\right)}\,\xi_{\upsilon+1}(\tau_c) - \xi_\upsilon(\tau_c) / \sqrt{1 + \lambda_p^2/\sigma_p^2},\ \tilde{\lambda}_2 = \left(\lambda_2 - \beta\lambda_p\right) / \sqrt{1 + \lambda_p^2 \sigma_p^2}$$

The non-linearity in the model is contained in the term $\tilde{\Delta}_p$. This term is non-zero for all values of τ and υ, although its expected value is zero. As already noted above, $\boldsymbol{\delta}$ measures the linear sensitivity of returns to return on the market portfolio, but its definition is different from the conventional beta.

Case 3: $\lambda_c = 0$

The final form of market model arises when $\boldsymbol{\lambda}_c = 0$. For this case, the market model is linear in r_p and the measure of risk is $\boldsymbol{\beta}$, as defined above:

$$E(\mathbf{R}_2 \mid r_p) = E(\mathbf{R}_2) + \boldsymbol{\beta}\{r_p - E(R_p)\}.$$

The conditional distribution of $\mathbf{R}_2$ is the elliptically symmetric modification of the multivariate Student-t described in Section 3.2. The vector $\boldsymbol{\beta}$ has the usual meaning as long as the degrees of freedom υ are greater than one. The interpretation of this result is that the conditioning variable R_p accounts for all the skewness in $\mathbf{R}_2$. It implies that a linear regression model may conceal skewness in the dependent or independent variables or in both.

In general, the conditional expected value is a non-linear function of r_p. Figure 1 shows a sketch, which illustrates the effect of skewness on the conditional expectation. In this context R_2 is the scalar return on an asset or a portfolio of assets. As the sketch indicates, the effect of increasing λ is to change the level of $E(\mathbf{R}_2 \mid r_p)$ and to introduce curvature as λ increases. The straight line corresponding to $\lambda = 0$ corresponds to the security market line or SML. For negative values of λ the corresponding graphs would be convex functions plotting below the SML

Figure 1 about here

The curves shown in Figure 1 are similar to the well known model for market timing introduced by Treynor and Mazuy (1966). In the notation of this paper, their model may be written as

$$E(R_2 \mid r_p) = \gamma_0 + \gamma_1\{r_p - E(r_p)\} + \gamma_2\{r_p - E(r_p)\}^2$$

According to the Capital Asset Pricing Model, the quadratic coefficient γ_2 must be identically zero and the conditional expected return on a portfolio must lie on the SML. Thus, if conditional expected return is found to be significantly greater than zero, the portfolio exhibits better performance than that predicted by theory. The implication is that the manager of such a fund exhibits superior skill and should expect to be rewarded. Conversely, the manager of a fund with a value γ_2 significantly less than zero might expect some form of sanction. However, the implication of the curves shown in Figure 1 is that performance above the security market line may be a consequence of positive skewness in asset returns and not necessarily due to manager skill at all. Conversely, performance below the security market line may be a consequence of negative skewness in asset returns. Thus, neither rewards nor sanctions may be warranted.

5. Portfolio Selection and Efficient Set Mathematics

The first order conditions for portfolio selection depend on expressions of the form:

$$\frac{\partial E[u()]}{\partial w_i} = \int r_i u'(\boldsymbol{w}^T \boldsymbol{r}) f(\boldsymbol{r}) d\boldsymbol{r} = \int cov\{r_i, u'(r_p)\} f(r_i, r_p) dr_i dr_p - E(r_i) E\{u'(r_p)\},$$

where $\boldsymbol{r}$ is the vector of asset returns, r_i is the return on asset *i*, r_p is the return on the portfolio, $f(\boldsymbol{r})$ is the multivariate probability density function of asset returns and $u(r_p)$ is the utility function. When returns follow a multivariate normal distribution the issue of the choice of utility function does not arise because all well-behaved utility functions lead to a point on Markowitz' mean-variance efficient frontier. This result is a consequence of Stein's lemma, Stein(1981), which is applied to general portfolio selection in Kallberg and Ziemba(1983). As noted in the introduction, papers by Liu (1995), Landsman (2006) and Landsman and Nešlehová (2008) present extensions to Stein's lemma for elliptically symmetric distributions. When asset returns follow the multivariate skew normal distribution at (1.), Lemma 1 of Adcock (2007) shows that there is a single mean-variance-skewness surface. Subject to some regularity conditions on $u(.)$ all expected utility maximising investors will be located on this surface regardless of their individual utility functions. Adcock (2008) shows that a similar result obtains for the multivariate extended skew-Student-t distribution $\text{MEST}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \tau, \upsilon)$. Ignoring the Lagrange multiplier of the budget constraint, the first order conditions are a vector equation of the form

