An efficient method for market risk management under multivariate extreme value theory approach

Miloš Božović
Center for Investments and Finance
Belgrade, Serbia

NBS Research Seminar
Belgrade, Dec 17, 2010

Introduction

- Market risk assessment typically relies on quantile-based measures of risk:
  - Incoherent measures (e.g. Value at Risk)
  - Coherent measures (e.g. Expected Shortfall)
- Traditional methods of VaR and ES estimation:
  - Analytical method
  - Historical simulation
  - Monte Carlo simulation
- Problem with traditional methods:
  - They try to reconstruct the entire distribution of returns.
  - However, extreme losses matter the most.

Extreme Value Theory (EVT)

- Characterizes the tail behavior of the distribution of returns.
- By focusing on extreme losses, the EVT successfully avoids tying the analysis down to a single parametric family fitted to the whole distribution.
- The empirical results show that VaR and ES estimates obtained by using EVT-based models outperform the ones based on analytical and historical methods.
  - McNeil (1997), Nyström & Skoglund (2002), Harmantzis, Chien & Miao (2005), and Marinelli, d'Addona & Rachev (2007)

Univariate EVT: Theorem 1

- Fisher & Tippett (1928), Gnedenko (1943)

Let $\{X_i\}_{i=1}^n$ be a set of n independent and identically distributed random variables with distribution function F and suppose that there are sequences of normalization constants, $\{a_n\}$ and $\{b_n\}$, such that, for some non-degenerated limit distribution F^*, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} [F(a_n x + b_n)]^n = F^*(x), \quad x \in \mathbb{R}.$$

Then, there exist $\xi \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$ such that $F^*(x) = \Gamma_{\xi,\mu,\sigma}(x)$ for any $x \in \mathbb{R}$, where

$$\Gamma_{\xi,\mu,\sigma}(x) := \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right]$$

is the so-called *generalized extreme value (GEV)* distribution.

Univariate EVT: Theorem 2

- Picklands (1975)

Let $\{X_i\}_{i=1}^n$ be a set of n independent and identically distributed random variables with distribution function F. Define

$$F_u(y) := \mathbb{P}\left(X \leq u + y \mid X > u\right) = \frac{F(u+y) - F(u)}{1 - F(u)}, \quad y > 0$$

to be the distribution of excesses of X over the treshold u. Let x_F be the end of the upper tail of F, possibly a positive infinity. Then, if F is such that the limit given by Theorem 1 exists, there are constants $\xi \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$ such that

$$\lim_{u \to x_F} \sup_{u < x < x_F} |F_u(x) - G_{\xi,\beta}(x - u)| = 0,$$

where

$$G_{\xi,\beta}(y) := 1 - \left(1 + \xi \frac{y}{\beta}\right)_+^{-1/\xi} \tag{1}$$

is known as the *generalized Pareto (GP)* distribution.

