Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation

Javier Mencía  
<javier.mencia@bde.es>  
Banco de España

Enrique Sentana  
<sentana@cemfi.es>  
CEMFI

XXXIII Symposium on Economic Analysis, Zaragoza, December 12th, 2008
Outline of the presentation

1. Introduction
2. Distributional assumptions
3. Portfolio allocation
4. Maximum likelihood estimation
5. Empirical Application
6. Conclusions
Motivation

- Mean-variance analysis remains the most widely used asset allocation framework:
  - It allows a graphical comparison of risk and return,
  - The frontier is spanned by only two funds,
  - It is consistent with elliptical distributions.

- However, it neglects the effect of higher order moments on asset allocation:
  - Chabi-Yo, Leisen and Renault (2007), Harvey et al. (2002), and Patton (2004) have shown the empirical importance of asymmetries in portfolio selection.
  - Athayde and Flôres (2004) obtain the mean-variance-skewness frontier by simulation and derive some of its properties.
  - Briec, Kerstens and Jokung (2007) propose an optimisation algorithm that obtains the “closest” mean-variance-skewness efficient portfolio to a given portfolio.
Contributions of this paper

- We make mean-variance-skewness analysis fully operational by working with location-scale mixtures of normals (LSMN), a family of multivariate asymmetric distributions that nests many important particular cases (Gaussian, Student $t$, Generalised Hyperbolic, finite mixtures of normals,...)
- We derive a standardised multivariate parametrisation, obtain analytical formulas for the score and evaluate the unconditional information matrix by simulation.
- We show that any portfolio is uniquely characterised by its mean, standard deviation and skewness, derive the mean-variance-skewness frontier in closed form and show that its efficient part can be spanned by three funds.
- In an empirical application we find that the Datastream World-ex US index is able to improve the investment opportunity set of US investors if they care about skewness but not if they only seek mean-variance efficiency.
Location-scale mixtures of normals (LSMN)

- Consider the $N \times 1$ vector $u \sim LSMN_N(\alpha, \beta, \Upsilon, \tau)$:

$$u = \alpha + \xi^{-1} \Upsilon \beta + \xi^{-1/2} \Upsilon^{1/2} r,$$

where $\alpha$ and $\beta$ are $N \times 1$ vectors, $\Upsilon$ is a $N \times N$ positive definite matrix, $r \sim N(0, I_N)$, and $\xi \sim F(\cdot; \tau)$ is an independent positive scalar mixing variable.

- $\alpha$ and $\Upsilon$ play the roles of location vector and dispersion matrix, respectively, $\tau$ allows for flexible tail modelling, while $\beta$ introduces skewness.

- This distribution is closed under linear combinations, which can be exploited in risk management applications, such as Value at Risk.

- We need to restrict $\alpha$ and $\Upsilon$ to model standardised residuals in heteroskedastic models, which leaves $\beta$ and $\tau$ as free shape parameters.
Tail dependence in the LSMN family
Bivariate densities and their contour plots

Normal
\[ \xi = 1, \beta = 0 \]

Asymmetric t
\[ \xi \sim \text{Gamma}, \beta = -3 \iota \]

Mixture
\[ \xi \sim \text{Bernoulli}, \beta = -3 \iota \]
Dynamic econometric specifications

- We will analyse investments in a risk-free asset and a set of $N$ risky assets with excess returns $y_t$, such that

$$y_t = \mu_t(\theta) + \Sigma_t^{1/2}(\theta)\varepsilon_t^*,$$

where $E(y_t|I_{t-1}; \theta_0) = \mu_t(\theta_0)$, $V(y_t|I_{t-1}; \theta_0) = \Sigma_t(\theta_0)$ and $\varepsilon_t^*$ is a standardised LSMN conditional on $I_{t-1}$.

- If $\beta \neq 0$, the likelihood will depend on the “square root” of $\Sigma_t(\theta_0)$. To avoid this problem, we parametrise $\beta$ as:

$$\beta_t(\theta, b) = \Sigma_t^{1/2}(\theta)b.$$
Portfolio allocation

- Consider an investor with wealth $A_{t-1}$ at $t - 1$. Then

$$A_t = A_{t-1} \left( 1 + r_t + w_t'y_t \right),$$

where $r_t$ is the risk free rate, and $w_t$ is the vector of allocations to the risky assets.

