Non-Gaussian Dependence using Lévy Processes

Dilip B. Madan
Robert H. Smith School of Business

Tata Institute of Fundamental Research

International Center for Theoretical Sciences

School in Mathematical Finance
January 16-27 2012, Mumbai, India

Motivation

- A number of applications in financial modelling call for the description of the joint law of asset returns over some horizon of interest.

- For many of these applications it is well recognized that the marginal return distributions of each asset return taken individually is not Gaussian (Jondeau, Poon, and Rockinger (2007), Menn, Fabozzi and Rachev (2005), McNeil, Frey and Embrechts (2005), Boyarchenko and Levendorskii (2002)).

- Hence the focus on non-Gaussian multivariate return distributions.

Applications

- Applications include the design of optimal portfolios where the interest is in the physical multivariate return distribution,

- the pricing of options on a basket of stocks for which the relevant return distribution is risk neutral.

- The marginal distributions reflect varying degrees of skewness and excess kurtosis, features that may be inherited and even exaggerated in portfolios.

- We investigate and report on the comparative performance of three tractable multivariate models for asset returns that have recently appeared in the literature.

Tractability and Alternatives

- The particular feature of tractability for the three chosen models is the ability to estimate the models in dimensions as high as 50, by reduction to a suitable sequence of univariate estimation problems.

- There are a variety of multivariate elliptical distributions like the multivariate $t - distribution$ Kotz and Nadarajah (2004), or the multivariate variance gamma (Madan and Seneta (1990), Schoutens and Cariboni (2009) that impose a uniform tail structure across the different assets that we do not study here. The methods developed could however be applied to these models as well.

Model FGC

- The first of these models is a full rank Gaussian copula (FGC) that has been proposed and studied by Malevergne and Sornette (2005).

- In this model each asset return is a nonlinear transform of a set of standard normal variates that are correlated with a correlation matrix C of possibly full rank.

Model LML

- The second model follows the idea implicit in a multivariate normal model where all the variables are linear transformations of independent Gaussian variates.

- We now consider a linear mixture of independent but non-Gaussian variates, that like the Gaussian variable, are infinitely divisible and associated with the unit time distribution of a Lévy process.

- The model was implemented for portfolio design in asset allocation by Madan and Yen (2008) using independent components analysis (ICA, Hyvärinen, Karhunen and Oja (2001)) to identify the independent variables.

- It was also used by Madan (2006) in an equilibrium asset pricing model. We denote this model LML for Lévy mixture of independent Lévy.

Model VGC

- The third model writes the marginals as following the variance gamma law (Madan and Seneta (1990), Madan, Carr and Chang (1998)) . The marginals are gamma time changed Brownian motion at unit time and we correlate the Brownian motions.

- The model was proposed by Eberlein and Madan (2008) and employed by Madan (2009) in a study pricing options on a basket of stocks. We term this model VGC for correlated variance gamma.

Choice of Data

- In order to investigate models of dependence it is helpful to consider data where there is some presumption of the presence of dependencies.

- Though this is expected of stock returns in general as they presumably share exposure to common macro movements of the economies in which they trade, one would expect such dependencies to be even greater for sector specific exchange traded funds ($ETF's$) that constitute diversified portfolios of similar collections of stocks.

- Additionally we have daily data on the market values of these funds, thereby providing us with a fertile environment in which to test our models of multivariate dependence.

- With these considerations in mind the three models are estimated on a number of ETF returns partitioned by economic sector, as well as one set that selects a single ETF from each of nine sectors.

Model Evaluation Methods

- The question then arises as to how one may evaluate model performance on this data. For many applications one is interested in the return on portfolios and so we ask how well the models explain the univariate distribution of arbitrary portfolios.

- For this evaluation we construct a thousand arbitrary randomly generated long short portfolio returns on the unit sphere of dimension matching the number of $ETF\prime s$.

- We construct both the actual portfolio return in our data and the distribution of this return as predicted by each of our three estimated models.

- We then construct the $p - value$ on a chi-square test for whether the actual return comes from each of the three models in turn.

- Finally we graph the proportion of portfolios with a $p - value$ greater than x for a range of x values.

- A model with a higher proportion of high $p - values$ for each candidate probability level does a better job in explaining the univariate laws of arbitrary portfolios and is therefore a superior model for the data set in question.

Enquiry into Model Structure

- Different models appear to dominate on different occasions. For example, within sector one gets a better performance from LML while across sectors FGC and VGC dominate LML.

- These observations lead us to enquire deeper into the structure of dependence in the different models.

- We follow the ideas of Longin and Solnik (2001) related to extreme correlation and localize further.

Local Correlation Concept

- For this purpose, we develop the concept of local correlation and observe that for the LML model there is greater correlation in the tails of the distribution than in the center while for FGC and VGC correlation drops of in the tails.

- We call the map of local correlation the correlation signature of the model and we present the correlation signatures for our three models as it is estimated for the energy sector and cross sector grouping.

- A richer understanding of correlations is being called for in recent researches and we note in this regard Embrechts (2009).

The Models Studied

FGC

- From one perspective it is uninformative to compute correlations of non-Gaussian variates as the result does not lead us to any ability at writing down the joint probability law.

- We merely have correlation estimates and plenty of them if the dimension is high, but there we stop.

- However, if the data are transformed to standard normal variates first, before the correlation is computed then the computed correlations may be used to write down the joint multivariate normal law of these transformed Gaussian variates with the original data being a non-linear transform thereof.

- The result is joint multivariate probably element.

FGC Details

- Let $X = (X_1, ..., X_N)$ be a vector of dimension N with marginal distributions for each X_i given by

$$P(X_i \leq x) = F_i(x).$$

- One may transform the marginal laws to standard normal variates by

$$Z_i = \Phi^{-1}(F_i(X_i)),$$

where Φ is the standard normal distribution function.

- By construction Z_i is a standard normal variate and one may recover X_i as

$$X_i = F_i^{-1}(\Phi(Z_i)).$$

- It is supposed that the vector $Z = (Z_1, \cdots, Z_N)$ is standard multivariate normal with correlation matrix C.

- The joint probably density of X may be expressed in terms of the multivariate normal density for Z by a simple change of variable.

- In our application we shall take the marginal distributions F_i to come from the variance gamma class of distributions.

LML Details

- The Lévy mixture model postulates that

$$X = AY,$$

for a mixing matrix A with each variable Y_i being independent of $(Y_j,\ j \neq i)$. We further suppose that each Y_j has a variance gamma distribution.

- Given characteristic functions

$$\phi_j(u) = E\left[\exp\left(iuY_j\right)\right]$$

the joint characteristic function of X may be easily derived as

$$\phi_X(u) = \prod_{j=1}^{N} \phi_j\left((A'u)_j\right).$$

VGC Details

- The marginal distributions are here postulated to be in the centered variance gamma class with

$$X_i = \theta_i (g_i - 1) + \sigma_i \sqrt{g_i} Z_i,$$

  where Z_i are standard normal variates and the $g_i's$ are a sequence of independent gamma variates with unit mean and variance ν_i.

- In the VGC specification we now further suppose that Z is multivariate normal with correlation matrix C.

- The joint probability density and characteristic functions are not available in closed form as one has to integrate out a large number of independent gamma densities but they appear as products of square roots that do not separate out in either the density or the characteristic function.

- The joint law, however, is easily simulated from a multivariate normal simulation coupled with drawings from gamma densities.

Comparative remarks on the three models

- The model FGC creates dependence by taking a nonlinear transform of correlated Gaussian variates.