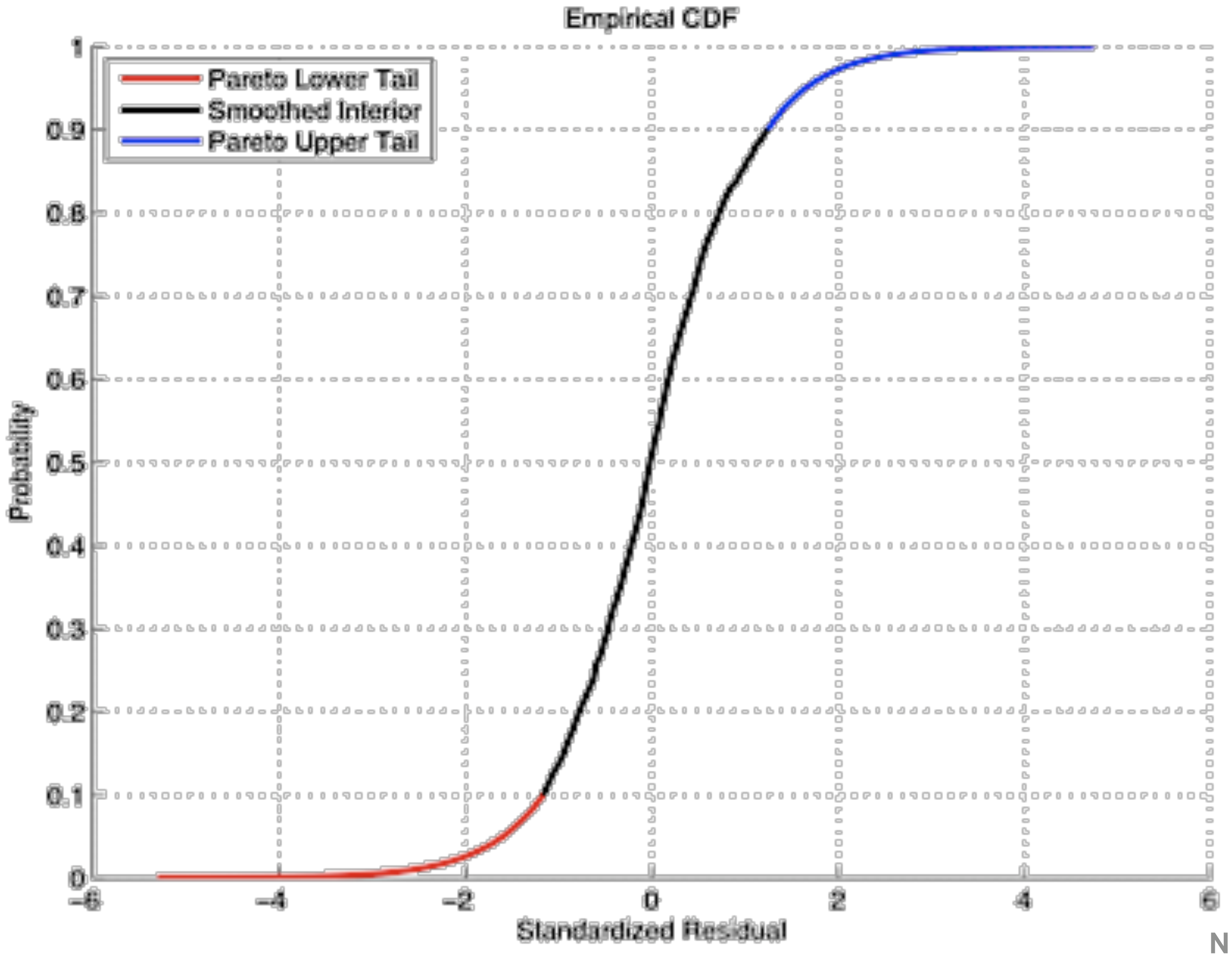
Multivariate EVT

- The results of the univariate EVT hold for i.i.d. random numbers.
- Multidimensional limiting relations are also available (e.g. Smith 2000), but model complexity increases greatly with the number of risk factors.
- Joint distributions of extreme returns often modeled by copulas.
- The proposed multivariate EVT method is based on **separate estimations of the univariate model**.
  - Works with *n* orthogonal series of conditional residuals that are approximately i.i.d.

Estimation methodology

- Choose a reasonable treshold *u*.
- Use one of the conventional methods to fit the interior of the distribution (e.g. historical simulation).
- Fit the upper and lower tail separately with GP distribution:
  - Using the Hill estimator (Danielsson & de Vries, 1997), or
  - Maximum likelihood estimator
- Nyström and Skoglund (2002) show that the ML outperforms the Hill estimator, and is almost insensitive to the choice of treshold *u*.

Estimation methodology



Estimation methodology

- Total distribution beyond the treshold is given by

$$F(x) = (1 - F(u))\, G_{\xi,\beta}(x-u) + F(u)$$

- From the estimated parameters of the GP distribution we find

$$\widehat{\mathrm{VaR}}_{q\pm} = u_\pm \pm \frac{\widehat{\beta}_\pm}{\widehat{\xi}_\pm}\left[\left(\frac{T}{N_{u\pm}}(1-q_\pm)\right)^{-\widehat{\xi}_\pm} - 1\right]$$

$$\widehat{\mathrm{ES}}_{q\pm} = \frac{1}{1-\widehat{\xi}_\pm}\left(\widehat{\mathrm{VaR}}_{q\pm} \pm \widehat{\beta}_\pm - \widehat{\xi}_\pm u_\pm\right)$$

where + (–) stands for the upper (lower) tail.

Orthogonalization

- In *n* dimensions, use the principal components of the unconditional VCV matrix

$$
\begin{aligned}
\mathbf{\Lambda} &:= \operatorname{diag}(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n) \\
\mathbf{V}_\infty &= \mathbf{P}\mathbf{\Lambda}\mathbf{P}' \\
\mathbf{L} &:= \mathbf{P}\mathbf{\Lambda}^{1/2} \\
\mathbf{z}_t &= \mathbf{L}^{-1}\boldsymbol{\varepsilon}_t \\
\mathbb{E}\left(\mathbf{z}_t\right) &= \mathbf{0} \\
\operatorname{var}\left(\mathbf{z}_t\right) &= \mathbf{1}_n
\end{aligned}
$$

- This gives a set of (at most) *n* orthogonal standardized coordinates.
- Univariate EVT can be applied separately on each.

Filtering

- Log-returns: ARMA(*r*,*m*)

$$\begin{aligned} y_{t,i} &= \mu_i + \sum_{s=1}^{r} b_{s,i} y_{t-s,i} + \varepsilon_{t,i} + \sum_{s=1}^{m} \theta_{s,i} \varepsilon_{t-s,i} \\ \mathbb{E}\left(\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1}\right) &= \mathbb{E}\left(\boldsymbol{\varepsilon}_t\right) = [0 \quad 0 \quad \dots \quad 0]' =: \mathbf{0}, \\ \operatorname{var}\left(\boldsymbol{\varepsilon}_t | \mathcal{F}_{t-1}\right) &= \mathbb{E}\left(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' | \mathcal{F}_{t-1}\right) =: \mathbf{V}_t \end{aligned}$$

- Conditional variance: GJR-GARCH(*p*,*q*)

$$\begin{aligned} \mathbf{V}_t &= \boldsymbol{\Omega} + \sum_{s=1}^{p} \mathbf{A}_s \mathbf{E}_{t-s} + \sum_{s=1}^{p} \boldsymbol{\Theta}_s \mathbf{I}_{t-s} \mathbf{E}_{t-s} + \sum_{s=1}^{q} \mathbf{B}_s \mathbf{V}_{t-s} \\ \mathbf{E}_t &:= \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \\ \mathbf{I}_t &:= \operatorname{diag}(\operatorname{sgn}(-\varepsilon_{t,1})_+,\ \operatorname{sgn}(-\varepsilon_{t,2})_+,\ \dots,\ \operatorname{sgn}(-\varepsilon_{t,n})_+) \end{aligned}$$

