Emerging Markets Review 20 (2014) 23–41

Contents lists available at ScienceDirect

Emerging Markets Review

journal homepage: www.elsevier.com/locate/emr

VaR performance during the subprime and sovereign debt crises: An application to emerging markets

Esther B. Del Brio [a], Andrés Mora-Valencia [b], Javier Perote [a,b,c,*]

[a] Faculty of Economics and Business, Department of Business Economics, University of Salamanca, Spain
[b] School of Economics and Finance, Department of Finance, EAFIT University, Colombia
[c] Faculty of Economics and Business, Department of Economics, University of Salamanca, Spain

ARTICLE INFO

Article history:
Received 18 July 2013
Received in revised form 9 May 2014
Accepted 15 May 2014
Available online 22 May 2014

JEL classification:
G17

Keywords:
Value-at-risk
Backtesting
Skewed Student's t
Extreme value theory
Gram–Charlier expansion
Hedge funds

ABSTRACT

Highly volatile scenarios, such as those provoked by the recent subprime and sovereign debt crises, have questioned the accuracy of current risk forecasting methods. This paper adds fuel to this debate by comparing the performance of alternative specifications for modeling the returns filtered by an ARMA-GARCH: Parametric distributions (Student's t and skewed-t), the extreme value theory (EVT), semi-nonparametric methods based on the Gram–Charlier (GC) expansion and the normal (benchmark). We implement backtesting techniques for the pre-crisis and crisis periods for stock index returns and a hedge fund of emerging markets. Our results show that the Student's t fails to forecast VaR during the crisis, while the EVT and GC accurately capture market risk, the latter representing important savings in terms of efficient regulatory capital provisions.

1. Introduction

Recent literature has extensively focused on studying the causes and consequences of the subprime and sovereign debt crisis (Kolb, 2010; Shiller, 2008). The origin of the crisis was the sharp decline of the U.S. housing prices in 2006, triggered by the enormous amount of subprime mortgages contracted in a decade characterized by low interest rates and irrational expectations about the sustainability of real estate market prices. The crisis was amplified by different mechanisms such as the highly leveraged banks,

* Corresponding author at: University of Salamanca, Campus Miguel de Unamuno (Edif. FES), 37007 Salamanca, Spain. Tel.: +34 923 294400x3512.

E-mail address: perote@usal.es (J. Perote).

http://dx.doi.org/10.1016/j.ememar.2014.05.001
1566-0141/

the deregulation of the financial system, the growth of securitization, the misbehavior of rating agencies and the so-called 'credit crunch', and it caused the bankruptcy of many investment banks (Bear Stearns and Lehman Brothers) and the bailout of insurance companies (AIG). This situation caused panic in the financial markets, dramatically increased systemic risk and turned into a global financial crisis and a major recession (Mishkin, 2011).

In an attempt to restore the trust in the financial system and impulse the recovery, governments implemented expansive fiscal policies and programs to bail out banks and buy toxic assets. These policies sharply increased public deficits and risk premia in different economies, triggering the European sovereign public debt crisis that rapidly spread to the world economy. Therefore, the subprime and sovereign debt crises had their origin in the United States and the European Union economies but had an impact on emerging markets (EMs hereafter). In this paper we focus on studying the impact of both crises on the EM volatility and risk and testing the performance of alternative methodologies to measure it (Huang and Tseng, 2009).

From a risk management perspective, the crisis increased volatility in the financial markets and demanded new methodologies capable of accurately estimating the regulatory capital of financial institutions. Risk measures are commonly obtained by quantile-based methods (Dowd and Blake, 2006), and among these the most widely used is the Value-at-Risk (hereafter, VaR). Different approaches have been used to compute VaR (Jorion, 2006; Kuester et al., 2006) but there is no consensus on the most appropriate methodology. Former VaR models have been criticized because the normality assumption involves risk underestimation, and thus skewed and heavy-tailed distributions have been proposed (Bali and Theodossiou, 2008; Harmantzis et al., 2006).

The current paper expands on this issue by comparing the performance of VaR forecasts obtained by the normal distribution (benchmark) to four natural candidates that account for the heavy tails of stock returns: the Student's t, a skewed variant of the Student's t distribution (Hansen, 1994), the extreme value theory (EVT) approach (Embrechts et al., 1997; and McNeil et al., 2005) and the semi-nonparametric approach based on the Gram–Charlier (GC) density, which is an expansion around the normal density allowing for skewness and excess kurtosis (Edgeworth, 1907).

The VaR forecasting performance of the models is analyzed for the high volatility scenario of the recent subprime and sovereign debt crises. Furthermore, we compare how VaR measures are affected by the occurrence of extreme events in the areas where the crises started but also on different EM. For this purpose, we analyze six leading world stock indices (MSCI Europe, MSCI USA, MSCI EM, MSCI EM Latin America, MSCI EM Europe and MSCI EM Asia) as well as a hedge fund on EM (Dow Jones Credit Suisse EM). For these indices, historical daily losses are compared to the maximum loss forecasted for each method considering a one-day-ahead horizon. VaR forecasts are computed by assuming an ARMA–GARCH model for the conditional mean-variance and estimating the quantile of the assumed distribution at the 99% confidence level. According to this backtesting procedure (Zumbach, 2006), it is expected that for 1% of the cases (days of the sample) the historical losses will fall outside the estimated VaR. This idea allows a straightforward implementation of VaR backtesting or forecasting performance tests (Christoffersen, 2003). As we focus on measuring the impact of the recent crises on the VaR methodology performance, the backtesting period is divided into two subperiods: (i) pre-crisis and (ii) crisis, the latter including the subprime and the sovereign debt crises periods.

The results show that both normal–GARCH and Student's t–GARCH are inadequate for high confidence levels and high volatility periods, although the skewed Student's t– (skewed-t hereafter) GARCH outperforms the Student's t–GARCH. The GC–GARCH and EVT–GARCH, however, produce accurate VaR forecasts in these contexts. Therefore the optimal VaR model depends not only on the assumed confidence level but also on the market conditions observed. The main contribution of the paper is then the comparison of a wide variety of VaR forecasting methods with the scarcely used GC–GARCH during the recent crisis and, particularly, for EM stock indices. We also include a hedge fund on EM and analyze the savings incurred by fund managers in terms of the accuracy of regulatory capital provisions when using our proposed GC–GARCH model.

The rest of the paper is organized as follows: Section 2 presents the models and VaR methodology, Section 3 analyzes the data and the empirical results on VaR forecasting, and Section 4 summarizes the main results of the paper.