$$-\partial E\{u(r_p)\}/\partial \boldsymbol{w} = (\boldsymbol{\Sigma} + \boldsymbol{\lambda\lambda}^T)\boldsymbol{w} - \theta_1 \boldsymbol{\mu} - \theta_2 \boldsymbol{\lambda}, \theta_{1,2} \geq 0,$$

where the scalar quantities $\theta_{1,2}$, which represent preferences for $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ respectively, are functions of the expectations of derivatives of the utility function $u(.)$. When they exist, this equation may be re-expressed in term of the expected return vector, the covariance matrix and $\boldsymbol{\lambda}$ the vector of skewness coefficients. A sketch of the efficient surface is shown in Figure 2. The equations that generate it are essentially the same as those in A&S and are summarised in Appendix B.

Figure 2 about here

It is well known that the moment generating function does not exist for random variables that follow Student distributions. Consequently, when asset returns follow the MEST distribution, the use of the negative exponential function is not available to expected utility maximisers. However, the implication of the regularity conditions is that utility functions may be constructed which are continuous and for which the first two derivatives exist. For given values of $\theta_{1,2}$ portfolio selection may be carried out using quadratic programming.

6. Concluding Remarks

This paper presents the extended multivariate extended skew-Student-t distribution as a model, for asset pricing and portfolio selection. The properties of the distribution are summarised, a non-linear market model is derived and the first order conditions for portfolio selection are presented.

The properties of the market model may be summarised as follows. Under the MEST distribution, there are three forms of market model. First, when the market portfolio does not possess skewness, the market model is non-linear in market portfolio return r_p. However, the non-linearity vanishes as the degrees of freedom increase. In this case, the sensitivity of asset returns to the market is similar to the conventional beta. Secondly, if the market portfolio does possess skewness, then the market model is always non-linear and the sensitivity of asset returns to the market is not equivalent to beta. Finally, in the case where conditional skewness equals zero, the market model is always linear and the sensitivity, as above, is similar to beta. The conditional variance of asset returns is time varying and is dependent on the squared deviation of market portfolio return from its location parameter.

Although there are some restrictions on the choice of utility function, all expected utility maximisers will select portfolios from an efficient surface, which is the analogue of the familiar mean-variance frontier.

The treatment of skewness depends on the use of a single skewness variable. A development of this work will be to extend the model so that multiple skewness shocks may be incorporated. As already noted, this extension is expected to be of mainly theoretical interest, since the practical difficulties associated with such a model are substantial.

Acknowledgements

This is a revised version of a paper originally presented at the Multi-moment Capital Asset Pricing Conference held in Paris in April 2002. It benefitted from the comments of several conference participants, with particular thanks to Renato Flores and Steve Satchell as well as to Emmanuel Jurczenko and Bertrand Maillet, the organisers of the meeting. More recently this work has been presented at the workshop on skew-elliptical distributions held in Bertinoro, Italy in April 2008 organised by Adelchi Azzalini and Antonella Capitanio and at a seminar at the University of Padua. Both this paper and intended future developments have been greatly assisted by the comments of participants at the workshop and at the seminar. I am very grateful to the two organisers of the workshop for the opportunity to participate in the meeting and for their helpful guidance. I am also grateful to Marc Genton for a pre-print of his 2008 paper with Reinaldo Arellano-Valle. Finally, this paper has benefitted greatly from the comments of the anonymous referee and from recent correspondence from Professor Azzalini. Remaining errors and shortcomings are mine alone.

Appendices

A – The Conditional Expected Value

Three cases of the conditional distribution are considered: (i) $\boldsymbol{\lambda}_1 = \mathbf{0}$, (ii) at least one element of $\boldsymbol{\lambda}_1$ is non-zero and (iii) the special case $\boldsymbol{\lambda}_{2|1} = \mathbf{0}$.