Filtering

- In the basis of principal components:

$$\begin{aligned}
\widehat{\mathbf{V}}_t &= \widehat{\mathbf{\Omega}} + \sum_{s=1}^{p} \widehat{\mathbf{A}}_s \widehat{\mathbf{E}}_{t-s} + \sum_{s=1}^{p} \widehat{\mathbf{\Theta}}_s \widehat{\mathbf{I}}_{t-s} \widehat{\mathbf{E}}_{t-s} + \sum_{s=1}^{q} \widehat{\mathbf{B}}_s \widehat{\mathbf{V}}_{t-s} \\
\widehat{\mathbf{E}}_t &:= \mathbf{L}^{-1} \mathbf{E}_t \mathbf{L}^{-1\prime} = \mathbf{z}_t \mathbf{z}_t' \\
\widehat{\mathbf{I}}_t &:= \mathbf{L}^{-1} \mathbf{I}_t \mathbf{L}^{-1\prime} = \mathrm{diag}(\mathrm{sgn}(-z_{t,1})_+,\ \mathrm{sgn}(-z_{t,2})_+,\ \ldots,\ \mathrm{sgn}(-z_{t,n})_+)
\end{aligned}$$

- Estimate *n* separate univariate GJR-GARCH(*p*,*q*) models:

$$\widehat{V}_{t,i} = \widehat{\Omega}_i + \sum_{s=1}^{p} \widehat{A}_{s,i} \widehat{E}_{t-s,i} + \sum_{s=1}^{p} \widehat{\Theta}_{s,i} \widehat{I}_{t-s,i} \widehat{E}_{t-s,i} + \sum_{s=1}^{q} \widehat{B}_{s,i} \widehat{V}_{t-s,i}$$

GMM estimation

- Use GMM to avoid:
  - having a unique likelihood function
  - specific distributional assumptions
- The estimator (Skoglund, 2001):

$$
\begin{aligned}
\mathbf{e}_t &:= \left[z_t \quad z_t^2 - \widehat{V}_t \right]' \\
\mathbf{g}_t(\psi) &:= \mathbf{F}_t'(\psi)\mathbf{e}_t \\
\mathbb{E}\left[\mathbf{g}_t(\psi)\right] &= \mathbf{0} \\
\mathbf{m}(\psi) &:= \frac{1}{T}\sum_{t=1}^{T} \mathbf{g}_t(\psi) \\
\widehat{\psi} &= \arg\min_{\psi} \ \mathbf{m}(\psi)' \ \mathbf{W} \ \mathbf{m}(\psi)
\end{aligned}
$$

Forecasting

- Confidence interval for the forecast of the value of *i*-th principal component, *h* steps ahead:

$$z^{\pm}_{t+h,i} = F_i^{-1}(q_{\pm})\sqrt{\widehat{V}_{t+h,i}}$$

- Forecasts of multivariate VaR and ES:

$$\begin{aligned}
\mathrm{VaR}_{q_{\pm}} &= \mathbf{a}'\mathbf{L}\left[\mathbb{E}\left(\widehat{\mathbf{y}}_{t+h}|\mathcal{F}_t\right) + \mathbf{z}^{\pm}_{t+h}\right] \\
\mathrm{ES}_{q_{\pm}} &= \mathbf{a}'\mathbf{L}\widetilde{\mathbf{z}}^{\pm}_{t+h} \\
\widetilde{z}^{\pm}_{t+h,i} &= \widetilde{F}_i^{-1}(q_{\pm})\sqrt{\widehat{V}_{t+h,i}} \\
\widetilde{F}_i^{-1}(q_{\pm}) &= \frac{1}{1-\xi_{\pm}}\left[F_i^{-1}(q_{\pm}) \pm \beta_{\pm} - \xi_{\pm}u_{\pm}\right]
\end{aligned}$$

Data

- Daily averages of four interbank exchange rates
  - Base currency: USD
  - Term currencies: EUR, GBP, JPY, CHF
- Sample 1: January 4, 1999 - December 31, 2007
- Sample 2: January 4, 1999 – September 30, 2008
- Source: Thomson's Datastream