2. Models and methodology

Since Mandelbrot (1963), the normality assumption of stock returns is deemed inappropriate, revealing the following stylized empirical regularities (Cont, 2001): (1) a sharp peak at the mean; (2) heavy tails; (2) skewness; (4) volatility clustering; (5) slow decay in the autocorrelation function of the absolute returns. To account for the leptokurtosis implied by the first two properties, the use of non-Gaussian distributions is proposed, out of which the Student's t is the most widely used. However, incorporating asymmetries requires the use of other densities such as the skewed-t (Giot and Laurent, 2003). Alternatively, for purposes of measuring risk, the EVT has directly focused on reproducing the behavior at the tails. Within the EVT framework, different approaches have been proposed, such as the generalized extreme value distribution or the generalized Pareto distribution (GPD). In this paper, we compare the VaR performance of both the Student's t and skewed-t to the so-called peaks over threshold (POT) method, which is based on the GPD (Davison and Smith, 1990). Furthermore, we also incorporate the semi-nonparametric estimation that is based on the asymptotic properties of the GC type A series when approximating a frequency function (see Kendall and Stuart, 1977, p. 168—72). Most of the financial literature about semi-nonparametric methodologies is devoted to price derivatives, following Jarrow and Rudd (1982). Only a few papers focus on the Gram–Charlier application to risk management (Mauleón, 2003; Mauleón and Perote, 2000).

On the other hand, stock returns also seem to have a small predictable component in the conditional mean that has traditionally been modeled according to simple ARMA structures. Nevertheless squared returns exhibit particular dynamics (conditional heteroskedasticity, volatility clustering and long memory) that have been extensively studied since Engle (1982) and Bollerslev (1986) introduced ARCH and GARCH models. As we focus on VaR performance due to the distributional hypotheses, the model implemented in this paper incorporates an ARMA(1,1) and a GARCH(1,1) for modeling the conditional mean (μ_t) and variance (σ_t^2), respectively, in accordance with the common use in risk management applications. We define the complete model in Eqs. (1) to (4) below.

$$r_t = \mu_t + \sigma_t z_t, \tag{1}$$

$$\mu_t = \varphi + \phi\mu_{t-1} + \theta\varepsilon_{t-1} + \varepsilon_t, \tag{2}$$

$$z_t = \varepsilon_t/\sigma_t, z_t \sim G(0,1), \tag{3}$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{4}$$

where φ, ϕ and θ are the parameters of ARMA(1,1); ω, α and β the parameters of GARCH (1,1); and ε_t is white noise. For the sake of comparison, different standardized (i.e. zero mean and unit variance) density specifications are considered for G. In particular, we consider the four following probability density functions (pdfs).

(i) The normal pdf:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \tag{5}$$

(ii) The Student's t pdf:

$$t(z) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{z^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \tag{6}$$

where Γ is the gamma function and ν is the degrees of freedom parameter.

Table 1
Descriptive statistics for stock index returns.

	Total period	Pre-crisis period	Crisis period
Panel A: MSCI Europe			
Mean	0.0025	0.0079	−0.0028
Median	0.0322	0.0311	0.0326
Standard deviation	1.4648	1.1567	1.7189
Excess kurtosis	6.1617	2.3405	5.5684
Skewness	−0.0963	−0.2128	−0.0492
Minimum	−10.1783	−5.6814	−10.1783
Maximum	10.6979	5.1938	10.6979
Panel B: MSCI USA			
Mean	0.0036	−0.0071	0.0143
Median	0.0258	0.0000	0.0489
Standard deviation	1.3131	1.1375	1.4681
Excess kurtosis	7.8484	2.7571	8.8208
Skewness	−0.1829	0.0964	−0.3154
Minimum	−9.5137	−6.1609	−9.5137
Maximum	11.0426	5.6104	11.0426
Panel C: MSCI EM			
Mean	0.0268	0.0323	0.0213
Median	0.1094	0.1128	0.0957
Standard deviation	1.2850	1.0174	1.5060
Excess kurtosis	7.7059	2.3419	7.2308
Skewness	−0.5113	−0.6116	−0.4406
Minimum	−9.9944	−5.8312	−9.9944
Maximum	10.0729	4.0564	10.0729
Panel D: MSCI EM Latin America			
Mean	0.0423	0.0577	0.0269
Median	0.1019	0.1221	0.0781
Standard deviation	1.7858	1.4250	2.0853
Excess kurtosis	9.3492	1.8679	9.2117
Skewness	−0.4225	−0.3578	−0.4046
Minimum	−15.0600	−7.5013	−15.0600
Maximum	15.3531	5.8709	15.3531
Panel E: MSCI EM Europe			
Mean	0.0246	0.0551	−0.0059
Median	0.1379	0.1594	0.0889
Standard deviation	1.9751	1.5888	2.2974
Excess kurtosis	10.4955	3.0064	10.3593
Skewness	−0.4296	−0.5310	−0.3496
Minimum	−19.9242	−8.8964	−19.9242
Maximum	18.6010	7.0912	18.6010
Panel F: MSCI EM Asia			
Mean	0.0198	0.0154	0.0241
Median	0.0731	0.0626	0.0959
Standard deviation	1.4406	1.2808	1.5848
Excess kurtosis	5.2376	2.2548	6.0090
Skewness	−0.3065	−0.3619	−0.2723
Minimum	−8.6225	−7.5269	−8.6225
Maximum	12.6541	5.4962	12.6541
*Panel G: DJCSEM HF**			
Mean			0.0197
Median			0.0564
Standard deviation			0.5958
Excess kurtosis			9.7667
Skewness			−0.6615
Minimum			−4.6737
Maximum			4.1343

(iii) The skewed-t pdf by Fernández and Steel (1998):

$$g(z) = \begin{cases} -\dfrac{2}{\gamma+\dfrac{1}{\gamma}} t(\gamma z) & \text{for} \quad x<0, \\ \dfrac{2}{\gamma+\dfrac{1}{\gamma}} t\left(\dfrac{z}{\gamma}\right) & \text{for} \quad x \geq 0, \end{cases} \tag{7}$$

where γ is the shape parameter, which incorporates the skewness, and $t(z)$ is the Student's t pdf in Eq. (6).

(iv) The GC Type A density is given by:

$$f(z, \boldsymbol{d}) = \left(1 + \sum\nolimits_{s=1}^{n} d_s H_s(z)\right) \phi(z), \tag{8}$$

where $\phi\ (z)$ denotes the normal pdf in Eq. (5), $\boldsymbol{d}' = (d_1, \ldots, d_n) \in \mathbb{R}^n$ and H_s is the Hermite polynomial (HP) of order s, which can be defined in terms of the derivatives of $\phi\ (z)$ as

$$\frac{d^s \phi(z)}{dz^s} = (-1)^s H_s(z) \phi(z). \tag{9}$$

In particular, the first eight HP are: $H_1(z) = z$, $H_2(z) = z^2 - 1$, $H_3(z) = z^3 - 3z$, $H_4(z) = z^4 - 6z^2 + 3$, $H_5(z) = z^5 - 10z^3 + 15z$, $H_6(z) = z^6 - 15z^4 + 45z^2 - 15$, $H_7(z) = z^7 - 21z^5 + 105z^3 - 105z$, and $H_8(z) = z^8 - 28z^6 + 210z^4 - 420z^2 + 105$.