Case 1: $\lambda_1 = 0$

When $\boldsymbol{\lambda}_1 = \mathbf{0}$, the conditional expected value of $\mathbf{R}_2$ given $\mathbf{r}_1$ is

$$E(\mathbf{R}_2 \mid \mathbf{r}_1) = \boldsymbol{\mu}_2 + \boldsymbol{\lambda}_2 \tau + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{r}_1 - \boldsymbol{\mu}_1) + \boldsymbol{\lambda}_2 \sqrt{\upsilon(1+Q_1/\upsilon)/(\upsilon+n_1)}\,\xi_{\upsilon+n_1}(\Xi)\,,$$

where $\xi_\upsilon(\ )$ is as defined above and $\Xi = \tau \big/ \sqrt{\upsilon(1+Q_1/\upsilon)/(\upsilon+n_1)}$. Using the formulae in Section 3.7 for the expected value of $\mathbf{R}_2$ and noting that, for $\upsilon > 1$, $E(\mathbf{R}_1) = \boldsymbol{\mu}_1$, the above conditional expectation may be written as

$$E(\mathbf{R}_2 \mid \mathbf{r}_1) = E(\mathbf{R}_2) + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\{\mathbf{r}_1 - E(\mathbf{R}_1)\} + \boldsymbol{\lambda}_2 \left\{\sqrt{\upsilon(1+Q_1/\upsilon)/(\upsilon+n_1)}\,\xi_{\upsilon+n_1}(\Xi) - \xi_\upsilon(\tau)\right\}.$$

Since $\boldsymbol{\lambda}_1 = \mathbf{0}$, $\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} = \mathbf{B}$ is the $n_2 \times n_1$ matrix of theoretical coefficients for the multivariate regression of $\mathbf{R}_2$ on $\mathbf{R}_1$. The regression above has two components. The first is linear in the conditioning vector $\mathbf{r}_1$. The matrix $\mathbf{B}$ has almost the same interpretation as in regression[1]. The residual term

$$\Delta = \sqrt{\upsilon(1+Q_1/\upsilon)/(\upsilon+n_1)}\,\xi_{\upsilon+n_1}(\Xi) - \xi_\upsilon(\tau)\,,$$

is non-linear in $\mathbf{r}_1$. As the degree of freedom parameter υ increases without limit, $\Delta \to 0$ with probability one. The conditional expectation therefore has a limiting linear form as $\upsilon \to \infty$. Taking expectations gives $E(\Delta) = 0$ [2] for all υ. The term Δ acts like a residual; it contributes to variability, but not to unconditional expected return. The implication is that the use of OLS would result in an unbiased estimator of $\mathbf{B}$, but that the results would be inefficient.

Case 2: At least one element of λ_1 not equal to 0

When $\boldsymbol{\lambda}_1 \neq \mathbf{0}$, that is when one or more of the conditioning variables possesses skewness, the conditional expected value may be written as

$$E(\mathbf{R}_2 \mid \mathbf{r}_1) = E(\mathbf{R}_2) + \tilde{\mathbf{B}}\{\mathbf{r}_1 - E(\mathbf{R}_1)\} + \tilde{\Delta}\tilde{\boldsymbol{\lambda}}_2$$

where:

$$\tilde{\mathbf{B}} = (\boldsymbol{\Sigma}_{21} + \boldsymbol{\lambda}_2\boldsymbol{\lambda}_1^T)(\boldsymbol{\Sigma}_{11} + \boldsymbol{\lambda}_1\boldsymbol{\lambda}_1^T)^{-1},\ \tilde{\boldsymbol{\lambda}}_2 = \frac{\boldsymbol{\lambda}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}{\sqrt{1+\boldsymbol{\lambda}_1^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}},$$

$$\tilde{\Delta} = \sqrt{\upsilon(1+Q_1/\upsilon)/(\upsilon+n_1)}\,\xi_{\upsilon+n_1}(\tau_{2|1}) - \xi_\upsilon(\tau) \Big/ \sqrt{1+\boldsymbol{\lambda}_1^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\lambda}_1}\,.$$

The matrix $\tilde{\mathbf{B}}$ still measures the linear sensitivity of the elements of $\mathbf{R}_2$ to elements of $\mathbf{r}_1$. However, the interpretation of $\tilde{\mathbf{B}}$ is different. A second difference is that $\tilde{\Delta}$

[1] It has the same interpretation as long as $\upsilon > 2$. For smaller values of υ, a slightly different interpretation of **B** is needed, although it still determines the sensitivity of elements of $\mathbf{R}_2$ to elements of $\mathbf{R}_1$.