- On the other hand in VGC the transformation is linear as seen in equation but both the intercept and slope are stochastic but simultaneously generated by a single gamma variate. Hence here we have a stochastic linear transformation of correlated Gaussian variates.

- In the model LML Gaussian variates do not appear at all, as we now take a multivariate linear transform of independent non-Gaussian variates.

- The three models create dependence in apparently quite different ways.

- We employ the variance gamma model for our univariate model here, but one could easily extend to the case of the generalized hyperbolic distribution (Eberlein (2001)) or its numerous special cases.

Estimation Procedures

- We suppose we have data $X_t = (X_{1t}, \cdots, X_{Nt})$ for $t = 1, \cdots T$ of independent draws from the relevant distributions.

- We suppose this data has been centered to a zero sample mean.

FGC

- For FGC one first estimates the marginal distribution functions on the univariate data and we employ distributions in the variance gamma class for this purpose.

- This gives us a matrix of marginal VG parameters

$$\sigma_i, \nu_i, \theta_i, i = 1, \cdots, N.$$

- We then form the univariate data

$$Z_{it} = \Phi^{-1}(F_{VG}(X_{it}; \sigma_i, \nu_i, \theta_i))$$

and we then estimate the correlation matrix C by

$$C_{jk} = \frac{1}{T} \sum_{t=1}^{T} Z_{jt} Z_{kt}.$$

- We then simulate $10,000$ readings from a multivariate normal density $Z_s = (Z_{1s}, \cdots Z_{Ns})$, $s = 1, \cdots, 10,000$, with this correlation matrix and generate simulated readings

$$X_{js} = F_{VG}^{-1}\left(\Phi\left(Z_{js}\right), \sigma_j, \nu_j, \theta_j\right).$$

This matrix of simulated draws from the estimated model will be used subsequently in our analysis of model quality.

LML

- For the LML model we first identify the mixing matrix following Madan and Yen (2008) and employ independent components analysis for this purpose.

- The hypothesis of independent components analysis is precisely the statement that one is observing a linear mixture of independent variates and this procedure first performs a principal components analysis (PCA) to generate a set of unit variance orthogonal random variables constructed as linear combinations of the original observed variables.

- It is then observed that an equivalent PCA is obtained on multiplication by any rotation matrix.

- The procedure is based on recognizing that a mixing of non-Gaussian signals induces a convergence to Gaussianity and hence the path back to the original signals amounts to maximizing a metric of non-Gaussianity.

- Such a criterion is employed to search over the class of rotation matrices to construct the matrix A that is the product of the matrix delivering the PCA followed by the non-Gaussianity maximizing rotation matrix.

- The specific criterion used is the maximization of the expected logarithm of the hyperbolic cosine (Hyvärinen (1999)).

- Once the matrix A has been identified, one obtains data on the independent components on premultiplication of the observed data matrix by the inverse of A.

- We further postulate that these independent components are variance gamma random variables and we estimate the parameters of the variance gamma model on the data for these components.

- We then generate a simulated matrix of 10000 draws from this estimated dependence model. First we generate an $N \times 1$ vector of independent variance gamma random variables 10000 times from the N estimated variance gamma laws.

- We then multiply each $N \times 1$ vector by the $N \times N$ matrix A to sample a draw of N observations from this probability law.

- The result is a matrix of N by 10000 readings from the LML dependence model.

VGC

- For VGC we employ the same VG marginal laws estimated in FGC and then infer the correlations between the Gaussian variates from the observed matrix of covariances between observed returns.

- This procedure inflates Gaussian component correlations relative observed correlations by a factor of decorrelation induced by the independent gamma time changes that depends on just the marginal laws.

- This inflation factor is explicitly described in Eberlein and Madan (2008). On occasion these inflation factors can lead to an estimated correlation matrix with some entries above unity.

- In this case we construct the closest correlation to our symmetric matrix using the procedures of Qi and Sun (2006).

- We then generate 10000 readings from this law by generating correlated Gaussian random variables and independent gamma variates to form a reading on an N vector.

- The result is a N by 10000 matrix of draws from the VGC law.

Investigating model quality

- We have estimated three joint laws on asset returns in dimensions ranging from 3 to 7.

- It is of interest to enquire into the quality of the estimated models, or their ability to describe the data.

- We do not have available in closed form the relevant joint densities and hence we cannot compute likelihoods and the models are not nested in any case.

- We also do not have estimates of asymptotic distributions of parameter estimates or likelihoods and hence cannot employ the procedures of non-nested tests either.

- We consider a performance based evaluation as opposed to testing whether the data comes from the proposed model.

- In fact these are a tractable class of models available to us and the data may well not come from any of them as the modeling of multidimensional financial return data is a fairly complex exercise.

- We enquire instead into how well these models of dependence explain the univariate laws of randomly chosen linear mixtures.

- A multivariate model that explains well all linear mixtures is clearly a good candidate model for the joint law.

The Specific Procedure

- With such a performance evaluation in mind, we randomly generate 1000 linear combinations with coefficients located on the unit sphere of N dimensional space.

- For each linear combination we construct readings on this linear combination in the data and for each of our three models with 10000 simulated paths we construct 10000 simulated readings for the same linear combination.

- The simulated readings are employed to construct the expected number of observations in 20 equally spaced cells covering the interquantile range from 5 to 95 percent.

- We also obtain the observed number of readings in each of these cells for the same linear combination applied to the observations in the data.

- For each linear combination we construct a chi-square test $p - value$ based on the observed and expected number of readings in each of the 20 cells.

- We then graph against x, a candidate $p - value$ between zero and unity, the proportion of portfolios with an observed $p - value$ above x.

- There are three such graphs for each of our three models.

- A model whose graph dominates that for another model clearly has a higher proportion of portfolios with high p-values than the dominated model and hence provides us with a superior explanation of the univariate laws of arbitrary linear combinations.

- The dominating model therefore constitutes a better candidate model of dependence for this data.

Local Correlation

- With a view towards taking a deeper look at how dependence is modeled in an arbitrary joint density we consider the formulation of local correlation in the neighbourhood of an arbitrary point in space.

- For this purpose we consider an arbitrary joint density for two random variables $q(x, y)$.

- We now define

$$h(x, y) = -2 \ln q(x, y).$$

- Consider expanding the function h to second order around the point (a, b) to obtain

$$
\begin{aligned}
h(x,y) \approx\ & h(a,b) + h_a(x-a) + h_b(y-b) \\
& + \frac{1}{2} h_{aa}(x-a)^2 + \frac{1}{2} h_{bb}(y-b)^2 \\
& + h_{ab}(x-a)(y-b).
\end{aligned}
$$

- A bivariate normal distribution has such an exact quadratic expression for the log likelihood where we identify

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} h_{aa} & h_{ab} \\ h_{ba} & h_{bb} \end{pmatrix}$$

- In this case we would have

$$\mathbf{\Sigma} = \frac{1}{h_{aa}h_{bb} - h_{ab}^2} \begin{pmatrix} h_{bb} & -h_{ab} \\ -h_{ba} & h_{aa} \end{pmatrix}$$

  and the correlation would be

$$\rho_{ab} = \frac{-h_{ab}}{\sqrt{h_{aa}h_{bb}}}.$$

- We might consider defining this value generally as the local correlation.

- It is then a question as to whether this value is between -1 and 1. For this we require the square to be less than one or

$$h_{ab}^2 \leq h_{aa}h_{bb}.$$

- This is precisely the condition for the negative of the log likelihood to be a convex function.