These polynomials form a normal basis and thus satisfy, the orthogonality property,

$$\int H_s(z) H_j(z) \phi(z) dz = 0 \forall s \neq j, \tag{10}$$

which is the basis to proof that the expansion integrates one and that the even non-central moments depend on the even d_s parameters (e.g. d_2 accounts for the variance, d_4 for the excess kurtosis and the rest of the even parameters capture higher order moments). Furthermore the skewness of the distribution is incorporated in the odd parameters. A well-known shortcoming of this expansion is that it is not necessarily positive for all the values of the parameters, which calls for the implementation of positive transformations (Gallant and Nychka, 1987; León et al., 2009) or positivity constraints (Jondeau and Rockinger, 2001). However, we use the original GC Type A expansion in Eq. (8), which is simpler and more useful for VaR applications.

Most of the empirical studies that implement GC density in finance (mainly for option pricing) truncate the expansion in $n = 4$ and employ only two terms of the expansion, d_3 and d_4. We initially follow this strategy and estimate the corresponding GC density by the method of moments (GC–MM) and maximum likelihood (GC–ML1). For the sake of comparison, we also implement larger expansions (up to the eighth order) estimated by maximum likelihood (GC–ML2). In this latter case the "optimal" truncation order is identified by using the Akaike Information Criterion (AIC). Furthermore, we follow the procedure proposed by Del Brio and Perote (2012) to estimate the density parameters in two steps. First, we estimate the conditional mean and variance using quasi maximum likelihood (QML) and obtain the standardized residuals; second, we estimate the d_s parameters for the standardized residuals. For the GC–ML1 and GC–MM models, the estimates are obtained by maximizing the log-likelihood function and applying directly Eqs. (39)–(46) from Del Brio and Perote (2012, p. 534-5), respectively.

Notes to Table 1:
This table displays descriptive statistics for the series of MSCI indices (Europe, MSCI and EM) and indices of different EM areas (Latin America, Europe and Asia) for the pre-crisis period (December 1997–July 2006) and the crisis period (July 2006–March 2013). For the crisis period the table includes an EM Hedge Fund index (DJCSEM HF).

* DJCSEM HF series comprises data from July 2006 to March 2013.

In this model, the estimated VaR with a confidence level α is computed as the estimated α-quantile, $\hat{q}_\alpha(z)$, of the assumed G distribution. Therefore, the predicted VaR for the variable r at the time horizon $t+1$ and with confidence level α is given in Eq. (11).

$$\mathrm{VaR}^{\alpha}_{t+1} = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\hat{q}_\alpha(z), \tag{11}$$

where $\hat{\mu}_{t+1}$ and $\hat{\sigma}_{t+1}$ are the predictions for the mean and standard deviation conditioned by the available information at time t, Ω_t, based on the ARMA–GARCH model described in Eqs. (2) and (4).

Alternatively, VaR can also be computed by the EVT methodology. Particularly, we implement the two-step procedure proposed in McNeil and Frey (2000). In the first step, the ARMA(1,1)–GARCH(1,1) model is fitted using QML and in the second step the so-obtained standardized residuals are used to implement the POT methodology using 10% of the tail of the distribution as the threshold. Thus, $\hat{q}_\alpha(Z)$ is given by

$$\hat{q}_\alpha(Z) = u + \frac{\delta}{\xi}\left(\left(\frac{1-\alpha}{N_u/n}\right)^{-\xi} - 1\right), \tag{12}$$

where u is the estimated threshold, N_u is the number of exceedances over the threshold, n is the sample size (thus N_u/n is a non-parametric estimator of the empirical distribution tail) and δ and ξ are the scale and shape parameters of the GPD. The cumulative distribution function (cdf) of the GPD distribution is given by

$$F(x) = \begin{cases} 1-(1+\xi x/\delta)^{-1/\xi}, & \xi \neq 0, \\ 1-e^{-x/\delta}, & \xi = 0. \end{cases} \tag{13}$$

The weakness of the EVT lies in the threshold selection, which involves a tradeoff between bias and variance in the estimation of the parameters, especially the shape parameter ξ. This parameter can be estimated by bootstrapping or by graphical techniques, but there is no optimal method to choose the appropriate threshold. Some empirical studies have shown that a good approximation is to choose 5% or 10% of the data in the tail of the distribution as a threshold. We adopt the latter, as in McNeil and Frey (2000).

3. Empirical application

3.1. Data and in-sample results

In this Section, we compare the performance of the above-mentioned models for computing VaR for different world stock indices: MSCI EM, MSCI EM Latin America (EMLA), MSCI EM Europe (EM Europe), MSCI EM Asia (EM Asia), MSCI Europe and MSCI USA, and in two different volatility scenarios — which we call the pre-crisis and crisis periods. Moreover, we analyze the Dow Jones Credit Suisse EM Hedge Fund (DJCSEM HF) Index for the crisis period, thus analyzing the applicability of our results to real investment funds. Most of the EM Hedge Fund indices are available on a monthly frequency. Then, the style attribution technique (Sharpe, 1992) – which consists in identifying the styles that may be attributed with the main risk factors in a portfolio – is applied to obtain the data on a daily basis. Particularly, we used the MSCI Asia Pacific Index, MSCI EM Latin America Index, MSCI North America Index, MSCI EM Europe, Middle East and Africa Index, and US Cash Indices LIBOR Total Return 3 month as factors.

All the data were obtained from Datastream and Bloomberg (for further details on the data see Appendix A). The sample comprises almost 16 years of daily data from December 1997 to the first quarter of 2013 for all data, with the exception of the DJCSEM HF index series that ranges from August 2004 to the first quarter of 2013 (approximately 8 years). We split this sample into two sub-samples and choose July 2006 as the crisis starting date, one year before the date when Bear Stearns hedge funds reported massive losses. Table 1 displays the descriptive statistics for continuously compounded returns of these series, defined as $r_t = 100\ \log(P_t/P_{t-1})$, where P_t represents the corresponding price index. Panels A–G gather the main statistics for each of the series.

Descriptive statistics show that the mean return for all the EM indices and the European stock index is positive in the pre-crisis period, but for the latter and MSCI EM Europe becomes negative in the crisis period. The measures of dispersion reveal that the volatility is much higher in the crisis period, almost double in contrast with the pre-crisis period, but also that the European and Latin American EM are much more volatile in the crisis period. The DJCSEM HF, however, is much less volatile than the other indices. This is not only due to the fact that the data correspond to a hedge fund but also to the different period under analysis. Regarding the skewness, the MSCI USA seems to be positive in the pre-crisis period and negative in the crisis period, which means that lower (higher) returns were more likely to be obtained in the pre-crisis (crisis) period. The MSCI Europe and all the EM indices, however, exhibit negative skewness in both periods. The deviations of the median from the mean and the values of the excess kurtosis justify the use of VaR based on skewed and leptokurtic distributions.