[2] It is straightforward to verify that this is the case.

does not tend to zero as the degrees of freedom increase, although it is of course true that $E(\widetilde{\Delta}) = 0$. Thus, even in the limiting case of multivariate skew-normality, the regression model is non-linear.

Case 3: $\lambda_{2|1} = 0$

A model that is exactly linear with respect to $\mathbf{R}_1$ arises when $\boldsymbol{\lambda}_{2|1} = \widetilde{\boldsymbol{\lambda}}_2 = \mathbf{0}$. In this case, the conditional expected value of $\mathbf{R}_2$ given $\mathbf{r}_1$ is

$$E(\mathbf{R}_2 \mid \mathbf{r}_1) = E(\mathbf{R}_2) + \mathbf{B}\{\mathbf{r}_1 - E(\mathbf{R}_1)\}.$$

Where $\mathbf{B}$ is as defined above. Although this is a special case, it merits a comment. The vector variables $\mathbf{R}_1$ and $\mathbf{R}_2$ have a MEST distribution both jointly and individually. When $\boldsymbol{\lambda}_{2|1} = \widetilde{\boldsymbol{\lambda}}_2 = \mathbf{0}$, the conditional distribution of $\mathbf{R}_2$ given that $\mathbf{R}_1 = \mathbf{r}_1$ is the elliptically symmetric modification of the multivariate Student in Section 3.2. The interpretation of this result is that the conditioning variable $\mathbf{R}_1$ accounts for all the skewness in $\mathbf{R}_2$. In view of the linear transformations property in Section 3.1, this raises the possibility of creating linear combinations of $\mathbf{R}$ which account for the skewness present in all variables. Conversely, it raises the possibility that a linear regression model may conceal skewness in the dependent or independent variables or in both.

B – Efficient Set Mathematics

For given values of θ_1 and θ_2, the solution to the standard efficient set problem $\max_w \mathrm{E}\{u(\boldsymbol{w}^{\mathrm{T}}\mathbf{R})\}$ s.t. $\mathbf{1}^{\mathrm{T}}\boldsymbol{w} = 1$ is the set of equations

$$(\boldsymbol{\Sigma} + \boldsymbol{\lambda}\boldsymbol{\lambda}^T)\boldsymbol{w} - \theta_1\boldsymbol{\mu} - \theta_2\boldsymbol{\lambda} - \zeta\mathbf{1} = \mathbf{0},$$

where ζ is the Lagrange multiplier of the budget constraint, which is the n-vectors of ones, $\mathbf{1}$. A&S show that this leads to an efficient surface as follows. Let the solution be $\boldsymbol{w}^*$ and define

$$E^* = \boldsymbol{w}^T\boldsymbol{\mu},\ \mathrm{V}^* = \boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w},\ \mathrm{S}^* = \boldsymbol{w}^T\boldsymbol{\lambda}.$$

Some algebra gives the following

$$E^* - \alpha_0 = \gamma_0(S^* - \gamma_2)/\gamma_1 + \sqrt{(\alpha_1 - \gamma_0^2/\gamma_1)\{V^* - \alpha_2 - (S^* - \gamma_2)^2/\gamma_1\}},$$

where

$$\alpha_0 = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}/\mathbf{1}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}, \quad \alpha_1 = \boldsymbol{\mu}^T \left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T \boldsymbol{\Sigma}^{-1}/\mathbf{1}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}\right)\boldsymbol{\mu},$$

$$\alpha_2 = 1/\mathbf{1}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}, \quad \gamma_0 = \boldsymbol{\mu}^T \left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T \boldsymbol{\Sigma}^{-1}/\mathbf{1}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}\right)\boldsymbol{\lambda},$$

$$\gamma_1 = \boldsymbol{\lambda}^T \left(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}^T \boldsymbol{\Sigma}^{-1}/\mathbf{1}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}\right)\boldsymbol{\lambda}, \quad \gamma_2 = \boldsymbol{\lambda}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}/\mathbf{1}^T \boldsymbol{\Sigma}^{-1}\mathbf{1}.$$

References

Adcock, C. J. (2005)Exploiting Skewness To Build An Optimal Hedge Fund With A Currency Overlay, *The European Journal of Finance*, **11**, 445-462.