- We must have $h_{aa}, h_{bb} > 0$ along with the condition for absolute local correlation below unity.

- For many models with unimodal joint densities we have the concavity of the density near the mode and hence in this region the negative of the log likelihood is convex.

- The proposed definition for local correlation will yield magnitudes dominated by unity in absolute value in this region.

- There is therefore quite generally a local domain in which one may investigate the shape of local correlation.

- We call this map in this local domain the correlation signature of the model.

- Additionally there are models with universally log concave densities and for these models the local correlation is universally well defined.

- This is an important class of densities, much studied in its own right Barlow and Proschan (1981), Prèkopa (1973).

- We now investigate the nature of this local correlation surface for our three models.

- Consider first FGC. Here our joint law is that of nonlinear transforms of correlated Gaussians. The joint density is now

$$f(x, y) = b(g(x), h(y))g'(x)h'(y)$$

where $b(z_1, z_2)$ is a bivariate normal density.

- The negative of twice the log density

$$\begin{aligned} \Phi(x, y) &= -2\ln(f(x, y)) \\ &= -2\ln b(g(x), h(y)) - 2\ln g'(x) - 2\ln h'(y) \end{aligned}$$

- Define $\widetilde{b}(z_1, z_2) = -2 \ln b(z_1, z_2)$. It follows that

$$\Phi_x = \widetilde{b}_x g'(x)$$

and

$$\Phi_{xy} = \widetilde{b}_{xy} g'(x) h'(y)$$

- We have $\widetilde{b}_{xy} = \rho$ a constant and g', h' are high in the center and low in the tails.

- Hence this model gives correlation in the neck and lower correlation in the tails.

- If we consider the VGC structure we have

$$\begin{aligned} x &= \theta_x(g_x - 1) + \sigma_x \sqrt{g_x} Z_x \\ y &= \theta_y(g_y - 1) + \sigma_y \sqrt{g_y} Z_y \end{aligned}$$

- The joint density is now

$$E\left[\frac{1}{\sigma_x\sqrt{g_x}} b\left(\frac{x-\theta(g_x-1)}{\sigma_x\sqrt{g_x}}, \frac{y-\theta(g_y-1)}{\sigma_y\sqrt{g_y}}\right)\frac{1}{\sigma_y\sqrt{g_y}}\right]$$

- The critical function now is

$$\Phi(x,y) = -2\ln E\left[\begin{array}{c}\frac{1}{\sigma_x\sqrt{g_x}}\times \\ b\left(\frac{x-\theta(g_x-1)}{\sigma_x\sqrt{g_x}},\frac{y-\theta(g_y-1)}{\sigma_y\sqrt{g_y}}\right)\times \\ \frac{1}{\sigma_y\sqrt{g_y}}\end{array}\right]$$

and we may consider in its place the expectation of the log or

$$\widetilde{\Phi}(x,y) = -2E\left[\ln\left(\begin{array}{c}\frac{1}{\sigma_x\sqrt{g_x}}\times \\ b\left(\frac{x-\theta(g_x-1)}{\sigma_x\sqrt{g_x}},\frac{y-\theta(g_y-1)}{\sigma_y\sqrt{g_y}}\right)\times \\ \frac{1}{\sigma_y\sqrt{g_y}}\end{array}\right)\right]$$

and this again gives central correlation as opposed to tail correlations for reasons comparable to the FGC model.

- Finally we consider LML with joint density

$$f(x, y) = g(ax + by)h(cx + dy)\kappa$$

- In this case we get

$$\Phi(x, y) = -2 \ln g(ax+by) - 2 \ln h(cx+dy) - 2 \ln \kappa$$

- If we compute the cross partial we get

$$\begin{aligned}
\Phi_x &= \tilde{g}' a + \tilde{h}' c \\
\Phi_{xy} &= \tilde{g}'' ab + \tilde{h}'' cd \\
\Phi_{xx} &= \tilde{g}'' a^2 + \tilde{h}'' c^2 \\
\Phi_{yy} &= \tilde{g}'' b^2 + \tilde{h}'' d^2
\end{aligned}$$

- For our convexity condition we require that

$$\Phi_{xx}\Phi_{yy} \geq \Phi_{xy}^2$$

- or equivalently that

$$\left(\tilde{g}''a^2 + \tilde{h}''c^2\right)\left(\tilde{g}''b^2 + \tilde{h}''d^2\right) \geq \left(\tilde{g}''ab + \tilde{h}''cd\right)^2$$

- and this yields the condition

$$\tilde{g}''\ \tilde{h}''\left(a^2d^2 + b^2c^2 - 2abcd\right) \geq 0$$

- or

$$\tilde{g}''\ \tilde{h}''\,(ad - bc)^2 \geq 0$$

- and hence we just need that

$$\tilde{g}''\ \tilde{h}'' \geq 0.$$

- This condition is satisfied if the marginal laws are themselves log convex.

- We note that the second derivatives pick up in the tails and the center and are small in the middle.

- This model leads to tail and central correlations with flat correlations in the middle .

- The LML structure is a fundamentally different model with respect to the associated correlation surfaces.

The Data Employed

- We obtained data on the time series of Exchange Traded Funds (henceforth ETF) that follow various sectors of the US economy.

- As we are also interested in risk neutral laws we focused attention of funds that also have options trading on the ETF and for which we had a time series exceeding 700 days of daily data ending on July 21 2009. There are nine industry groups and the $ETF's$ in the group are displayed in Table 1.

TABLE 1

Sectors	ETF Tickers
Cons. Disc.	xly,rth,xrt,itb,xhb
Energy	xle,iye,ieo,oih,xop
Financials	xlf,iyf,iai,kbe,kre,rkh,kce
Health and Pharm.	xlv,bbh,pph
Ind. and Tech.	xli,iyt,iyw,xlk
Int. Netw. Semicond. Softw.	hhh, bdh, igw, smh, swh
Mat., RE, Telecomm.	xlb, iyr, iyz, tth
Natural Resources	ige, gdx, slx, xme
Utilities	idm, xlu, uth
Cross Sectors	xly,xlp,xle,xlf,xlv,xli,xlk,xlb,xlu

Results

- For each of these nine groups and each of three models we present nine graphs with three curves each for FGC in black, LML in blue and VGC in red displaying the proportion of a thousand random linear combinations with a univariate law explained by the model with a p-value exceeding a candidate value given by the x axis.

- Figure (1) displays the result for the Consumer Discretionary sector where FGC and VGC perform equally well and dominate LML.

- In Figure (2) we have the results for the Energy sector where LM dominates followed by VGC and the FGC.

- For the Financial sector, Figure (3), that saw a lot of movement, VGC dominates by far the other two models.

- In the Health related sector Figure (4) LM and VGC criss cross and dominate FGC.

- For the industrial sector, Figure (5) all three models are equivalent.

- The technology sector Figure (6) like Energy has LML dominating followed by VGC and FGC.

- All three models are equivalent for Natural Resources Figure (7).

- Telecom, Figure (8) sees the order VGC followed by LML and FGC.

- Finally the Utility sector Figure (9) has LM followed by VGC and FGC.

- We observe from focusing in some cases around the 10% point that in five of the nine groups we have LML dominating VGC that dominates FGC.

- In a further two cases all three models are equivalent.

- In one case, Financials, VGC dominates the other two by far.

- There is some broad preference for LML followed by VGC and then FGC.

- We next consider the cross sector group Figure (10) with one ETF from each of the nine sectors.

- In this grouping we have a clear domination by VGC over FGC and LML that are somewhat equivalent.