Next we proceed to select our model among the three plausible models for conditional mean – white noise, AR(1) and ARMA(1,1) – according to accuracy criteria. Table 2 shows the log-likelihood values of these three models combined with a GARCH(1,1) for modeling conditional variance and under different distributional hypotheses, either normal (Panel A), Student's t (Panel B) or skewed-t (Panel C). The results show clear evidence in favor of the autoregressive models and among them we select the latter ARMA(1,1) since it has slightly higher log-likelihood values and it nests the AR(1).

Table 3 presents the Maximum Likelihood (ML) estimates of the parameters of the ARMA(1,1)–GARCH(1,1) model under the three distributional hypotheses: Normal (Panel A), Student's t (Panel B) and Skewed-t (Panel C). P-values for testing the significance of each parameter are given in parentheses. The values for the ARMA structure (ϕ and θ) are insignificant for the EM series, which is in line with the 'small predictable component of the conditional mean' stylized fact featured by stock returns. The values of the GARCH(1,1) parameters, however, are statistically significant and $\alpha + \beta$ is close to one for all series, reflecting the persistence and clustering in volatility. The estimates for the shape (degrees of freedom), ν, and skew parameter (γ) of the Student's t distributions are also displayed. Both parameters are significant, which shows that the distribution is leptokurtic and asymmetric.

The parameters of the EVT and the semi-nonparametric VaR methodologies implemented in the paper are displayed in Table 4. In these cases, two-step estimation was implemented as in Del Brio and Perote (2012), i.e. returns were filtered according to the ARMA(1,1)–GARCH(1,1) estimated in the first step by QML. The shape parameter for EVT, ξ in Panel A, is only significantly different from zero at a 5% confidence for the USA index. This means that, after filtering the returns by an ARMA(1,1)–GARCH(1,1) model, the standardized residuals from most data exhibit medium-tailed distributions (note that if $\xi = 0$ then GPD becomes the exponential distribution). Regarding the GC densities three alternative models are estimated:

Table 2
Log-likelihood for different conditional mean-variance models and under different distributional assumptions.

	Europe	USA	EM	EMLA	EM Europe	EM Asia	DJCSEM HF
Panel A: Normal							
GARCH(1,1)	−6358.32	−5908.27	−6052.28	−7378.11	−7782.00	−6734.74	−1425.33
AR(1)–GARCH(1,1)	−6358.10	−5904.80	−5944.01	−7347.41	−7772.13	−6707.52	−1385.77
ARMA(1,1)–GARCH(1,1)	−6356.58	−5899.51	−5943.02	−7346.17	−7772.08	−6707.29	−1383.98
Panel B: Student's t							
GARCH(1,1)	−6337.14	−5833.96	−6002.94	−7290.13	−7673.15	−6682.24	−1387.08
AR(1)–GARCH(1,1)	−6336.95	−5829.49	−5899.35	−7260.94	−7666.76	−6659.31	−1349.28
ARMA(1,1)–GARCH(1,1)	−6336.81	−5822.88	−5897.60	−7259.58	−7666.76	−6659.15	−1347.34
Panel C: Skewed-t							
GARCH(1,1)	−6329.54	−5826.93	−5979.55	−7278.27	−7663.80	−6677.41	−1374.67
AR(1)–GARCH(1,1)	−6329.46	−5820.70	−5886.28	−7253.16	−7658.83	−6656.25	−1343.53
ARMA(1,1)–GARCH(1,1)	−6324.65	−5809.24	−5883.28	−7251.22	−7658.82	−6656.21	−1341.07

This table compares the log-likelihood of three mean-variance models (white noise, AR(1) and ARMA(1,1) combined with a GARCH(1,1) for capturing volatility). The comparison includes three different distributions for the disturbances (normal, Student's t or Skewed-t).

Table 3
Parameters of the ARMA(1,1)–GARCH(1,1).

	Europe	USA	EM	EMLA	EM Europe	EM Asia	DJCSEM HF
Panel A: Normal							
φ	0.0141	0.0111	0.0798	0.0970	0.0866	0.0681	0.0478
	(0.0374)	(0.0134)	(0.0000)	(0.0004)	(0.0134)	(0.0007)	(0.0000)
ϕ	0.7699	0.7682	0.1338	−0.0387	0.1551	0.2186	−0.0009
	(0.0000)	(0.0000)	(0.0975)	(0.7202)	(0.5537)	(0.1057)	(0.9934)
θ	−0.7896	−0.8124	0.1159	0.1744	−0.0823	−0.0968	0.2060
	(0.0000)	(0.0000)	(0.1549)	(0.1011)	(0.7567)	(0.4826)	(0.0508)
ω	0.0180	0.0139	0.0294	0.0719	0.0705	0.0282	0.0028
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)
α	0.0901	0.0783	0.1005	0.0989	0.0930	0.0902	0.1104
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.000)	(0.0000)
β	0.9017	0.9130	0.8793	0.8759	0.8881	0.8979	0.8838
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel B: Student's t							
φ	0.0836	0.0179	0.0913	0.1137	0.1164	0.0701	0.0542
	(0.0415)	(0.0019)	(0.0000)	(0.0000)	(0.0026)	(0.0007)	(0.0000)
ϕ	−0.3100	0.7198	0.0870	−0.0642	0.0685	0.1959	−0.0183
	(0.5733)	(0.0000)	(0.2712)	(0.5654)	(0.7861)	(0.1827)	(0.8632)
θ	0.3194	−0.7711	0.1506	0.1879	−0.0156	−0.0885	0.2108
	(0.5576)	(0.0000)	(0.0569)	(0.0869)	(0.9511)	(0.5538)	(0.0425)
ω	0.0163	0.0099	0.0257	0.0551	0.0571	0.0246	0.0027
	(0.0002)	(0.0008)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0023)
α	0.0849	0.0799	0.0952	0.0907	0.0897	0.0817	0.1023
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
β	0.9086	0.9169	0.8876	0.8914	0.8967	0.9079	0.8925
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ν	10	6.4203	8.7273	6.4819	5.9786	8.3116	6.8330
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel C: Skewed Student's t							
φ	0.0118	0.0127	0.0708	0.0833	0.0886	0.0628	0.0460
	(0.0119)	(0.0016)	(0.0002)	(0.0119)	(0.0112)	(0.0026)	(0.0000)
ϕ	0.7628	0.7160	0.0262	−0.1076	0.0048	0.1514	−0.0696
	(0.0000)	(0.0000)	(0.7447)	(0.3283)	(0.9859)	(0.3376)	(0.5320)
θ	−0.7962	−0.7821	0.2007	0.2238	0.0420	−0.0471	0.2473
	(0.0000)	(0.0000)	(0.0114)	(0.0377)	(0.8801)	(0.7679)	(0.0220)
ω	0.0165	0.0094	0.0239	0.0522	0.0546	0.0244	0.0025
	(0.0001)	(0.0008)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0021)
α	0.0808	0.0774	0.0885	0.0859	0.0847	0.0804	0.0964
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
β	0.9115	0.9183	0.8941	0.8959	0.9010	0.9087	0.8968
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ν	10	6.9753	9.2225	6.7944	6.2603	8.6431	7.3266
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
γ	0.8980	0.8945	0.8819	0.9139	0.9157	0.9475	0.9003
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