- Moments for $A_{t-1} = 1$:

$$E_{t-1}(A_t) = \left[ 1 + r_t + w_t'\mu_t(\theta) \right],$$

$$E_{t-1} \left\{ [A_t - E_{t-1}(A_t)]^2 \right\} = w_t'\Sigma_t(\theta)w_t,$$

$$E_{t-1} \left[ (A_t - E_{t-1}(A_t))^3 \right] = \varphi_t(w_t, \theta, b, \tau),$$

$$= \left( s_{1t} + 3s_{2t}s_{3t} \right) \left[ w_t'\Sigma_t(\theta)b \right]^3 + 3s_{2t} \left[ w_t'\Sigma_t(\theta)w_t \right] \left[ w_t'\Sigma_t(\theta)b \right].$$
Proposition 3

Let $\mathbf{y}_t$ be conditionally distributed as a $N \times 1$ LSMN random vector with conditional mean $\mu_t(\theta)$, conditional covariance matrix $\Sigma_t(\theta)$, and shape parameters $\tau$ and $b$.

Then, for any vector $\mathbf{w}_t \in \mathbb{R}^N$ known at $t - 1$, the conditional distribution of $\mathbf{w}_t' \mathbf{y}_t$ can be fully characterised as a function of its mean, variance and skewness parameter $\mathbf{w}_t' \Sigma_t \mathbf{b}$.

Remark

An investor who, ceteris paribus, prefers high expected returns and positive skewness but dislikes high variances would like to invest in portfolios on the mean-variance-skewness frontier.
The mean variance frontier is characterised by

\[ \mu_0 t = \sigma_0 t \sqrt{\mu'_t(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)}. \]

It can be spanned by two assets:
- The risk-free asset, with return \( r_t \),
- and a portfolio with weights proportional to

\[ \Sigma_t^{-1}(\theta) \mu_t(\theta). \]
We maximise the asymmetry for every feasible variance:

$$\max_{w_t \in \mathbb{R}^N} \varphi_t(w_t, \theta, b, \tau) \quad \text{s.t.} \quad w_t' \Sigma_t(\theta) w_t = \sigma_{0t}^2.$$  

(1)

Proposition 4

The portfolio $w^\dagger_t = \sigma_{0t} b / \sqrt{b' \Sigma_t(\theta) b}$ solves (1).

If

$$\frac{s_{2t}}{s_{1t} + 3s_{2t}s_{3t}} \left[ b' \Sigma_t(\theta) b + \frac{s_{2t}}{s_{1t} + 3s_{2t}s_{3t}} \right] < 0$$

there will be a continuum of portfolios $b' \Sigma_t(\theta) w^\ddagger_t = z^\ddagger$ that also solves (1), yielding the same skewness for each $\sigma_{0t}^2$. However, the one with maximum expected return is a linear combination of $\Sigma_t^{-1}(\theta) \mu_t(\theta)$ and $b$. 

We maximise asymmetry for every feasible pair of target expected return and variance

\[
\max_{w_t \in \mathbb{R}^N} \varphi_t(w_t, \theta, b, \tau) \quad \text{s.t.} \quad \begin{cases} 
    w_t' \mu_t(\theta) = \mu_{0t} \\
    w_t' \Sigma_t(\theta) w_t = \sigma_{0t}^2
\end{cases}
\]

The feasibility of the pair \( \mu_{0t} \) and \( \sigma_{0t}^2 \) is ensured by the mean-variance restriction

\[
\sigma_{0t}^2 \geq \frac{\mu_{0t}^2}{\mu_t'(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)}.
\]
Proposition 6

