Modeling Co-movements and Tail Dependency in the International Stock Market via Copulae

Katja Ignatieva, Eckhard Platen

MACQUARIE UNIVERSITY SYDNEY ~ AUSTRALIA

UNIVERSITY OF TECHNOLOGY SYDNEY

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Bachelier Finance Society World Congress
22-26 June 2010, Toronto

Linear Portfolio

- *value of portfolio* $w = (w_1, \ldots, w_d)^\top$ of assets $S_t = (S_{1,t}, \ldots, S_{d,t})^\top$:

$$V_t = \sum_{j=1}^{d} w_j S_{j,t}$$

- *profit and loss* (P&L) function:

$$L_{t+1} = (V_{t+1} - V_t) = \sum_{j=1}^{d} w_j S_{j,t}(e^{X_{j,t+1}} - 1)$$

$$X_{t+1} = (\log S_{t+1} - \log S_t)$$

- *Value-at-Risk* at level α:

$$VaR(\alpha) = F_L^{-1}(\alpha)$$

The VaR depends on
the distribution F_X of the risk factor increments $X = (X_1, \ldots, X_d)^\top$.

1. How to model the dependency among $X_1, \ldots, X_d$?
2. How does F_X and the dependency among $X_1, \ldots, X_d$ vary over time ?

Traditional approach: Riskmetrics

- the conditional distribution of log-returns is multivariate normal: $X_t \sim N(0, \Sigma_t)$
- the covariance matrix Σ_t is estimated by:

$$\widehat{\Sigma}_t = (e^{\lambda} - 1) \sum_{s<t} e^{-\lambda(t-s)} X_{t-s} X_{t-s}^T$$

- *decay factor* λ $(0 < \lambda < 1)$ is determined by backtesting
- $\lambda = 0.94$ provides best results (Morgan/Reuters, 1996)
- Drawbacks:
  - does not allow to generate tail dependence
  - does not allow heavy tails

Copula based approach

- the conditional distribution of log-returns is modelled with Copula C:

$$X_t \sim C\{F_{X_1}(x_1), \ldots, F_{X_d}(x_d), \theta_t\}$$

- $F_{X_1}, \ldots, F_{X_d}$ are marginal distributions
- θ_t dependence parameter

$$\Downarrow$$

- Specify marginal distributions
- Specify dependence structure

Outline

1. Motivation ✓
2. Copulae and Value-at-Risk
3. Copula Estimation
4. Empirical Analysis
   - Specify marginals
   - Specify dependence structure
5. Value-at-Risk applications
6. Conclusion

Copulae

Theorem (Sklar's Theorem)

Let F be a d-dimensional distribution function with marginals $F_1 \ldots, F_d$. Then there exists a copula C with

$$F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\} \tag{1}$$

for every $x_1, \ldots, x_d \in \overline{\mathbb{R}}$. If $F_1, \ldots, F_d$ are continuous, then C is unique. On the other hand, if C is a copula and $F_1, \ldots, F_d$ are distribution functions, then the function F defined in (1) is a joint distribution function with marginals $F_1, \ldots, F_d$.

Generating tail dependence

1. Elliptical Copulae
   - **Gaussian Copula** (no tail dependence)
     $C_{\Psi}^{Ga}(u_1, \ldots, u_d) = \Phi_{\Psi}\{\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\}$ where Φ_{Ψ} d-dimensional standard normal cdf
   - **Student-t Copula** (symmetric tail dependence)
     $C_{\nu,\Psi}^{t}(u_1, \ldots, u_d) = t_{\nu,\Psi}\{t_{\nu}^{-1}(u_1), \ldots, t_{\nu}^{-1}(u_d)\}$ where $t_d(\nu, 0, \Psi)$ is Student-t cdf, Ψ is the correlation matrix, ν df
2. Archimedean Copulae, Mixture Copula Models
   - **Clayton** (lower tail dependence) $\theta \in (0, \infty)$
   - **Gumbel** (upper tail dependence) $\theta \in (1, \infty)$
3. Survival Copulae

$$C^*(u_1, u_2) = 1 - u_1 - u_2 + C(1 - u_1, 1 - u_2)$$

   - **survival Clayton** (upper tail dependence)
   - **survival Gumbel** (low tail dependence)

Value-at-Risk with Copulae

The process $\{X_t\}_{t=1}^{T}$ of log-returns can be modelled as

$$X_{j,t} = \mu_{j,t} + \sigma_{j,t}\varepsilon_{j,t}$$

with $E[\varepsilon_{j,t}] = 0$, $E[\varepsilon_{j,t}^2] = 1, j = 1, \ldots, d$ and

$$E[X_{j,t} \mid \mathcal{F}_{t-1}] = \mu_{j,t}$$

$$E[(X_{j,t} - \mu_{j,t})^2 \mid \mathcal{F}_{t-1}] = \sigma_{j,t}^2$$

where $\mathcal{F}_t$ is the available information at time t.