This table gathers the ML estimates of the ARMA(1,1)–GARCH(1,1) model with either normal, Student's t or Skewed-t disturbances. φ, ϕ and θ are the parameters for the ARMA(1,1); ω, α and β the parameters of the GARCH(1,1); ν is the degrees of freedom of the Student's t and γ the skew parameter of the Skewed-t. P-values for the *t*-test are in parentheses.

the GC is expanded to the fourth term and estimated either by the method of moments (Panel B) or maximum likelihood (Panel C), and the GC is expanded to the eighth order (Panel D). In all the cases, parameters d_3 and d_4 confirm the presence of (negative) skewness and leptokurtosis since the former captures skewness and the latter excess kurtosis. Nevertheless, not all the parameters of the larger expansion (GC–ML2) are significantly different from zero (mainly the odd parameters, since skewness is accurately captured by d_3). Despite this fact we keep both polynomial structures in order to compare the effect of the expansion length in the VaR forecasting performance in the next section.

Table 4
Parameters of the standardized EVT and GC.

	EUROPE	USA	EM	EMLA	EM Europe	EM Asia	DJCSEM HF
Panel A: EVT							
ξ	−0.0290	−0.1963	−0.0674	−0.0650	−0.0244	−0.0543	−0.0908
	(0.2861)	(0.0000)	(0.0670)	(0.0766)	(0.2968)	(0.0786)	(0.057)
δ	0.4949	0.5836	0.5427	0.5293	0.5327	0.5362	0.5401
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel B: GC–MM							
d_3	−0.0338	−0.0759	−0.0499	−0.0665	−0.0667	−0.0360	−0.0590
d_4	0.0264	0.0675	0.0421	0.0717	0.0789	0.0433	0.0492
Panel C: GC–ML1							
d_3	−0.0345	−0.0449	−0.0370	−0.0331	−0.0336	−0.0225	−0.0411
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0022)	(0.0000)
d_4	0.0235	0.0347	0.0305	0.0235	0.0445	0.0314	0.0380
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel D: GC–ML2							
d_3	−0.0242	−0.0651	−0.0405	−0.0446	−0. 0447	−0.0294	−0.0450
	(0.0035)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0023)	(0.0008)
d_4	0.0270	0.0285	0.0417	0.0572	0. 0635	0.0410	0.0505
	(0.0000)	(0.0012)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
d_5	0.0056	−0.0150	−0.0031	−0.0075	−0. 0073	−0.0048	−0.0038
	(0.0418)	(0.0028)	(0.1944)	(0.0358)	(0.0358)	(0.1124)	(0.2391)
d_6	0.0021	0.0000	0.0049	0.0073	0. 0085	0.0041	0.0052
	(0.1232)	(0.5130)	(0.0095)	(0.0008)	(0.0009)	(0.0231)	(0.0401)
d_7	0.0000	−0.0027	−0.0008	−0.0009	−0. 0008	−0.0002	−0.0004
	(0.4540)	(0.0003)	(0.0598)	(0.0668)	(0.0668)	(0.3315)	(0.3096)
d_8	0.0004	0.0008	0.0005	0.0008	0.0009	0.0004	0.0007
	(0.0186)	(0.0032)	(0.0212)	(0.0004)	(0.0004)	(0.0163)	(0.0213)

This table displays the parameter estimates of the EVT–POT (Panel A), and alternative models for a GC expansion (Panels B, C and D). Panels B and C estimate two-parameter (skewness–excess of kurtosis) expansions either by MM or ML and Panel D a larger GC expansion. δ and ξ are the scale and shape parameters of the GPD and d_2–d_8 the parameters of the GC. P-values of the t-tests are in parentheses.

3.2. Backtesting

VaR forecasts are very sensitive to features such as the model employed, the volatility scenario, the confidence level or the time horizon considered. In this paper, we focus on the two first issues and assume a 99% confidence level (the common use for financial institutions) and a one-step-ahead (1-day) VaR. However, the computation of VaR for longer time horizons (h-day VaR) can be directly obtained by the so-called square-root-of-time rule (Daníelsson and Zigrand, 2006; Diebold et al., 1997), the α-root rule (de Vries, 2000) or the modified-squared-root rule (Wang et al., 2011), but at the cost of a loss in terms of misspecification. At the end of this Section, we provide an example of the performance of multiperiod VaR forecasting models.

Given a time horizon, a confidence level and a volatility scenario, VaR figures are clearly affected by the model specification. Fig. 1 shows an illustration of the distribution tails and the corresponding 99% 1-day VaR obtained for the standardized series (i.e. with zero mean and unit variance after filtering an ARMA(1,1)–GARCH(1,1) model) of the EM Europe under the normal, Student's t (with 5.9786 degrees of freedom) and the GC (see the estimates of the on Table 4). In particular the maximum expected loss for the next day will be 2.326%, 3%, and 3.146%, if we assume a normal, Student's t and GC, respectively. It is noteworthy that the tail of the GC is not strictly decreasing and presents thicker tails than the Student's t (in the area where outliers are likely to happen) and converges faster to zero than the Student's t. This is the reason why, unlike the Student's t, all moments of the GC exist even though it accounts more accurately for risk.

To test the validity of the different distributional assumptions on the returns of the indices (normal, Student's t, skewed-t, GC and EVT), the historical series, $r_1,\ldots,r_m$, are compared to the one-step-ahead

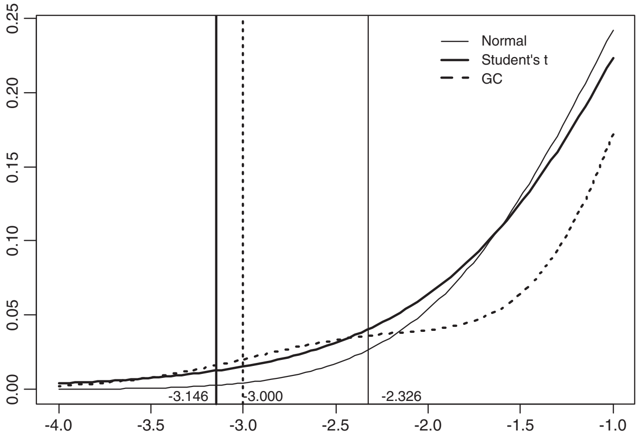

Fig. 1. VaR at 99% for the normal, Student's t and GC distribution. This figure illustrates the left tail of the distribution of the returns of the EM Europe and the corresponding VaR figures obtained under three different distributions (normal, Student's t and GC).