The maximum (and minimum) skewness portfolios for a given mean \( \mu_{0t} \) and variance \( \sigma^2_{0t} \) will be given by either

\[
\mathbf{w}_{1t}^* = \frac{\mu_{0t} + \Delta_t^{-1} \mu_t'(\theta) \mathbf{b}}{\mu_t'(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)} \Sigma_t^{-1}(\theta) \mu_t(\theta) - \frac{1}{\Delta_t} \mathbf{b},
\]

or

\[
\mathbf{w}_{2t}^* = \frac{\mu_{0t} - \Delta_t^{-1} \mu_t'(\theta) \mathbf{b}}{\mu_t'(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)} \Sigma_t^{-1}(\theta) \mu_t(\theta) + \frac{1}{\Delta_t} \mathbf{b},
\]

where

\[
\Delta_t = \sqrt{\left(\mathbf{b}' \Sigma_t(\theta) \mathbf{b}\right) \left(\mu_t'(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)\right) - \left(\mu_t'(\theta) \mathbf{b}\right)^2} \frac{\sigma^2_{0t} \left(\mu_t'(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)\right) - \mu_{0t}^2}{\sigma^2_{0t} \left(\mu_t'(\theta) \Sigma_t^{-1}(\theta) \mu_t(\theta)\right) - \mu_{0t}^2}.
\]
Proposition 6 shows that the mean-variance-skewness frontier can be spanned by three funds:

- The risk-free asset,
- The mean-variance efficient portfolio $\Sigma^{-1}_t(\theta)\mu_t(\theta)$,
- The skewness-variance efficient portfolio $b$.

Interestingly, if $b \propto \Sigma^{-1}_t(\theta)\mu_t(\theta)$, then the mean-variance and skewness-variance frontiers will coincide, so that there will be two fund spanning.
Mean-Variance-Skewness frontier. Proposition 4

Three dimensional representation

Mean vs. Standard Deviation

Mean vs. Asymmetry

Standard Deviation vs. Asymmetry
Mean-Variance-Skewness frontier. Proposition 4 (Contd)

Three dimensional representation

Mean vs. Standard Deviation

Mean vs. Asymmetry

Standard Deviation vs. Asymmetry
The score vector and the information matrix

- We can use EM algorithm - type arguments to obtain analytical formulae for the score function:

\[ s_t(\phi) = \mathbb{E} \left( \frac{\partial l(\mathbf{y}_t | \xi_t, I_{t-1}; \phi)}{\partial \phi} \middle| I_T; \phi \right) + \mathbb{E} \left( \frac{\partial l(\xi_t | I_{t-1}; \phi)}{\partial \phi} \middle| I_T; \phi \right) \]

- We derive explicit formulas for the GH distribution.

- For pure time series models, we propose the evaluation of the information matrix by simulating a long time series from the estimated D.G.P.

\[ \hat{\mathcal{I}}_{T_s}(\hat{\phi}_T) = \frac{1}{T_s} \sum_{t_s=1}^{T_s} s_{t_s}(\hat{\phi}_T)s'_{t_s}(\hat{\phi}_T), \]

- We assess the accuracy of \( T_s = 100,000 \) in a Monte Carlo exercise.
Data

- We consider the ten Datastream main sectoral indices for the US.
- We estimate conditional mean-variance-skewness frontiers for these indices and assess the diversification gains of adding the Datastream World ex-US index.
- Our database consists of daily data from January 4th, 1988 to October 12th, 2007 (4971 observations).
- Prior to estimation, we compute excess returns with respect to the Eurodollar overnight interest rate.
- We model the mean with a diagonal VAR(1) with drift and the covariance dynamics with a conditionally heteroskedastic single factor model
  \[ \Sigma_t(\theta) = cc'\lambda_t + \text{diag}(\gamma_t), \]
  in which both \( \lambda_t \) and the elements of \( \gamma_t \) follow GQARCH(1,1) processes.
## Distributional tests