- $\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{d,t})^\top$ are standardized $i.i.d.$ innovations with a joint distribution function F_ε
- ε_j, $j = 1, \ldots, d$ have continuous marginal distributions F_j

VaR with Copulae

For the log-returns $\{x_{j,t}\}_{t=1}^{T}$, $j = 1, \ldots, d$ Value-at-Risk at level α is estimated:

1. determination of the innovations $\hat{\varepsilon}_t$ (e.g. by deGARCHing)
2. specification and estimation of marginal distributions $F_j(\hat{\varepsilon}_j)$
3. specification of a copula C and estimation of dependence parameter θ
4. simulation of innovations ε and losses L
5. determination of $\widehat{VaR}(\alpha)$, the empirical α-quantile of F_L.

Copula estimation

The distribution of $X = (X_1, ..., X_d)^\top$ with marginals $F_{X_j}(x_j, \delta_j)$, $j = 1, ..., d$ is given by:

$$F_X(x_1, \ldots, x_d) = C\{F_{X_1}(x_1; \delta_1), \ldots, F_{X_d}(x_d; \delta_d); \theta\}$$

and its density is given by

$$f(x_1, \ldots, x_d; \delta_1, \ldots, \delta_d, \theta)$$

$$= c\{F_{X_1}(x_1; \delta_1), \ldots, F_{X_d}(x_d; \delta_d); \theta\} \prod_{j=1}^{d} f_j(x_j; \delta_j)$$

where c is a copula density.

Copula estimation

- For a sample of observations $\{x_t\}_{t=1}^{T}$ and $\vartheta = (\delta_1, \dots, \delta_d, \theta)^\top \in \mathbb{R}^{d+1}$ the likelihood function is

$$L(\vartheta; x_1, \dots, x_T) = \prod_{t=1}^{T} f(x_{1,t}, \dots, x_{d,t}; \delta_1, \dots, \delta_d, \theta)$$

  and the corresponding log-likelihood function $\ell(\vartheta; x_1, \dots, x_T) = \sum_{t=1}^{T} \log c\{F_{X_1}(x_{1,t}; \delta_1), \dots, F_{X_d}(x_{d,t}; \delta_d); \theta\} + \sum_{t=1}^{T} \sum_{j=1}^{d} \log f_j(x_{j,t}; \delta_j)$

- **Estimation methods:**
  - Exact Maximum Likelihood
  - Inference for Margins
  - Canonical Maximum Likelihood

Data Set

Data used

- for regional indices
  - S&P 500
  - Dow Jones EURO STOXX 50
  - FTSE 100
  - TOPIX
- Sample period from 01 January 1987 to 10 March 2006

Specify Marginal Distribution

Symmetric generalized hyperbolic (SGH) family of distributions:

$$f_X(x) = \frac{1}{\delta\sigma K_\lambda(\bar{\alpha})}\sqrt{\frac{\bar{\alpha}}{2\pi}}\left(1+\frac{x^2}{(\delta\sigma)^2}\right)^{\frac{1}{2}(\lambda-\frac{1}{2})} K_{\lambda-\frac{1}{2}}\left(\bar{\alpha}\sqrt{1+\frac{x^2}{(\delta\sigma)^2}}\right)$$

- $K_\lambda(\cdot)$ *Bessel function*
- λ and $\bar{\alpha}$ are the *shape parameters*: $\bar{\alpha} \neq 0$ if $\lambda \geq 0$ and $\delta \neq 0$ if $\lambda \leq 0$
  - *Variance Gamma* (VG) distribution: $\bar{\alpha} = 0$ and $\lambda > 0$
  - *Student-t* distribution: $\bar{\alpha} = 0$ and $\lambda < 0$ (consider $\lambda \leq -1$ for $\nu = -2\lambda \geq 2$, std.dev. $\sigma_X = \sigma\sqrt{\frac{\nu}{\nu-2}}$)
  - *Hyperbolic* (HYP) distribution: $\lambda = 1$
  - *Normal Inverse Gaussian* (NIG) distribution: $\lambda = -0.5$