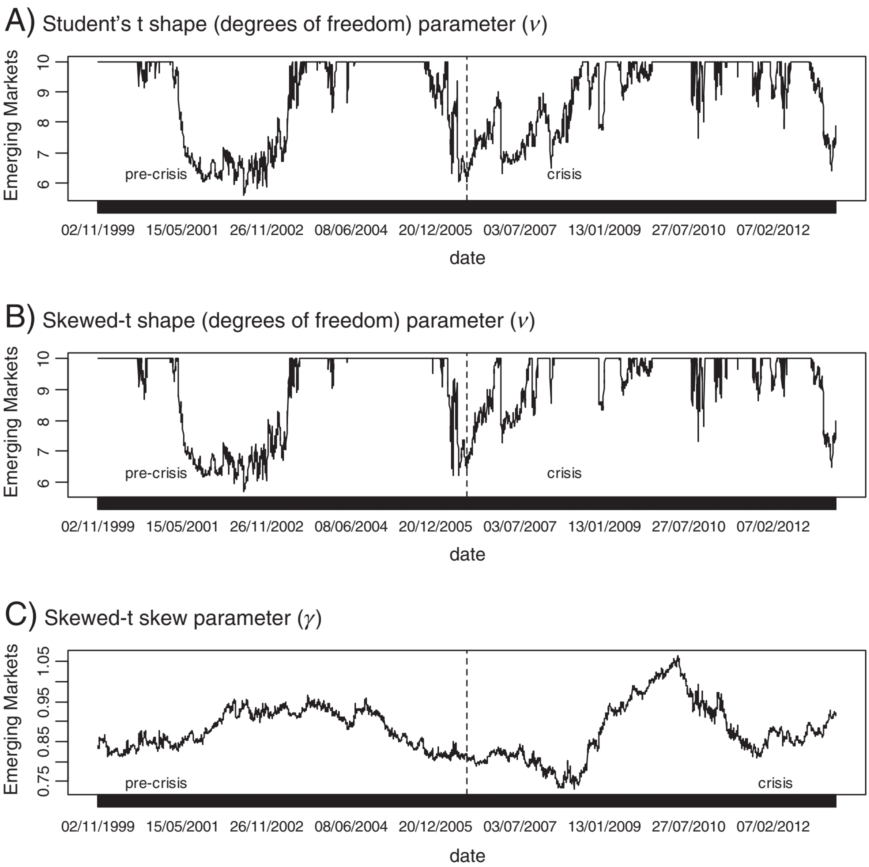

Fig. 2. Student's t distributions time-varying parameters for MSCI EUROPE index returns. A. Student's t shape (degrees of freedom) parameter (ν); B. Skewed-t shape (degrees of freedom) parameter (ν); C. Skewed-t skew parameter (γ). This figure shows the dynamics of estimated degrees of freedom (ν) and skew (γ) parameter for the Student's t (A) and Skewed-t (B and C) used in the backtesting. Dotted lines divide the pre-crisis and crisis period.

predicted VaR for the day $t = \{n + 1,\ldots,m\}$ and the confidence α, VaR_t^α. For this purpose we use a rolling window for estimation that includes the n previous days. The model performance is tested on different volatility scenarios (pre-crisis and crisis period) by implementing a backtesting procedure. We consider a time window of 500 days for computing every one-step-ahead VaR prediction and a total period of 3500 days as the backtesting or out-of-sample period. The backtesting period is divided into two identically sized sub-periods of 1750 days each: the pre-crisis period (November 1999–July 2006) and the crisis period (July 2006–1st quarter of 2013). For the DJCSEM HF case only the crisis period is analyzed. The predicted VaR is compared to the observed return at the 99% confidence level, thus we expect that in 1% of the backtesting days the negative returns will exceed VaR predictions. These values are referred to as 'violations' or 'exceptions'. More specifically, if I_t is the indicator function defined in Eq. (14),

$$I_t := 1_{\{r_{t+1} > \text{VaR}_t^\alpha\}}, \tag{14}$$

a 'violation' occurs when $r_{t+1} > \text{VaR}_t^\alpha$, and then the indicator function takes value 1 (otherwise, I_t takes value 0). Therefore, if the VaR methodology is adequate, it is expected that the violation indicator function values will behave as realizations of independent and identically distributed (iid) Bernoulli experiments with success probability equal to $1 - \alpha$, i.e. $\sum_{t=1}^{m} I_t \sim \text{Bin}(m, 1 - \alpha)$. Thus the null hypothesis that 'the model adequately estimates VaR' can be tested by a straightforward one-sided binomial test. The alternative hypothesis suggests that the method underestimates or overestimates the VaR depending on the number of expected violations.

The performance of every distribution depends on its flexibility to incorporate the new volatility scenario extreme events. Figs. 2–5 show how density parameters adapt to the time-varying scenario for the MSCI EM index (the corresponding figures for the other indices present similar patterns). In particular, Fig. 2A and B display the changes of the degrees of freedom parameter of Student's t and skewed-t over time. Note that the Y-axis is truncated at 10 since for bigger values the distribution is not very different from the standard normal. It is clear that this parameter decreases (increases) as the volatility increases (decreases) but it remains above 4 so that kurtosis is still well-defined. Fig. 2C displays the evolution of the

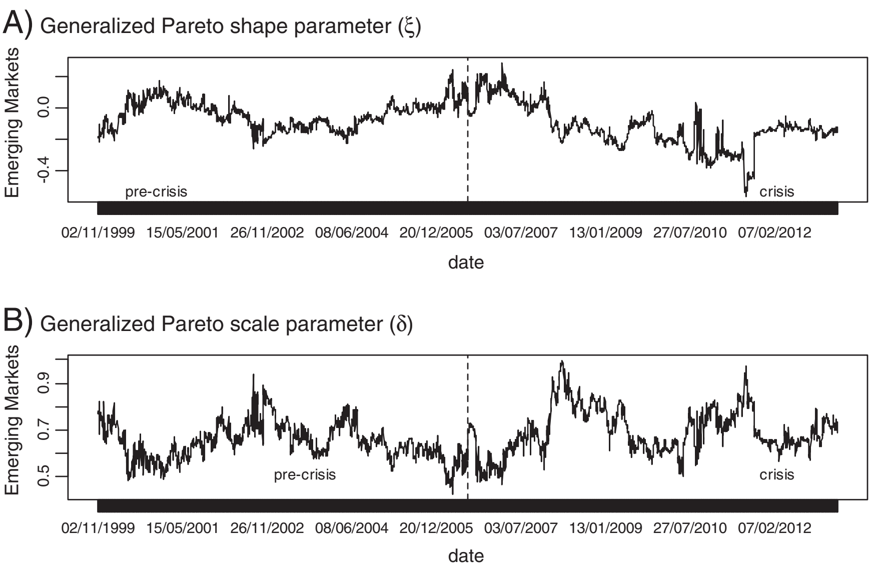

Fig. 3. Generalized Pareto time-varying parameters for MSCI EUROPE index returns. A. Generalized Pareto shape parameter (ξ); B. Generalized Pareto scale parameter (δ). This figure shows the dynamics of the estimates for the GDP shape (ξ) and scale (δ) parameter used in the backtesting for the ETV methodology. Dotted lines divide the pre-crisis and crisis period.