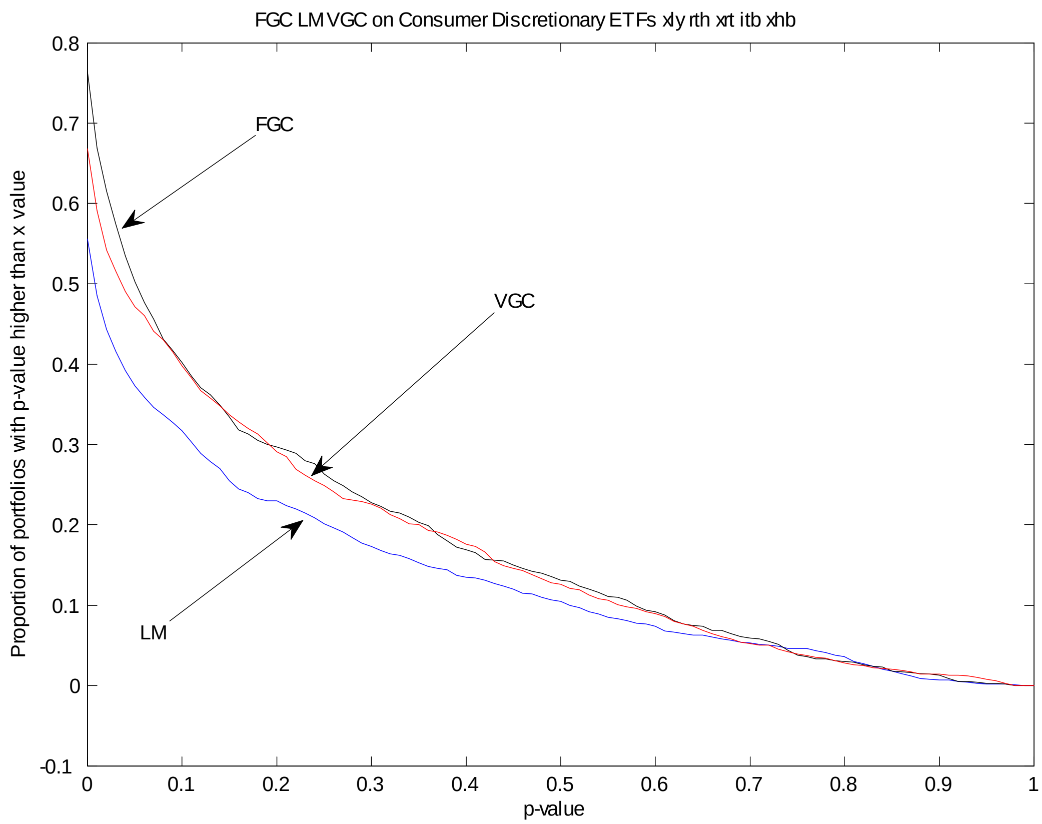

Figure 1: Portfolio Proportions with given probability values in the Consumer Discretionary sector.

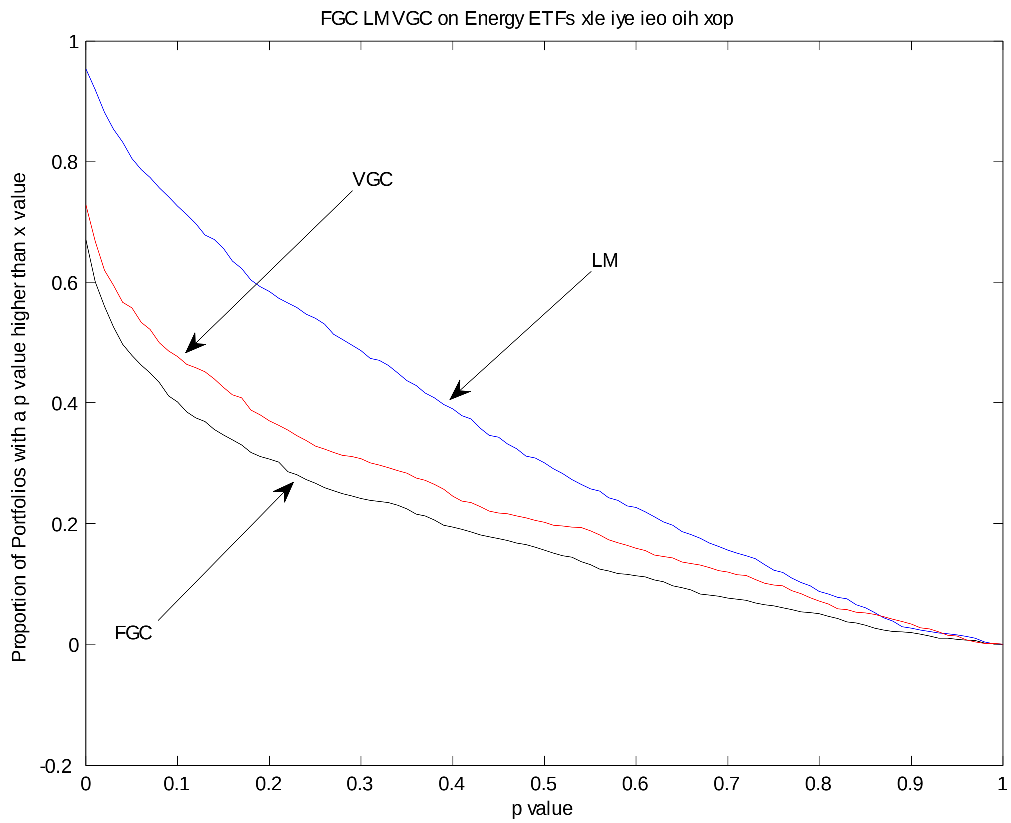

Figure 2: Portfolio Proportions with given probability values in the Energy sector.

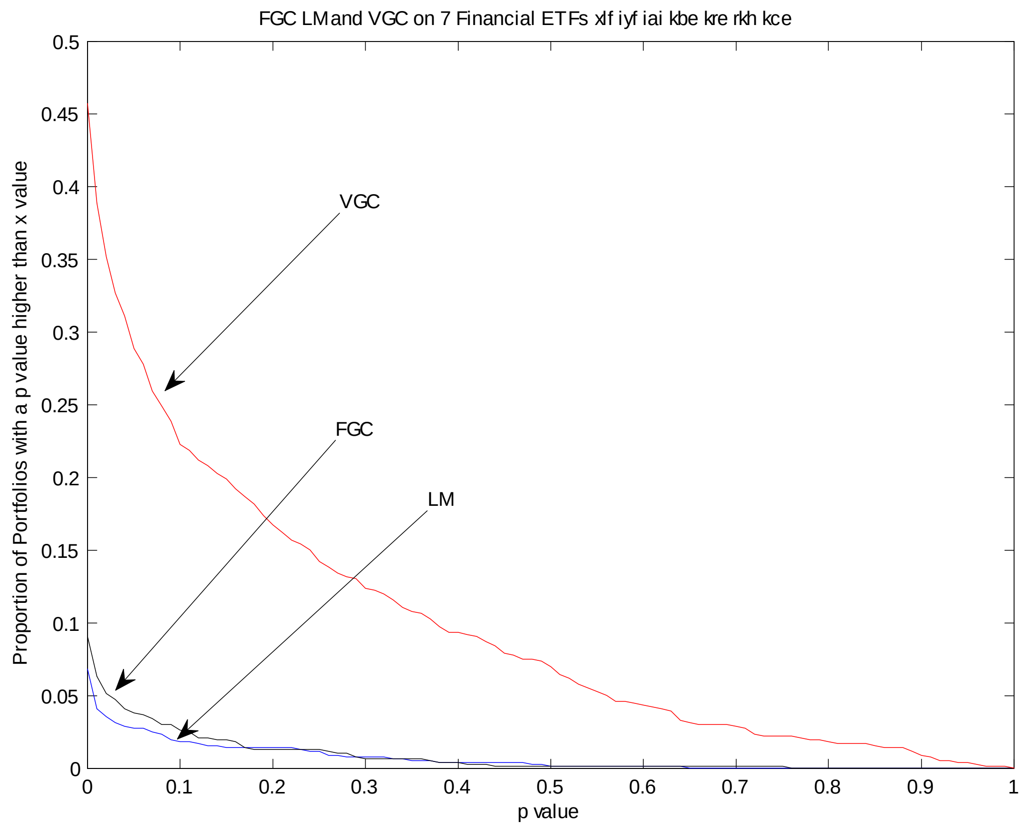

Figure 3: Portfolio Proportions with given probability values in the Financial sector.