### Normality

<table>
<thead>
<tr>
<th>Score based</th>
<th>Test</th>
<th>p-value</th>
<th>Asymptotic</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>9289.32</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Skewness</td>
<td>204.34</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Sup-LM</td>
<td>9493.66</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Kuhn-Tucker</td>
<td>9493.66</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LR</th>
<th>Asymptotic</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$: sym. GH</td>
<td>2246.63</td>
<td>0.000</td>
</tr>
<tr>
<td>$H_1$: asym. $t$</td>
<td>2270.09</td>
<td>0.000</td>
</tr>
<tr>
<td>$H_1$: asym. GH</td>
<td>2270.09</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Source: Mencía and Sentana (2008), Distributional tests in multivariate dynamic models with Normal and Student $t$ innovations
### Distributional tests

**Student $t$**

<table>
<thead>
<tr>
<th>Score based</th>
<th>Test</th>
<th>p-value</th>
<th>Asymptotic</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurtosis</td>
<td>0.00</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>25.35</td>
<td>0.005</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Joint</td>
<td>25.35</td>
<td>0.006</td>
<td>0.007</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LR</th>
<th>Asymptotic</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$: sym. GH</td>
<td>0.00</td>
<td>1.000</td>
</tr>
<tr>
<td>$H_1$: asym. $t$</td>
<td>23.45</td>
<td>0.012</td>
</tr>
<tr>
<td>$H_1$: asym. GH</td>
<td>23.45</td>
<td>0.012</td>
</tr>
</tbody>
</table>

Source: Mencía and Sentana (2008), Distributional tests in multivariate dynamic models with Normal and Student $t$ innovations
Spanning tests

- From the point of view of a US investor, an interesting question is whether investing outside the US can enlarge the mean-variance-skewness frontier.

- This test has two parts:
  - Mean-variance spanning: if the regression of the World ex. US index on the ten US sectoral returns has zero intercept, then this asset will not modify the mean-variance frontier.
  - Skewness variance spanning: the World ex. US index will only expand the skewness-variance frontier if its skewness parameter $b$ is different from zero when estimated jointly with the US series.
Spanning tests

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>Wald</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value</td>
<td>p-value</td>
</tr>
<tr>
<td>Mean-variance efficiency</td>
<td>1.00</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>0.317</td>
<td>0.306</td>
</tr>
<tr>
<td>Skewness-variance efficiency</td>
<td>9.64</td>
<td>9.79</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Joint</td>
<td>13.57</td>
<td>13.72</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

- Thus, a US investor that only cares about mean-variance efficiency will not obtain any gains from investing outside the US.
- In contrast, a US investor who dislikes negative skewness would find significant gains from investing overseas.
Mean-Variance-Skewness frontier

Three dimensional representation

Mean vs. Standard Deviation

Mean vs. Asymmetry

Standard Deviation vs. Asymmetry

Javier Mencía  Enrique Sentana  LSMN and mean-variance-skewness analysis
Conclusions

- We make mean-variance-skewness analysis fully operational by working with the flexible family of location-scale mixtures of normals.
- Under the LSMN assumption, the distribution of excess returns on any given portfolio depends only on its mean, standard deviation and skewness.
- We describe the mean-variance-skewness frontier in closed form and show that its efficient part can be spanned by combining the risk-free asset, a mean-variance efficient portfolio and a skewness-variance efficient portfolio.
- We study maximum likelihood estimation of our model parameters.
- We carry out an empirical application to the ten Datastream US sectoral indices, where we also compute mean-variance and asymmetry-variance spanning tests.
- We find that the Datastream World ex. US index does not modify the original US mean-variance frontier, but it provides a significant improvement on the skewness-variance frontier.
Directions for future research

- An interesting avenue for future research would be to analyse the pricing implications of our model, by relating our framework to the extensions of the CAPM based on the first three moments of returns (Barone-Adesi, 1985; Chabi-Yo, Leisen and Renault (2007); Kraus and Litzenberger, 1976; and Lim, 1989).

- It would also be useful to explore the implications of our model at different horizons. In this sense, we could exploit particular examples of the LSMN, such as the Variance Gamma model (Madan and Milne, 1991; Madan, Carr and Chang, 1998), which are closed under temporal aggregation.

- Finally, it would be interesting to check that our empirical results are robust to replacing the GH assumption by a non-parametric distribution for the mixing variable $\xi_t$, or a dynamic model for this variable.