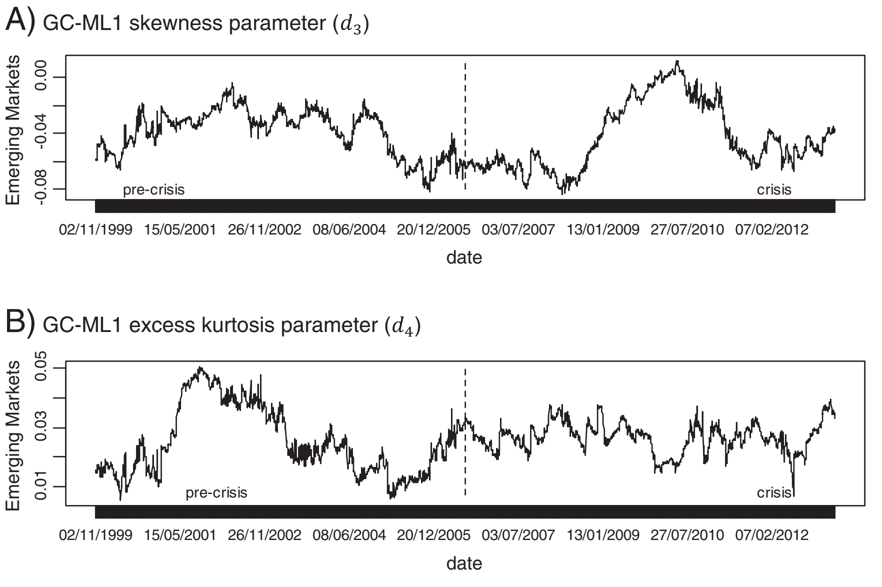

Fig. 4. GC–ML1 time-varying parameters for MSCI EUROPE index returns. A. GC–ML1 skewness parameter (d_3); B. GC–ML1 excess kurtosis parameter (d_4). This figure shows the dynamics of the estimates for the GC skewness (d_3) and excess kurtosis (d_4) parameter used in the backtesting. Dotted lines divide the pre-crisis and crisis period.

skew parameter, which is relatively stable throughout the pre-crisis period but has an increasing pattern during the crisis.

Fig. 3 displays the shape parameters for the GPD implemented in the EVT model. Parameter ξ in Fig. 3A seems to be negative (evidencing shorter tails) in the pre-crisis period and positive (heavier tails) at the beginning of the crisis period and becomes negative again at the end, and parameter δ in Fig. 3B is very volatile and seems to increase in the crisis period.

Fig. 4A. and B illustrate the dynamics for GC–ML1 parameters, d_3 and d_4, which account for skewness and excess kurtosis, respectively. Parameter d_3 estimates remain negative for the whole pre-crisis period, although they are positive within a specific time range in the crisis period. Regarding parameter d_4 estimates, they are positive and decreasing in the pre-crisis period but increasing with extreme values occurrence.

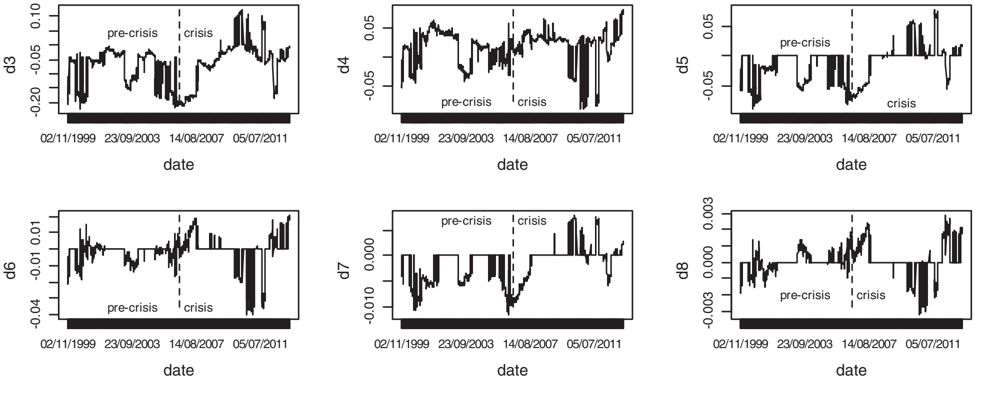

Fig. 5. GC–ML2 time-varying parameters for MSCI EM index returns. This figure shows the dynamics of the estimates for the GC parameters (d_3–d_8) used in the backtesting. Dotted lines divide the pre-crisis and crisis period.

Fig. 5 shows the evolution of the parameters of the GC expanded up to the eighth term (GC–ML2). Note that in the first part of the crisis period, only d_3 and d_4 are included according to the AIC criterion. In this context, the interpretation of these parameters is consistent with the expected values of skewness (negative) and excess kurtosis (positive). Nevertheless, when a larger expansion is chosen, skewness is captured by the interaction among the odd parameters and heavy-tailed patterns are featured by the interaction among the even parameters. Therefore, the insights about the combinations that may incorporate a certain degree of skewness and kurtosis are not straightforward. This is what happens, for instance, in the last part of the crisis period, where odd (even) parameters seem to be significantly positive (negative), although they also present an extreme volatility. Furthermore, the high presence of extreme values involves both asymmetries and heavy tails, and thus the relation between even and odd parameters also plays an important role in accounting for them.

Table 5 displays the number of exceptions at 99% confidence level and their binomial test p-value (in parentheses) for the seven models (ARMA–GARCH–normal, ARMA–GARCH–t, ARMA–GARCH–skewed-t, ARMA–GARCH–EVT, ARMA–GARCH–GC–MM, ARMA–GARCH–GC–ML1, ARMA–GARCH–GC–ML2) and for the seven indices, including a hedge fund. Table 5 shows two different panels, one for each backtesting period (pre-crisis and crisis periods).

For the pre-crisis period (Panel A), the ARMA–GARCH–normal distribution specification underpredicts the tail behavior (i.e. the number of exceptions is significantly higher than the expected value) for all cases, while the ARMA–GARCH–skewed-t and ARMA–GARCH–GC–ML2 fail for the MSCI USA case, and the t for the EM Europe case. The ARMA–GARCH–EVT, ARMA–GARCH–GC–MM and ARMA–GARCH–GC–ML1

Table 5
VaR forecasting performance of different models.