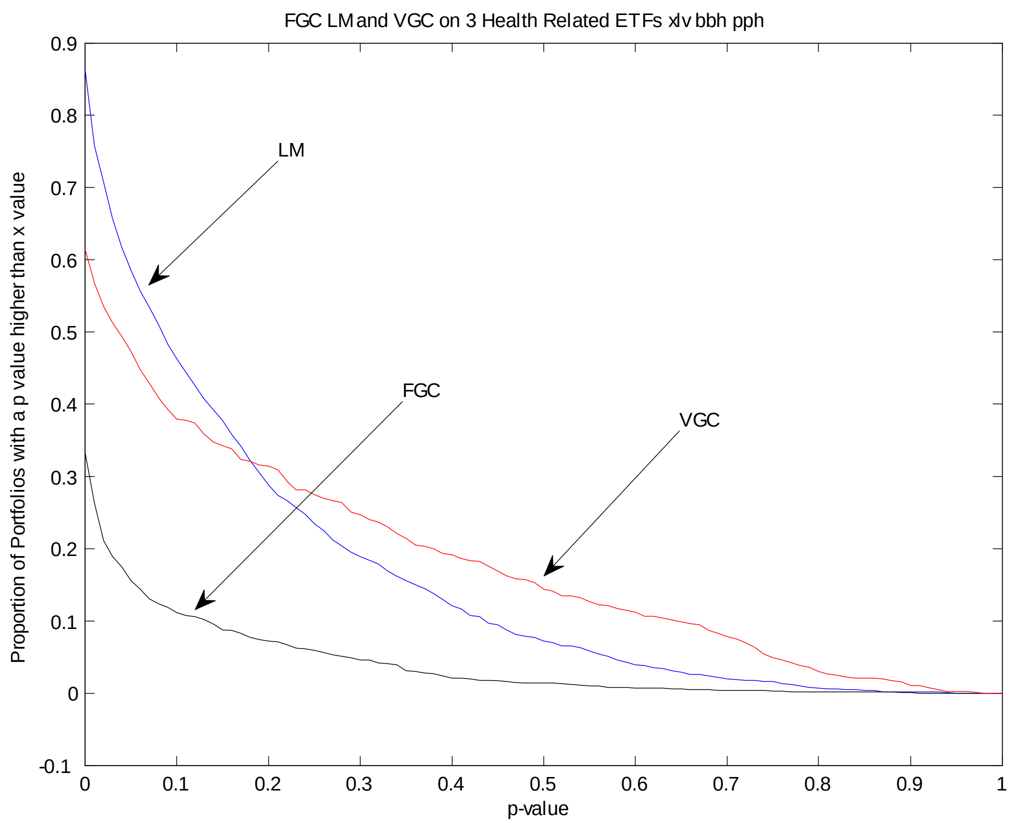

Figure 4: Portfolio Proportions with given probability values in the Health related sector.

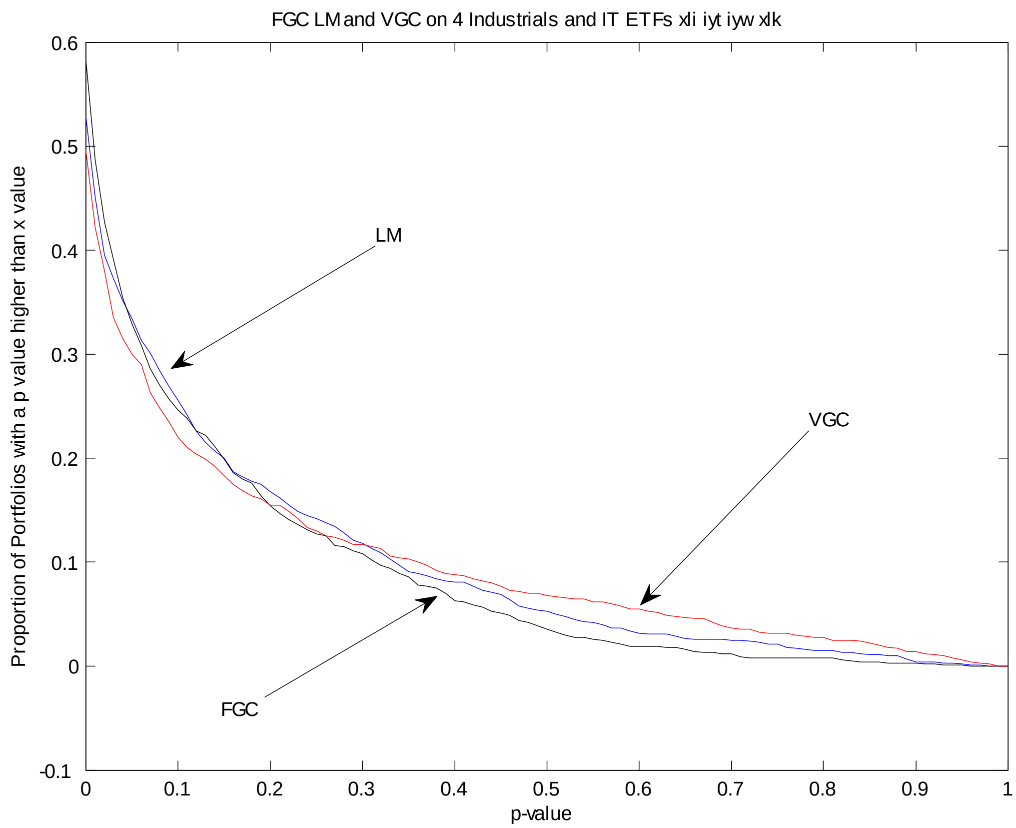

Figure 5: Portfolio Proportions with given probability values in the Industrial sector.