Adcock, C. J. (2007) Extensions of Stein's Lemma for the Skew-normal Distribution, *Communications in Statistics – Theory and Methods*, **36**, 1661-1672.

Adcock, C. J. (2008) Extensions of Stein's and Siegel's Lemmas for the Skew-Student Distribution, *Working Paper*.

Adcock, C. J. & K. Shutes (2001) Portfolio Selection Based on the Multivariate Skew-Normal Distribution in *Financial Modelling* (A Skulimowski, Editor), Krakow , Progress and Business Publishers.

Aparicio, F. & J. Estrada (2000) Empirical distributions of stock returns: European Securities markets, *European Journal of Finance*, **7**, 1-21.

Arrellano-Valle, R. B. and A. Azzalini (2006) On the unification of families of skew-normal distributions, *Scand. J. Statist*., **33**, 561-574.

Arellano-Valle, R. B., M. D. Branco & M. G. Genton (2006). A unified view on skewed distributions arising from selections. *Canad. J. Statist.*, **34**, 581–601.

Arellano-Valle, R. B. and M. G. Genton (2005). On fundamental skew distributions, *J. Multivariate Analysis*, **96**, 93-116.

Arrellano-Valle, R. B. & M . G. Genton (2008) Multivariate Extended Skew-*t* Distributions and Related Families with Applications, Technical Report # 080628, Department of Econometrics, University of Geneva

Azzalini, A. (1985) A Class of Distributions Which Includes The Normal Ones, *Scand. J. Statist*, **12**, 171-178.

Azzalini, A. (1986) Further Results on a Class of Distributions Which Includes The Normal Ones, *Statistica*, **46**, 199-208.

Azzalini, A. & A. Capitanio (2003) Distributions Generated by Perturbation of Symmetry With Emphasis on a Multivariate Skew t Distribution, *J Roy Statist. Soc Series B*, **65**, 367–389.

Azzalini, A. & A. Dalla-Valle (1996) The Multivariate Skew-normal Distribution, *Biometrika*, **83**, 715-726.

Azzalini, A. & M. G. Genton (2008) Robust Likelihood Methods Based on the Skew-*t* and Related Distributions, *International Statistical Review* (2008), **76**, 106–129

Bauwens, L. & S. Laurent (2005) A New Class of Multivariate Skew Densities, With Applications to GARCH Models, *Journal of Business and Economic Statistics,* **23***,* 346-354.

Bernardo, J. M. & A. F. M. Smith (1994) *Bayesian Theory*, New York, John Wiley and Sons Inc.

Blattberg, R. & N. Gonedes (1974) A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices, *Journal of Business*, **47**, 244-80.

Branco, M. D. and D. K. Dey (2001) A general class of multivariate skew-elliptical distributions, *J. Multivariate Anal.*, **79**, 99-113.

Cherubini, U., E. Luciano, & W. Vecchiato (2004) *Copula Methods in Finance.* New York, John Wiley.

Fama, E. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, **25**, 383-417.

Fang, K.-T., S. Kotz & K.-W. Ng, (1990). *Symmetric multivariate and related distributions*. Monographs on Statistics and Applied Probability, Vol. 36. London, Chapman & Hall, Ltd.

Genton, M.G. (2004) *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Edited volume, Boca Raton, Chapman & Hall/CRC Press.

Gonzàlez-Farías, G., A. Dominguez-Molína & A. K. Gupta (2004) Additive Properties of skew normal random vectors, *Journal of Statistical Planning and Inference*, **126**, 521-534.

Johnson, N. L. & S. Kotz (1972) *Distributions in Statistics, Continuous Multivariate Distributions*, New York, John Wiley and Sons Inc.

Jones, M. C. (2001) Multivariate t and Beta Distributions Associated with the Multivariate F Distribution, *Metrika*, **54**, 215-231.

Jones, M. C. (2002) Marginal Replacement in Multivariate Densities, with Application to Skewing Spherically Symmetric Distributions, *Journal of Multivariate Analysis*, **81**, 85-99.