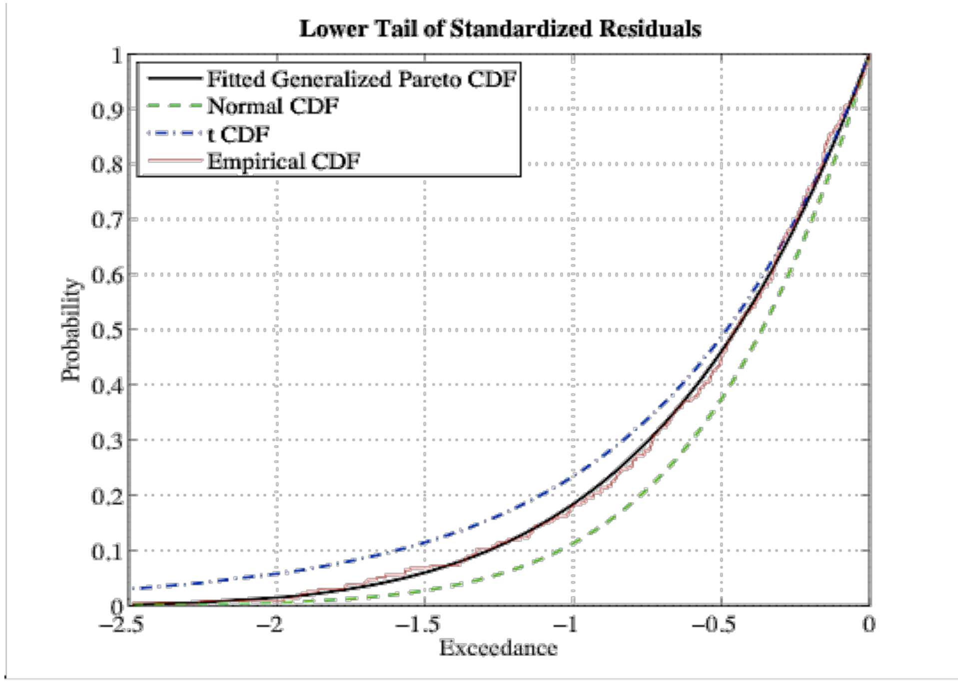
Parameters of the GP distribution

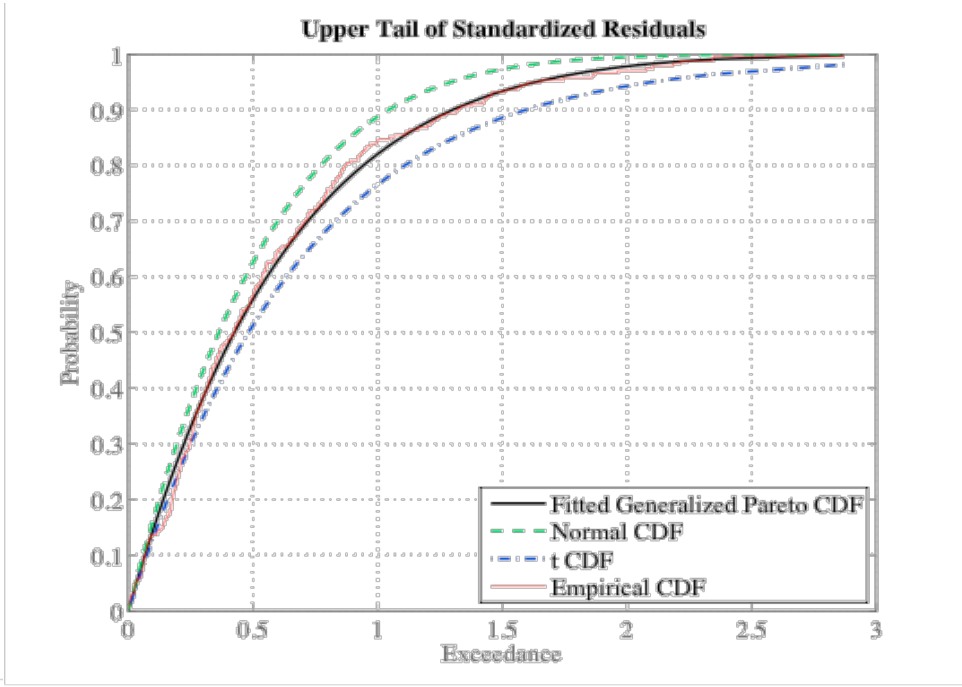
Upper tail				
Parameter	PC 1	PC 2	PC 3	PC 4
$\widehat{\xi}_+$	-0.1096	0.0544	0.0058	0.1804
	(0.0612)	(0.0703)	(0.0755)	(0.0752)
$\widehat{\beta}_+$	0.6397	0.5153	0.5724	0.5386
	(0.0573)	(0.0497)	(0.0575)	(0.0536)

Lower tail				
Parameter	PC 1	PC 2	PC 3	PC 4
$\widehat{\xi}_-$	-0.2030	0.0570	-0.0293	0.0239
	(0.0575)	(0.0714)	(0.0572)	(0.0625)
$\widehat{\beta}_-$	0.7013	0.6765	0.6379	0.6031
	(0.0605)	(0.0658)	(0.0556)	(0.0548)