Specify Marginal Distribution

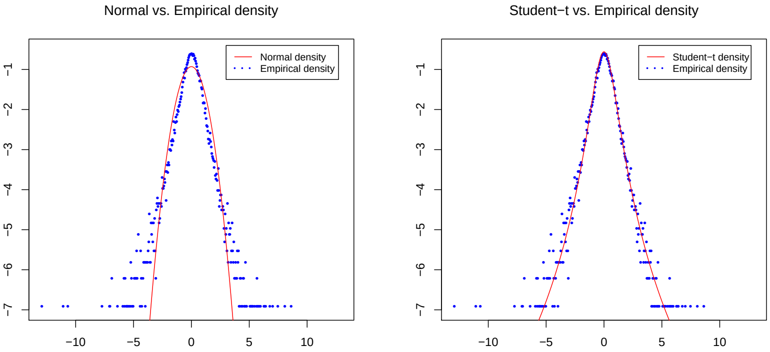

Figure: Logarithm of the histogram for the pooled data vs. normal density (left panel) and Student-t density (right panel). Pooled data is taken for indices S&P 500, Dow Jones EURO STOXX 50, FTSE100, TOPIX from 01 January 1987 to 10 March 2006. Estimated number of degrees of freedom for the Student-t distribution is $\nu = 3.15$.

Specify Marginal Distribution

- Goodness-of-fit testing: *Anderson-Darling* (AD) distance and the *Kolmogorov-Simirnov* (KS)

$$AD = \frac{\sup_{x \in \mathcal{R}} |F_s(x) - \hat{F}(x)|}{\sqrt{\hat{F}(x)(1-\hat{F}(x))}},$$

$$KS = \sup_{x \in \mathcal{R}} |F_s(x) - \hat{F}(x)|,$$

  - $F_s(x)$ denotes the empirical sample distribution
  - $\hat{F}(x)$ is the estimated distribution.

Specify Marginal Distribution

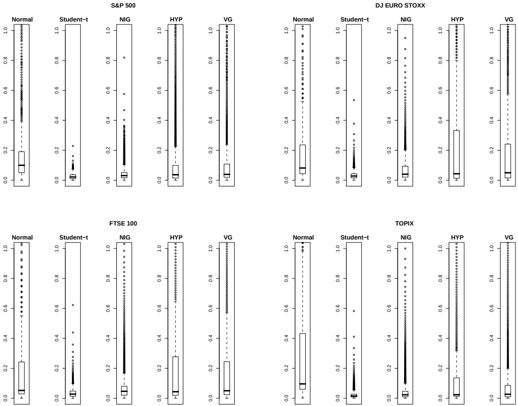

Figure: Box-plots for Anderson-Darling distance for modelling marginal distributions of the S&P 500, Dow Jones EURO STOXX 50, FTSE100, TOPIX with alternative residual distributions.

Model Selection: static case

- Akaike information criterion (AIC):

$$AIC = -2L(\alpha; x_1, \ldots, x_T) + 2q$$

favors:

- Student-t copula vs. mixture Gumbel & survival Gumbel for two-constituents portfolios where TOPIX is not included
- Student-t copula vs. mixture Gumbel & survival Gumbel model for a 3-constituent portfolio (S&P 500, DJ EURO STOXX 50, FTSE 100)
- Mixture Clayton & Gumbel model vs. Student-t copula for a 4-constituent portfolio (S&P 500, DJ EURO STOXX 50, FTSE 100, TOPIX)

Time-varying estimation

- Static case: estimate the dependence parameter at once based on the whole series of observations.
- Time-varying case:
  - Estimate the dependence parameter by using subsets of size n of log-returns, that is a moving window of size n, $\{\widehat{X}_t\}_{t=s-n+1}^{s}$ scrolling in time for $s = n, ..., T$
  - It generates a time-series for the dependence parameter $\{\widehat{\theta}_t\}_{t=n}^{T}$ and time-series of VaR: $\{\widehat{VaR}_t\}_{t=n}^{T}$.