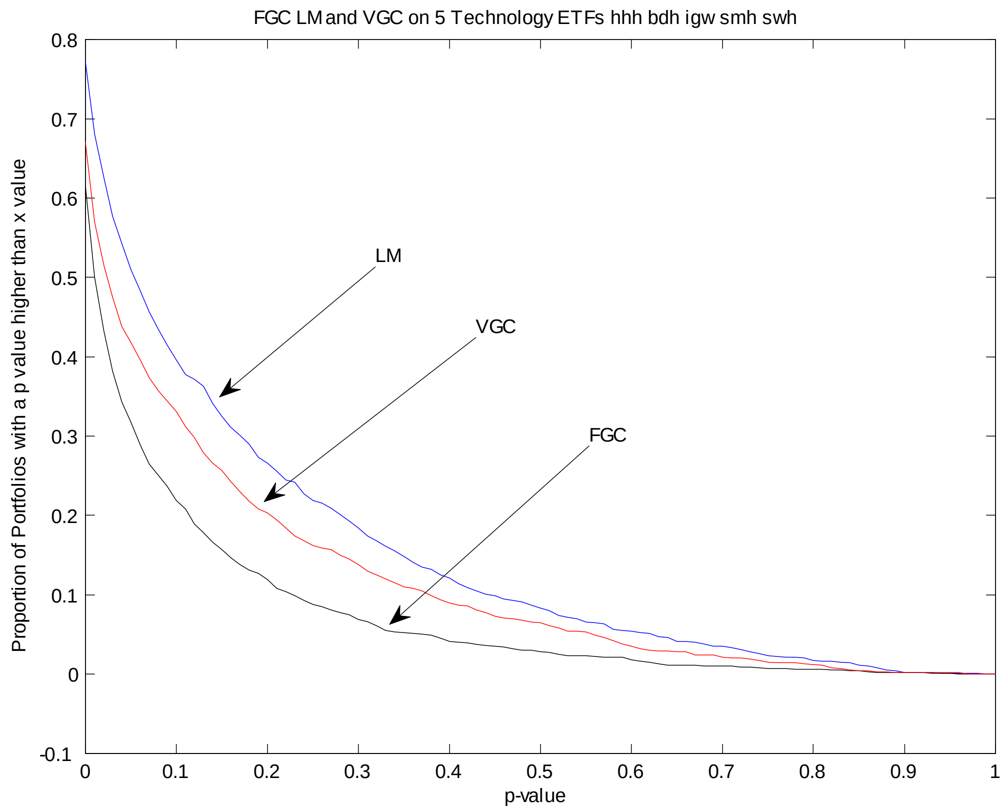

Figure 6: Portfolio Proportions with given probability values in the Technology sector.

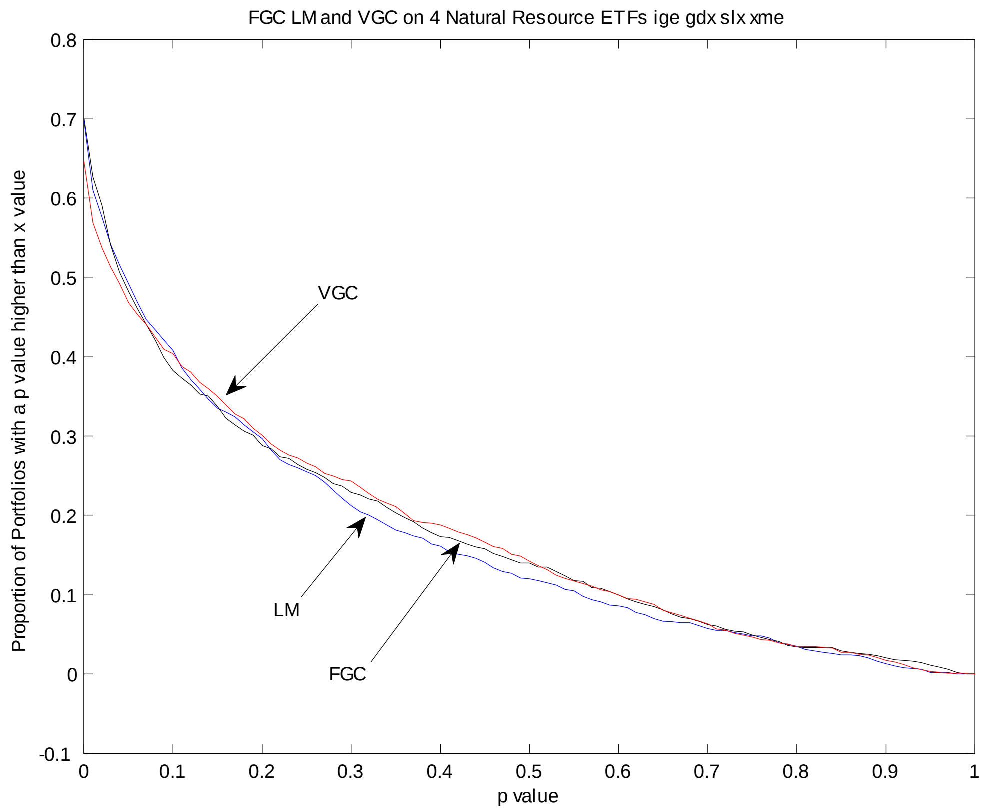

Figure 7: Portfolio Proportions with given probability values in the Natural Resource sector.

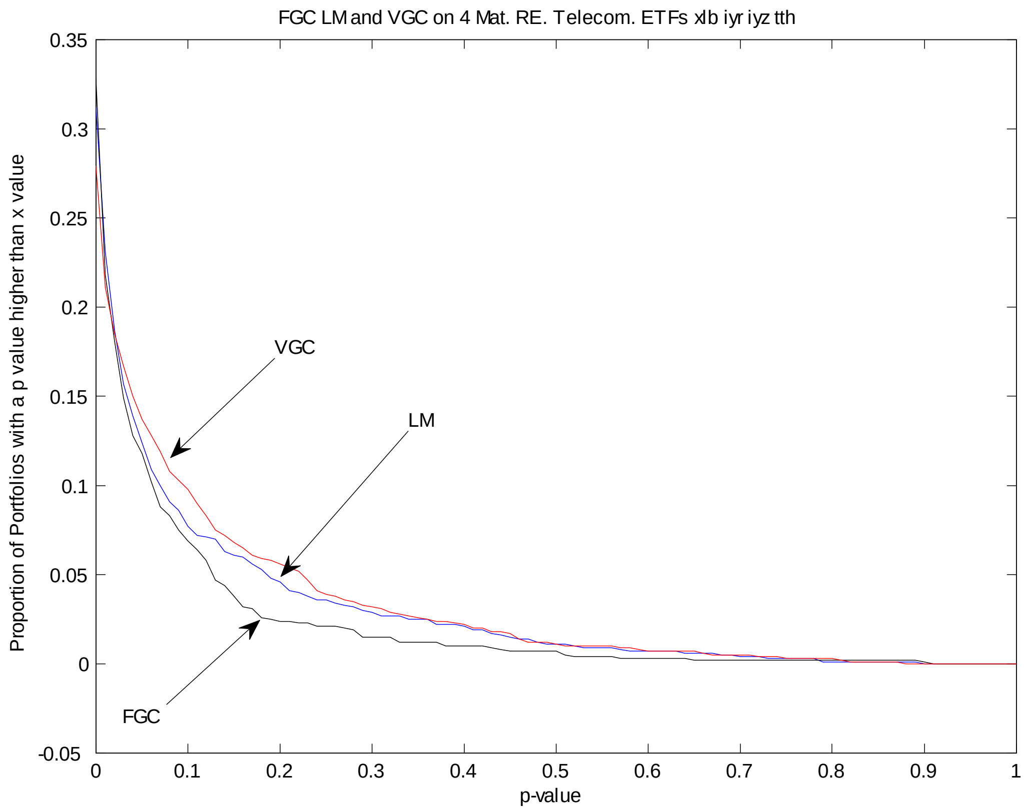

Figure 8: Portfolio Proportions with given probability values in the Telecom sector.