Tails of the standardized residuals

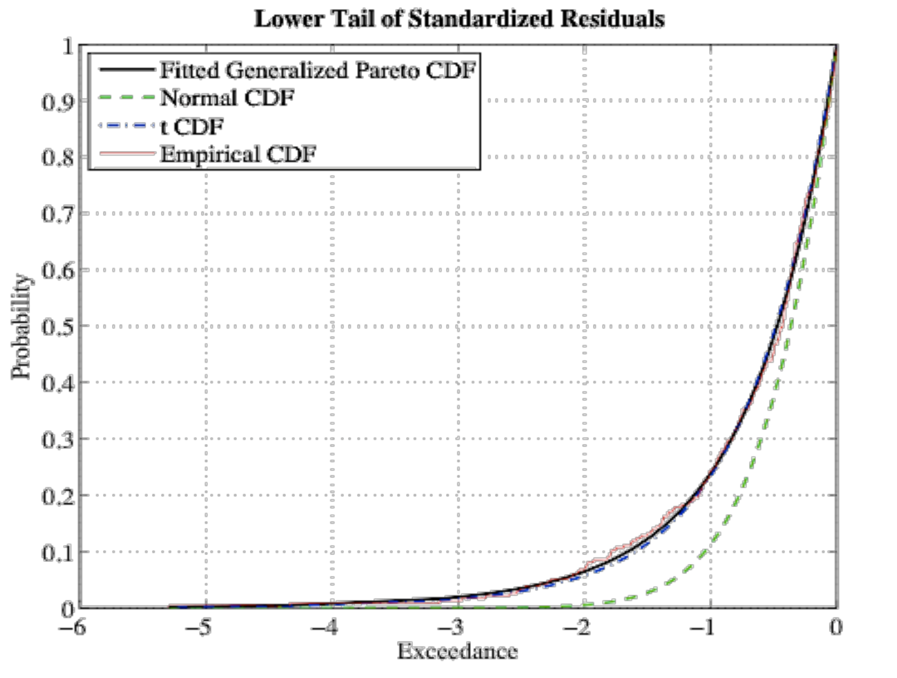
First principal component

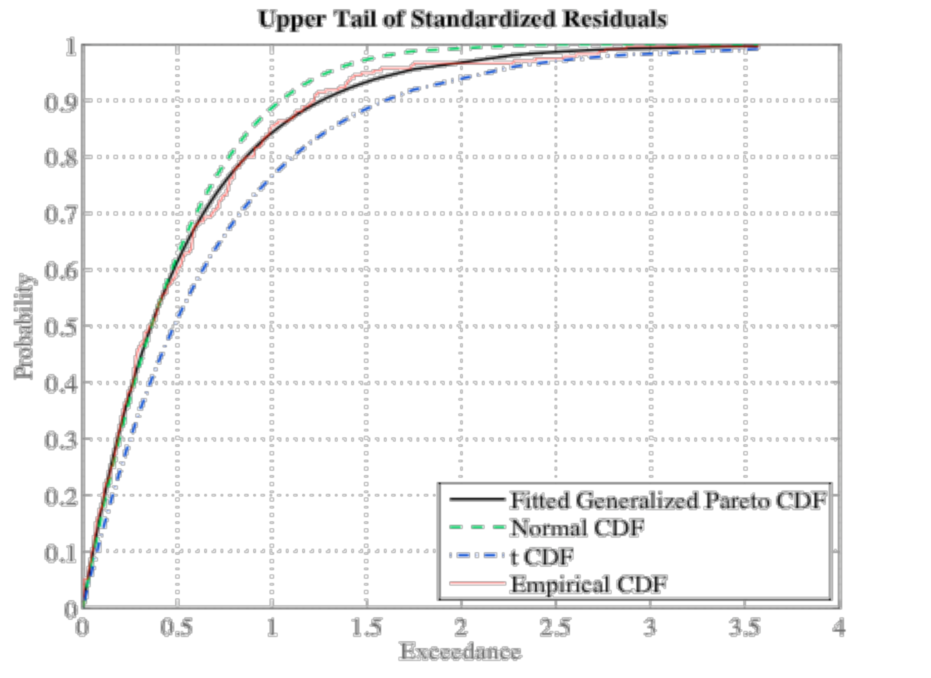




Tails of the standardized residuals

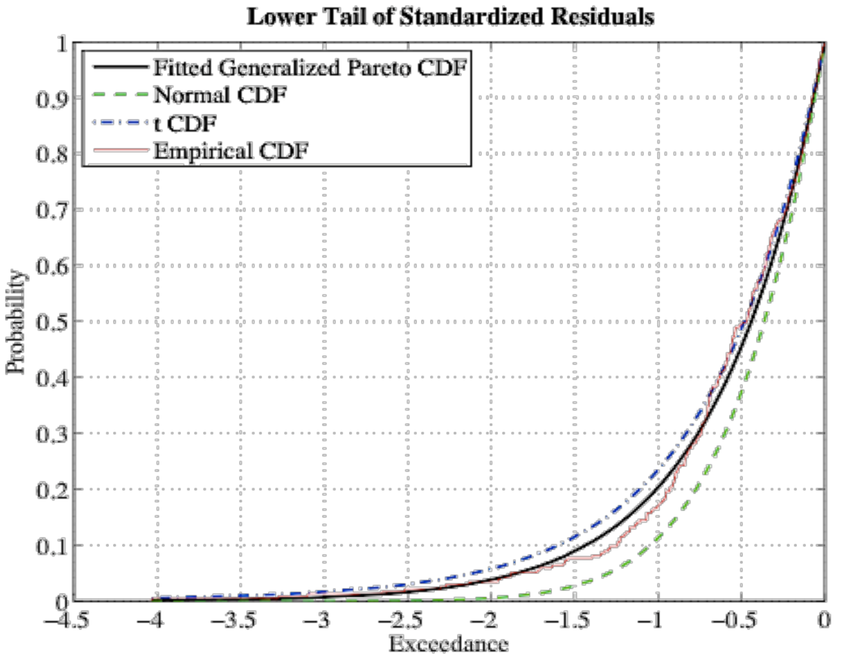
Second principal component

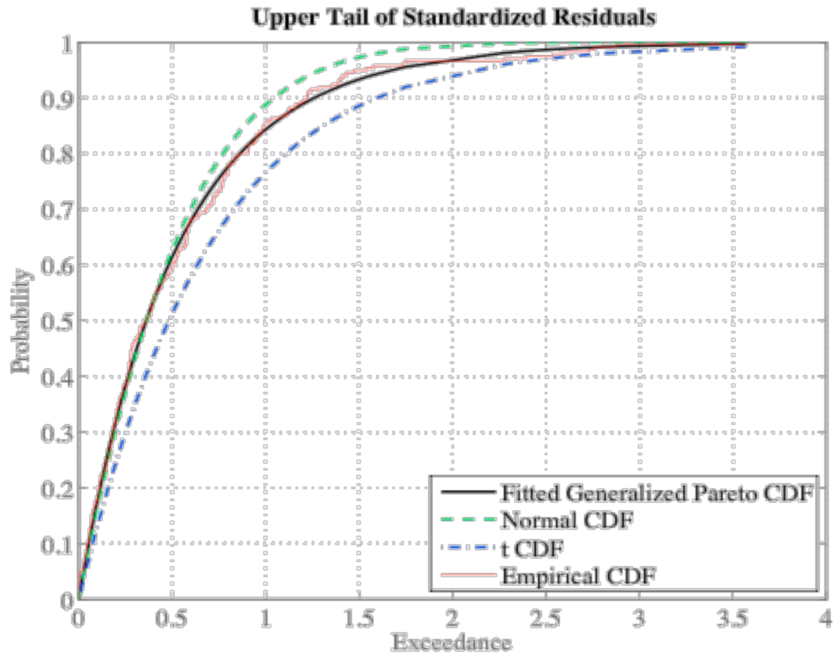




Tails of the standardized residuals

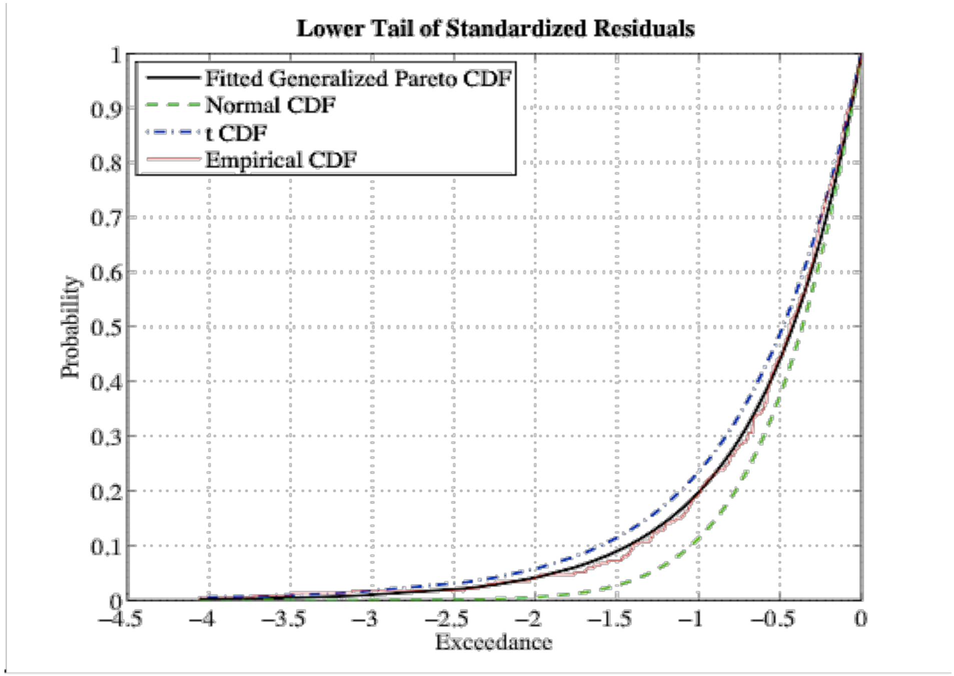
Third principal component