Jones, M. C. & M. J. Faddy (2003 A skew Extension of the t Distribution, with Applications, *J Roy Statist. Soc Series B*, **65**, 159-174.

Harvey, C. R., J. C. Leichty, M. W. Leichty & P. Muller (2004) *Portfolio Selection With Higher Moments*, Working Paper.

Horrace, W. C. (2005). Some results on the multivariate truncated normal distribution. *J. Multivariate Anal.* **94**, 209–221.

Kallberg, J. G. & W. T. Ziemba (1983) Comparison of Alternative Functions in Portfolio Selection Problems, *Management Science*, **11**, 1257 – 1276.

Landsman, Z. (2006) On the generalization of Stein's Lemma for elliptical class of distributions, *Statistics & Probability Letters*, **76,** 1012–1016

Landsman, Z. & J. Nešlehová (2008) Stein's Lemma for elliptical random vectors, *Journal of Multivariate Analysis*, **99**, 912 – 927

Liu, J. S. (1994) Siegel's Formula via Stein's Identities, *Statistics and Probability Letters,* **21**, 247-251.

McDonald, J. B. & Y. J. Xu (1995) A Generalization of the Beta Distribution with Applications, *Journal of Econometrics*, **66**, 133-152.

Meucci, A. (2006) Beyond Black-Littermann: views on non-normal markets, *Risk*, February 2006, 87-92.

Praetz, P. (1972) The Distribution of Share Price Changes, *Journal of Business*, **45**, 49-55.

Sahu, S. K., D. K. Dey & M. D. Branco (2003) A New Class of Multivariate Skew Distributions With Applications to Bayesian Regression Models, *Canad. J. Statist.*, **31,** 129-150.

Stein, C. (1981) Estimation of the Mean of a Multivariate Normal Distribution, *Annals of Statistics*, **9**, 1135-1151.

Theodossiou, P. (1998) Financial Data and the Skewed Generalized T Distribution, *Management Science*, **44**, 1650-1661.

Treynor, J. L. & K. Mazuy (1966) Can Mutual Funds Outguess the Market? *Harvard Business Review*, **44**, 131-136.

Figure 1 - A Sketch of Conditional Expected Value

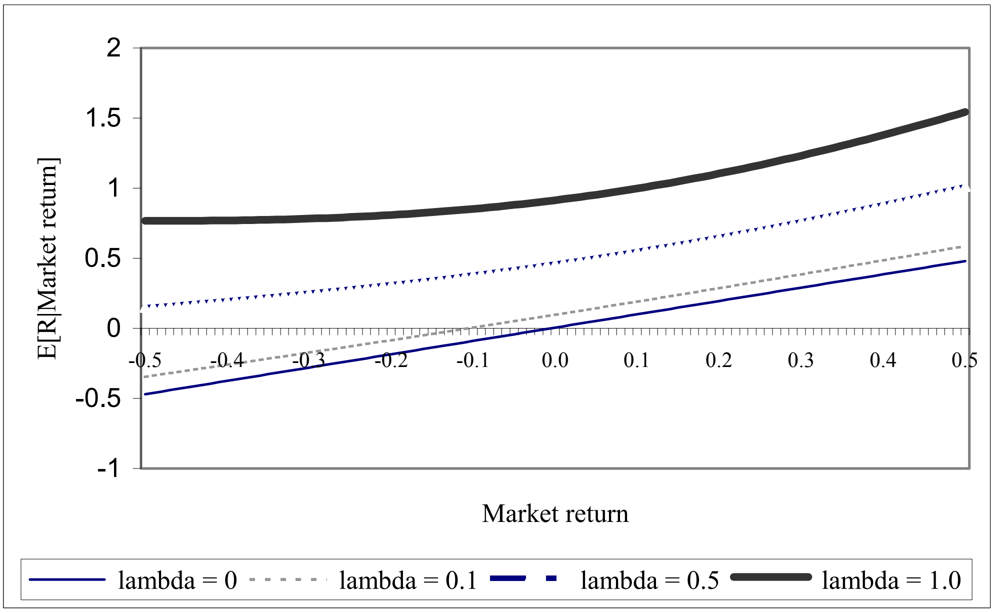

Figure 2 – Sketch of the Mean-Variance-Skewness Efficient Surface