Student-t dependence parameter time-varying

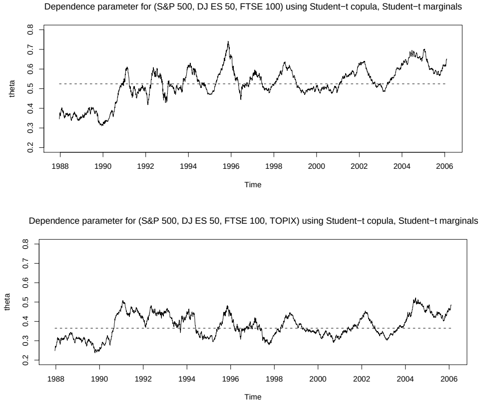

Figure: Copula dependence parameter $\hat{\theta}$ estimated for a 3-constituent portfolio (S&P 500, Dow Jones EURO STOXX 50, FTSE 100) (upper panel) and 4-constituent portfolio constructed of (S&P 500, Dow Jones EURO STOXX 50, FTSE 100, TOPIX) (lower panel) using Student-t copula with Student-t marginals.

VaR for portfolio

- The one-day VaR at time t and significance level α is given by the α-quantile of the distribution of the P&L:

$$VaR_t(\alpha) = F^{-1}_{L_{t+1}}(\alpha),$$

- The expected shortfall (ES) at time t is:

$$ES_t(\alpha) = \frac{1}{N_{t+1}} \sum_{i=1}^{N_{t+1}} L_{t+1,i} \mathbf{1}_{\{L_{t+1,i} \leq VaR_t(\alpha)\}},$$

  - N_{t+1} is the number of simulated portfolio returns with value less or equal than $VaR_t(\alpha)$ and $L_{t+1,i}$ is the i^{th} outcome of the N_{t+1} samples.

Backtesting

Compare the estimated values for the VaR with the true realizations $\{L_t\}$ of the P&L function

the *exceedances ratio* is given by

$$\hat{\alpha} = \frac{1}{T-w}\sum_{t=w}^{T} \mathbf{1}\{L_t < \widehat{VaR}_t(\alpha)\}$$

Table: Exceedances ratios

	(S&P 500, Dow Jones EURO STOXX 50, FTSE 100)					
Copula	0.1	0.05	0.01	0.005	0.001	$\sum_\alpha((\alpha-\widehat{\alpha})/\alpha)^2$
Student-t	0.094180	0.043981	0.004442	0.001111	0.000444	1.242913
Gumbel & surv. Gumbel	0.125499	0.077521	0.021990	0.011106	0.002221	4.788473
Riskmetrics	0.106525	0.063471	0.024190	0.016866	0.009099	73.31631
	(S&P 500, Dow Jones EURO STOXX 50, FTSE 100, TOPIX)					
Copula	0.1	0.05	0.01	0.005	0.001	$\sum_\alpha((\alpha-\widehat{\alpha})/\alpha)^2$
Student-t	0.096179	0.041315	0.003110	0.000888	0.000444	1.491430
Clayton & Gumbel	0.127277	0.062639	0.007552	0.024621	0.000222	1.064764
Riskmetrics	0.102308	0.058145	0.017310	0.010874	0.004882	17.01467

Estimated VaR

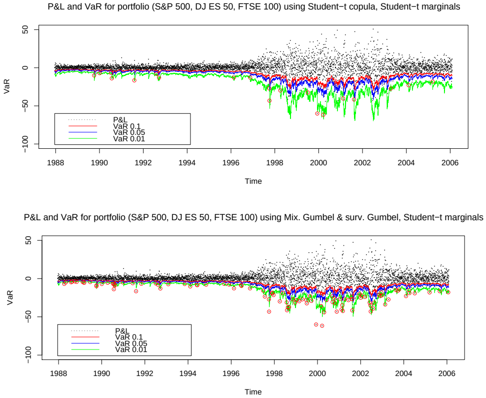

Figure: P&L, VaR estimated at different confidence levels using Student-t copula (upper panel) and mixture model Gumbel & survival Gumbel with (lower panel) for a 3-constituent portfolio of (S&P 500, Dow Jones EURO STOXX 50 FTSE 100); Student-t marginals; exceedances at level $\alpha = 0.01$.

Summarize Results

Summarize Results:

- Student-t assumption allows to better capture the dependent extreme values which can be observed in index log-returns
- Log-returns of the indices follow the Student-t distribution with about four degrees of freedom
- Dependence structure:
    - Student-t is preferred over mixture Gumbel & surv. Gumbel for (S&P 500, Dow Jones EURO STOXX 50, FTSE 100)
    - Mixture Clayton & Gumbel is preferred over Studnet-t for (S&P 500, Dow Jones EURO STOXX 50, FTSE 100, TOPIX) providing the best backtesting results.

Thank you very much!

Thank you very much for your attention!