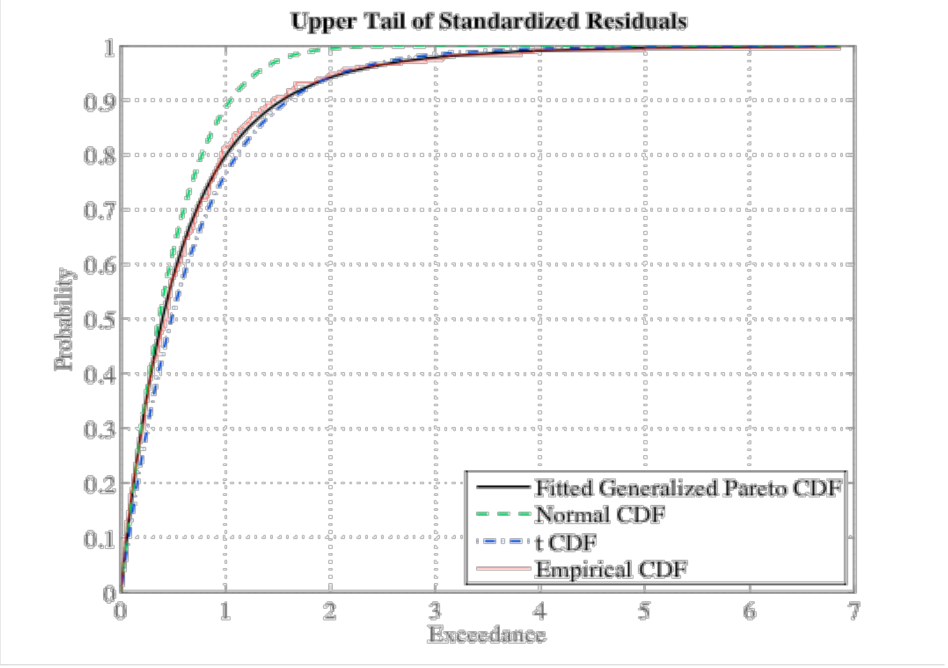




Tails of the standardized residuals

Fourth principal component





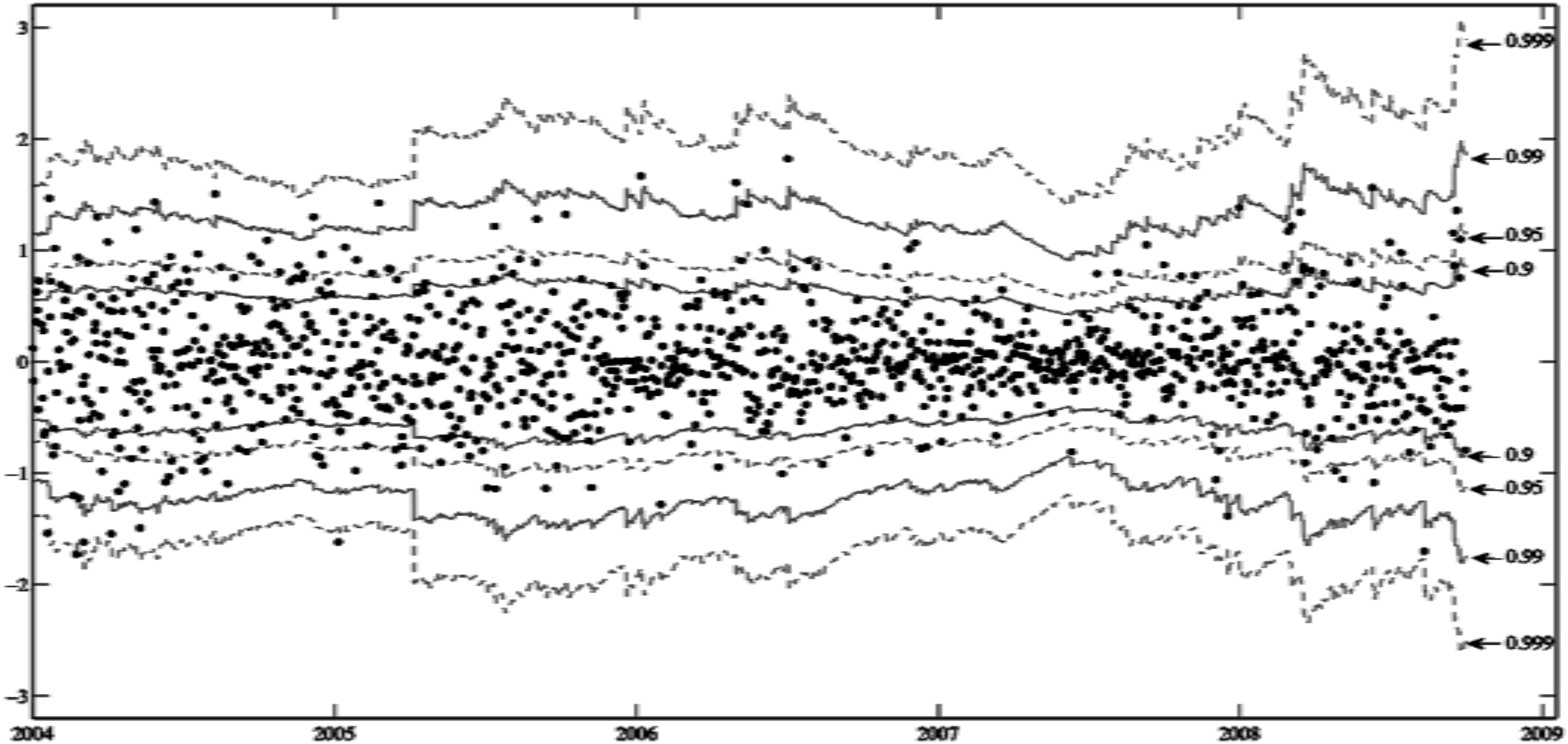
VaR and ES

- One-step-ahead forecasts for an equally-weighted portfolio (January 1, 2008)

Upper tail				
CL	0.90	0.95	0.99	0.999
VaR	0.5294	0.6947	1.0750	1.6101
ES	0.7284	0.8924	1.2694	1.7994

Lower tail				
CL	0.90	0.95	0.99	0.999
VaR	−0.4212	−0.5923	−0.9203	−1.2495
ES	−0.6790	−0.8251	−1.1026	−1.3743

Backtesting



Backtesting: Number of violations by quantiles

Upper tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	109	56	12	0
Normal	144	89	31	16
t	121	49	11	0
HS	235	166	70	21
Expected	123.9	61.95	12.39	1.239

Lower tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	88	48	9	1
Normal	118	67	19	8
t	86	40	7	0
HS	219	142	54	28
Expected	123.9	61.95	12.39	1.239