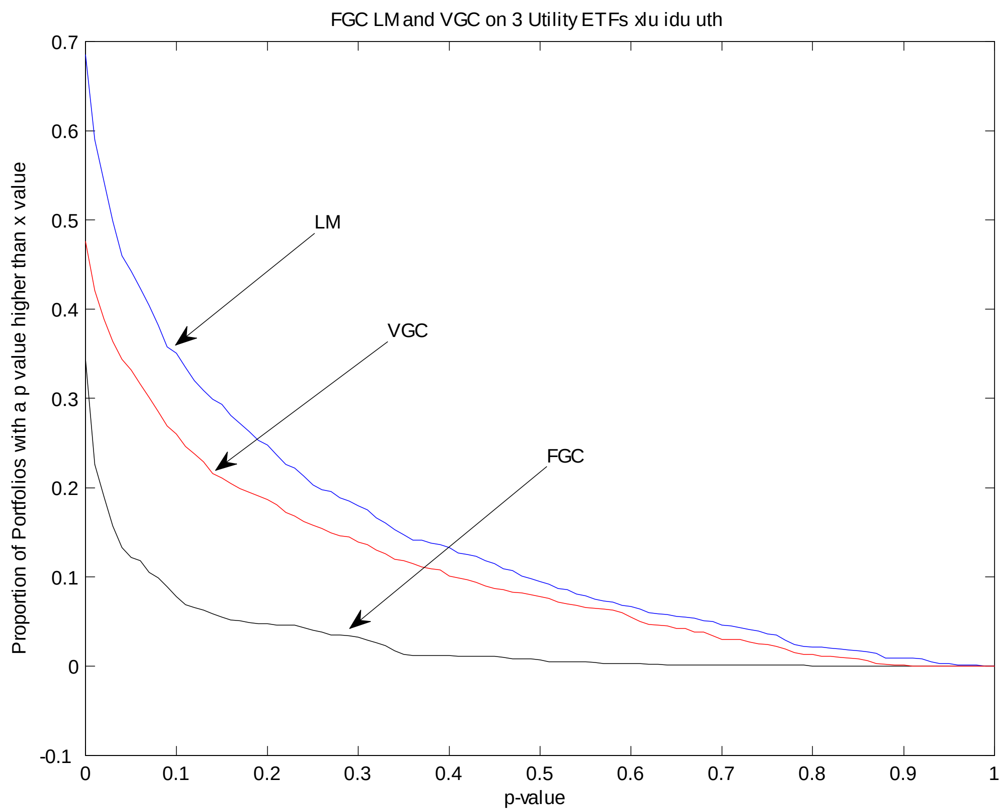

Figure 9: Portfolio Proportions with given probability values in the Utility sector.

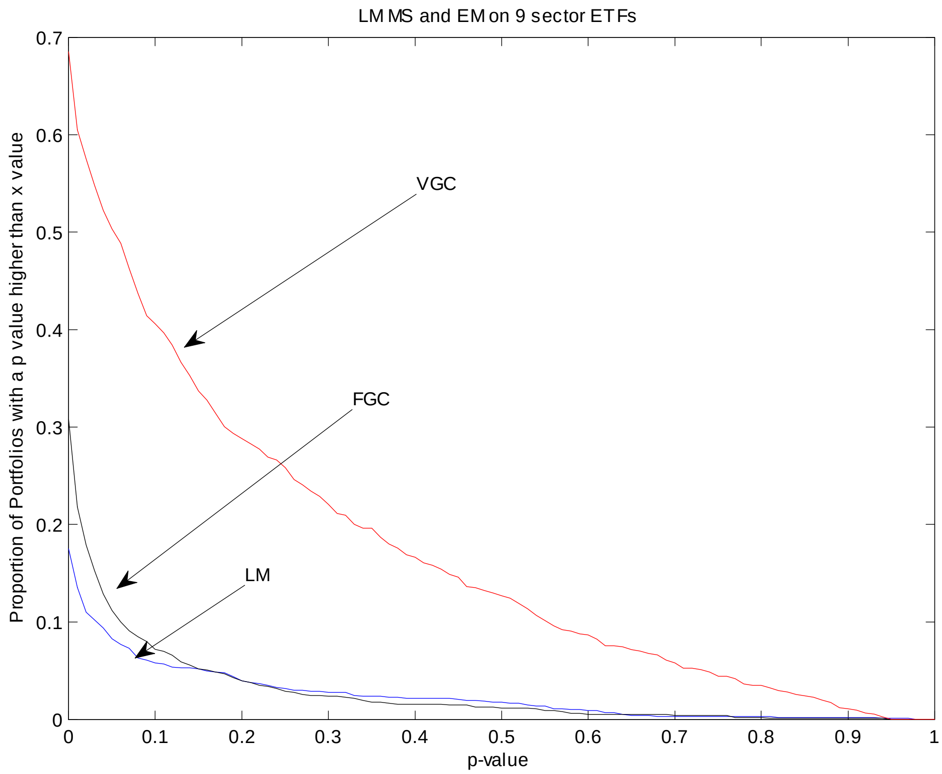

Figure 10: Portfolio Proportions with given probability values in the Cross Sector group.

Model Correlation Signatures

- We first present the details on how the correlation signatures are constructed for each model. In each case we extract the joint density for a pair of returns in the set of returns jointly modeled.

We then evaluate the local correlation numerically by evaluating the appropriate derivatives of the negative of the log likelihood.

FGC

- For the FGC model the joint density is obtained as follows.

- Let the marginal distribution functions be $F(x), G(y)$. We then have that

$$\begin{aligned} z_1 &= \Phi^{-1}(F(x)) \\ z_2 &= \Phi^{-1}(G(y)) \end{aligned}$$

are distributed bivariate normal with correlation ρ. The density of $z = (z_1, z_2)$ is

$$b(z_1, z_2)$$

It follows that the density of x, y is

$$q(x,y) = b(\Phi^{-1}(F(x)), \Phi^{-1}(G(y)))\frac{f(x)}{\phi\,(z_1)}\frac{g(y)}{\phi(z_2)}$$

Once we have q we may apply our localcorrelation surface construction to extract the correlation signature of this model.

- We now wish to incorporate scaling to unit variance. We may do this via the marginals as

$$X = \frac{x}{\sigma}$$

and

$$\begin{aligned} F_X(a) &= P(X \leq a) = P\left(\frac{x}{\sigma} \leq a\right) \\ & F_x(\sigma a). \end{aligned}$$

LML

- We wish to construct the correlation signatures of our models estimated for example on the energy sector. There are five $ETF\prime s$ for which the joint law was estimated and these are

  $$xle, iye, ieo, oih, xop$$

  We consider xle, iye and ieo, xop. We have modeled daily returns as

  $$\begin{aligned} x_i &= \phi_i' y \\ x_j &= \phi_j' y \end{aligned}$$

  and we have the joint characteristic functions but we shall compute the correlation signatures for standardized variates. The variance of x is given by

  $$\sigma_i^2 = \sum_k \phi_{ik}^2 \left(\sigma_k^2 + \theta_k^2 \nu_k \right).$$

- The standardized vector is

$$X_i = \frac{x_i}{\sigma_i} = \frac{\phi_i'}{\sigma_i} y = \mathbf{\Phi}_i' y$$

and it is centered of unit variance by construction.

- We easily obtain from the joint characteristic function the joint characteristic function of any 2 variates out of the full set modeled.

- We may invert this joint characteristic function using two dimensional Fourier inversion for the joint density.