Backtesting: Kupiec test

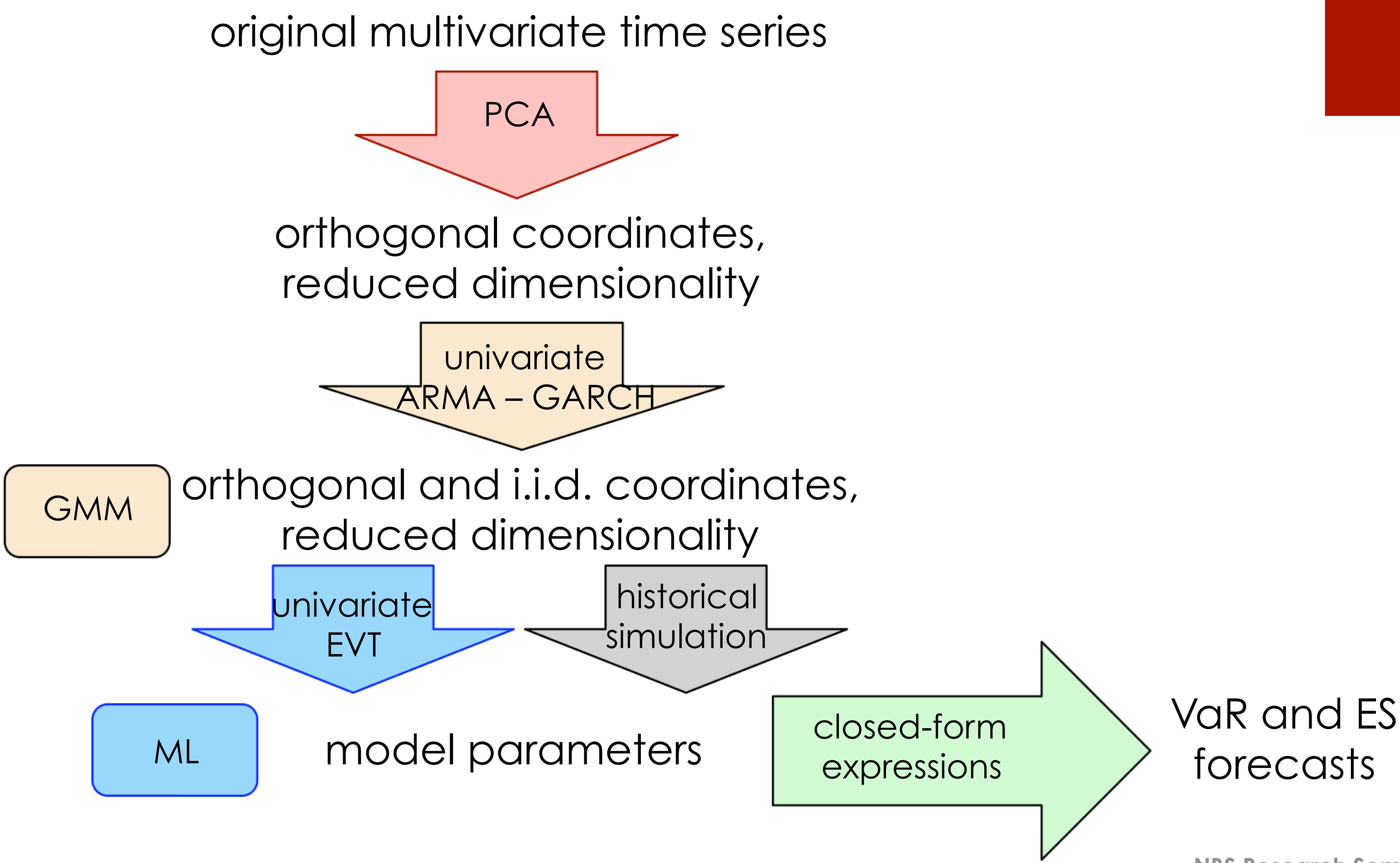
Upper tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	2.0665	0.6207	0.0125	2.4792
	$(< 10^{-4})$	$(< 10^{-4})$	$(< 10^{-4})$	−−
Normal	3.4620	11.0174	19.9238	52.5198
	$(< 10^{-4})$	(0.4564)	(0.0626)	(~ 1.0)
t	0.0759	3.0602	0.1637	2.4792
	$(< 10^{-4})$	$(< 10^{-4})$	(0.2022)	−−
HS	90.1083	128.6208	129.9539	79.6643
	$(< 10^{-4})$	(0.0142)	(~ 1.0)	(~ 1.0)

Lower tail				
Method	Number of violations			
	CL = 0.90	CL = 0.95	CL = 0.99	CL = 0.999
EVT	12.7273	3.5725	1.0354	0.0494
	$(< 10^{-4})$	$(< 10^{-4})$	(0.0006)	(0.1760)
Normal	0.3167	0.4226	3.0626	16.3572
	$(< 10^{-4})$	$(< 10^{-4})$	$(< 10^{-4})$	(0.9625)
t	14.2719	9.3110	2.8099	2.4792
	$(< 10^{-4})$	$(< 10^{-4})$	(0.0980)	−−
HS	67.6349	81.0498	77.1940	121.6632
	$(< 10^{-4})$	$(< 10^{-4})$	(0.9791)	(~ 1.0)

Backtesting: Pearson's test

Method	Lower tail	Upper tail
EVT	0.4170	1.3142
	(0.0189)	(0.1410)
Normal	38.5252	7.5773
	($\sim$ 1.0)	(0.8917)
t	2.2298	2.0934
	(0.3064)	(0.2814)
HS	123.1067	146.7974
	($\sim$ 1.0)	($\sim$ 1.0)

Summary of the method



Conclusion

- The proposed approach employs the notion that some key results of the univariate EVT can be applied separately to a set of orthogonal i.i.d. random variables.
- Such random variables can be constructed from the principal components of ARMA-GARCH conditional residuals of a multivariate return series.
- The estimation is free of any unnecessary distributional assumptions.
- The proposed approach tends to yield more precise VaR and ES forecasts than the usual methods based on conditional normality, conditional t-distribution or historical simulation, without losing efficiency.