VGC

- For the VGC model the construction is

$$\begin{aligned} X &= \theta_x(g_x - 1) + \sigma_x\sqrt{g_x}Z_x \\ Y &= \theta_y(g_y - 1) + \sigma_y\sqrt{g_y}Z_y \end{aligned}$$

- Hence given the density of $Z = (Z_x, Z_y)$ we may write

$$\begin{aligned} q(x,y) &= E\left[\begin{array}{c} b\left(\frac{x-\theta_x(g_x-1)}{\sigma_x\sqrt{g_x}}, \frac{y-\theta_y(g_y-1)}{\sigma_y\sqrt{g_y}}\right) \\ \times\frac{1}{\sigma_x\sqrt{g_x}}\frac{1}{\sigma_y\sqrt{g_y}} \end{array}\right] \\ &= \int_0^\infty\int_0^\infty b\left(\frac{x-\theta_x(g_x-1)}{\sigma_x\sqrt{g_x}}, \frac{y-\theta_y(g_y-1)}{\sigma_y\sqrt{g_y}}\right) \\ &\quad \times p_x(g_x)p_y(g_y)dg_xdg_y \end{aligned}$$

where p_x, p_y are the gamma densities for the two gamma time changes.

- We may write explicitly as

$$\int_0^\infty \int_0^\infty b\left(\frac{x-\theta_x(g_x-1)}{\sigma_x\sqrt{g_x}}, \frac{y-\theta_y(g_y-1)}{\sigma_y\sqrt{g_y}}\right) \times$$
$$\frac{1}{\sigma_x \nu_x^{\frac{1}{\nu_x}} \Gamma\left(\frac{1}{\nu_x}\right)} g_x^{\frac{1}{\nu_x}-\frac{3}{2}} e^{-\frac{g_x}{\nu_x}} \times$$
$$\frac{1}{\sigma_y \nu_y^{\frac{1}{\nu_y}} \Gamma\left(\frac{1}{\nu_y}\right)} g_y^{\frac{1}{\nu_y}-\frac{3}{2}} e^{-\frac{g_y}{\nu_y}} dg_x dg_y$$

- We make the change of variable to

$$\begin{aligned} w_x &= \frac{g_x}{\nu_x} \\ w_y &= \frac{g_y}{\nu_y} \end{aligned}$$

to get

$$\int_0^\infty \int_0^\infty b\left(\frac{x-\theta_x(\nu_x w_x - 1)}{\sigma_x\sqrt{\nu_x w_x}}, \frac{y-\theta_y(\nu_y w_y - 1)}{\sigma_y\sqrt{\nu_y w_y}}\right) \times$$

$$\frac{1}{\sigma_x\sqrt{\nu_x w_x}\Gamma\left(\frac{1}{\nu_x}\right)} w_x^{\frac{1}{\nu_x}-1} e^{-w_x} \times$$

$$\frac{1}{\sigma_y\sqrt{\nu_y w_y}\Gamma\left(\frac{1}{\nu_y}\right)} w_y^{\frac{1}{\nu_y}-1} e^{-w_y} dw_x dw_y$$

- We evaluate this as a double sum using Gauss-Laguerre quadrature for the construction of the joint density in two dimensions. This is then fed to the local correlation surface construction program for the correlation signature. We evaluate as

$$q(x,y) =$$
$$\sum_{ij} p_i p_j b\left(\frac{x-\theta_x(\nu_x w_i - 1)}{\sigma_x\sqrt{\nu_x w_i}}, \frac{y-\theta_y(\nu_y w_j - 1)}{\sigma_y\sqrt{\nu_y w_j}}\right) \times$$
$$\frac{1}{\sigma_x\sqrt{\nu_x w_i}\Gamma\left(\frac{1}{\nu_x}\right)} w_i^{\frac{1}{\nu_x}-1} \times$$
$$\frac{1}{\sigma_y\sqrt{\nu_y w_j}\Gamma\left(\frac{1}{\nu_y}\right)} w_j^{\frac{1}{\nu_y}-1}$$

where p_i are the Laguerre weights and w_i are the points.

Correlation Signature Results for Energy and the Cross Sector Group

- We present in Table 2 the correlation signatures for two pairs of stocks from the Energy sector, ieo,xop and xle,iye and three pairs of stocks from the cross sector group xly,xlp, xli,xlk and xly,xli.

- We observe that the local correlations in LM tend to be substantially higher and particularly so in the tails.

- The local correlations are computed at the center of the distributions and points 10% up and down from this level.

Conclusion

- Three models of dependence in asset returns with non-Gaussian marginals are investigated on ETF daily return data.

- The first is a full rank Gaussian copula also studied and proposed in Malvergne and Sornette (2005) termed FGC.

- The second is a linear mixture of independent Lévy processes as proposed in Madan and Yen (2008) and studied in Madan (2006) termed LML.

- The third correlates Gaussian components in a variance gamma representation of the marginals as proposed in Eberlein and Madan (2008) termed VGC.

- All three models are easily estimated in fairly high dimensions as most of the work is done at a univariate level. The models are evaluated on the basis of their ability to explain the univariate laws of randomly generated portfolios.

- It is observed that on a number of occasions all three models are at a comparable level of performance.

- In some cases we get a superior performance from the LML model followed by VGC and FGC. There are occasions when the VGC and FGC dominate.

- The three models are tractable in different ways with the LM model yielding closed form characteristic functions.

- With a view to exploring more deeply the different forms of dependence modeling the concept of local correlation is introduced. It is shown that the LML model displays higher levels of local correlation than that obtained in the FGC and VGC models.

Table 2

Correlation Signature for LM

	xop		
	0.663066	0.702658	0.725284
ieo	0.639887	0.694015	0.735871
	0.589141	0.664609	0.73184

	iye		
	0.709821	0.64253	0.543917
xle	0.725103	0.689161	0.627705
	0.712397	0.721196	0.707472

	xlp		
	0.870662	0.824027	0.789156
xly	0.80326	0.791869	0.800942
	0.737392	0.775675	0.828535

	xlk		
	0.88267	0.823123	0.768254
xli	0.831342	0.79638	0.779932
	0.790777	0.788804	0.809007

	xli		
	0.881328	0.833928	0.799507
xly	0.828022	0.800518	0.792809
	0.780602	0.784001	0.806905

Correlation Signature for FGC

	xop		
	0.151467	0.15983	0.179276
ieo	0.160283	0.168552	0.188268
	0.179742	0.188198	0.209177

	iye		
	0.073226	0.080281	0.092234
xle	0.076677	0.083968	0.096331
	0.083994	0.091854	0.105202

	xlp		
	0.518616	0.539077	0.591593
xly	0.539412	0.554894	0.601744
	0.590685	0.599728	0.641581

	xlk		
	0.692328	0.705675	0.746557
xli	0.724125	0.727674	0.758185
	0.795676	0.785003	0.803508

	xli		
	0.649106	0.675192	0.735665
xly	0.66263	0.680397	0.729319
	0.702033	0.71057	0.749299

Correlation Signature for VGC

	xop		
	0.206153	0.211492	0.2177
ieo	0.212761	0.21722	0.222705
	0.21984	0.223406	0.228671

	iye		
	0.184961	0.203322	0.188705
xle	0.162205	0.176877	0.163207
	0.162894	0.177886	0.164916

	xlp		
	0.586448	0.593867	0.598697
xly	0.591685	0.599928	0.604753
	0.599142	0.604244	0.608796

	xlk		
	0.686074	0.698671	0.704153
xli	0.681454	0.696105	0.704073
	0.688898	0.69427	0.702349

	xli		
	0.719777	0.713182	0.714502
xly	0.726997	0.725125	0.722426
	0.72911	0.730034	0.728464