	EUEurope	USA	EM	EMLA	EM Europe	EM Asia	DJCSEM HF
Panel A: Pre-crisis period (November 1999–July 2006)							
99% 1750 days	Expected number of exceptions = 17						
ARMA–GARCH	26	27	38	39	38	31	
Normal	(0.0331)	(0.0203)	(0.0000)	(0.0000)	(0.0000)	(0.0021)	
ARMA–GARCH	20	13	25	25	26	23	
Student's t	(0.3048)	(0.1685)	(0.0522)	(0.0522)	(0.0332)	(0.1174)	
ARMA–GARCH	16	10	17	18	19	20	
skewed-t	(0.4197)	(0.0380)	(0.5157)	(0.4842)	(0.3908)	(0.3048)	
ARMA–GARCH	21	22	19	18	21	19	
EVT	(0.2296)	(0.1670)	(0.3908)	(0.4842)	(0.2296)	(0.3908)	
ARMA–GARCH	21	18	19	16	19	18	
GC–MM	(0.2296)	(0.4842)	(0.3908)	(0.4197)	(0.3908)	(0.4842)	
ARMA–GARCH	21	18	20	21	22	22	
GC–ML1	(0.2296)	(0.4842)	(0.3048)	(0.2296)	(0.1670)	(0.1670)	
ARMA–GARCH	22	26	20	21	23	23	
GC–ML2	(0.1670)	(0.0332)	(0.3048)	(0.2296)	(0.1174)	(0.1174)	
Panel B: Crisis period (July 2006–first quarter of 2013)							
99% 1750 days	Expected number of exceptions = 17						
ARMA–GARCH	41	52	31	30	38	43	39
Normal	(0.0000)	(0.0000)	(0.0021)	(0.0039)	(0.0000)	(0.0000)	(0.0000)
ARMA–GARCH	29	29	22	25	25	31	29
Student's t	(0.0070)	(0.0070)	(0.1670)	(0.0522)	(0.0522)	(0.0021)	(0.0070)
ARMA–GARCH	12	13	12	16	15	12	14
Skewed-t	(0.1104)	(0.1685)	(0.1104)	(0.4197)	(0.3265)	(0.1104)	(0.2414)
ARMA–GARCH	21	21	14	19	16	21	20
EVT	(0.2296)	(0.2296)	(0.2413)	(0.3908)	(0.4197)	(0.2296)	(0.3048)
ARMA–GARCH	20	12	16	17	15	21	17
GC–MM	(0.3048)	(0.1104)	(0.4197)	(0.5157)	(0.3265)	(0.2296)	(0.5157)
ARMA–GARCH	20	21	17	17	19	23	21
GC–ML1	(0.3048)	(0.2296)	(0.5157)	(0.5157)	(0.3908)	(0.1174)	(0.2296)
ARMA–GARCH	25	19	15	19	19	23	23
GC–ML2	(0.0522)	(0.3908)	(0.3265)	(0.3908)	(0.3908)	(0.1174)	(0.1174)

This table displays the number of exceptions with respect to the forecasted VaR with the different models, series and periods (pre-crisis and crisis). P-values for the binomial test are in parentheses. EVT considers a 10% threshold.

models cannot be rejected at 5% confidence for any of the series. In the crisis period (Panel B), the ARMA–GARCH–normal significantly underpredicts the VaR for all series, while the ARMA–GARCH–Student's t cannot be rejected on EM, EMLA and EM Europe data. The ARMA–GARCH–skewed-t performance, however, tends to overpredict risk (i.e. it involves overly conservative risk measures) although VaR forecasts cannot be rejected in any of the series. The ARMA–GARCH–EVT approach and ARMA–GARCH–GC models perform well for all cases since these methods focus on modeling the extreme values. Within the ARMA–GARCH–GC the shorter expansions seem to provide more accurate VaR forecasts.

Regarding the comparison of the results between both periods, in the pre-crisis scenario the ARMA–GARCH–skewed-t and the ARMA–GARCH–GC–ML2 fail in one occasion each (MSCI USA), while in the crisis period they always perform accurately. Therefore, the outperformance of the more flexible methods is clearer in the higher volatility scenario.

These results are consistent with the usual evidence found in stock return VaR performances, i.e. the normal-GARCH is strongly rejected (especially for high confidence levels and volatile scenarios), the Student's t–GARCH might be useful only if kurtosis and skewness are not severe (see e.g. Lin and Shen, 2006) and the skewed-t–GARCH is an alternative to capture skewness although it might not be the best option for capturing kurtosis. The EVT–GARCH and the GC–GARCH involve accurate market risk measures since the former focuses on extreme values and the latter is very flexible to adapt to different scenarios. Nevertheless we find that, for prediction purposes, the shorter GC expansions seem to provide the best outcomes and the simpler MM estimation procedures involve the most accurate VaR forecasts. These results are consistent with Del Brio and Perote (2012), although this paper focuses on in-sample fit.

The information for the crisis period in Table 5 is summarized in Table 6, highlighting the sign of the exceptions, i.e. the cases when VaR is under- (positive sign) and over- (negative sign) estimated. The table emphasizes the fact that the ARMA–GARCH–GC–ML2 seems to outperform their competitors, although the ARMA–GARCH–EVT and the ARMA–GARCH–Skewed-t cannot be rejected (the latter tending to overestimate risk at 99% confidence). The performances of all these models improve for the EM indices, likely because the EM remained more stable during the sovereign debt crisis. The ARMA–GARCH–normal model underpredicts risk, as well as the ARMA–GARCH–Student's t in the cases of the MSCI USA and MSCI Europe, precisely where the subprime and sovereign debt crises had their origins.

An illustration of the MSCI EM index returns and the 1-day forecasted VaR at 99% confidence for the normal–GARCH, Skewed t–GARCH, EVT–GARCH and GC–ML1–GARCH and the whole backtesting period is plotted in Fig. 6.

In Fig. 7 we provide an illustration of the 10-day returns and their corresponding VaR at 99% obtained for the DJCSEM HF by assuming normal and Student's t innovations and through alternative procedures: the ARMA–GARCH (the methodology employed in the rest of the paper applied to 10-days forecasting, see e.g. Tsay, 2010) and the so-called square-root-of-time (Daníelsson and Zigrand, 2006; Diebold et al., 1997) and α-root rules (de Vries, 2000). As is traditionally found, the square-root rule seems to overestimate VaR compared to the ARMA–GARCH procedure. Nevertheless, when Student's t innovations are assumed, the multi-period VaR forecasting obtained by the α-root rule is less conservative than the square-root approach since $h^{1/2} > h^{1/\alpha}$, where h is the holding period (10 days) and $\alpha = 1/v$ the tail index. Along the backtesting, Student's t degrees of freedom (v) ranges between 4.89829 and 10, then $\alpha \in [0.1,0.2]$.

Table 6
Exceptions over the forecasted of VaR models.

	Normal	Student t	Skewed-t	EVT	GC–MM	GC–ML1	GC–ML2
Crisis period (July 2006–first quarter of 2013)							
MSCI Europe	**+24**	**+12**	−5	+4	+3	+3	+7
MSCI USA	**+35**	**+12**	−4	+4	−5	+4	+2
MSCI EM	**+14**	+5	−5	−3	−1	0	−2
EMLA	**+13**	+8	−1	+2	0	0	+2
EM Europe	**+21**	+8	−2	−1	−2	+2	+2
EM Asia	**+26**	**+14**	−5	+4	+4	+6	+6

This table gathers the times when realizations exceed the forecasted VaR for the different distributional assumptions and indices. Positive/negative values indicate under/over-prediction of VaR. The statistically significant values at 5% confidence are highlighted in bold. All the models in the columns include an ARMA–GARCH structure.

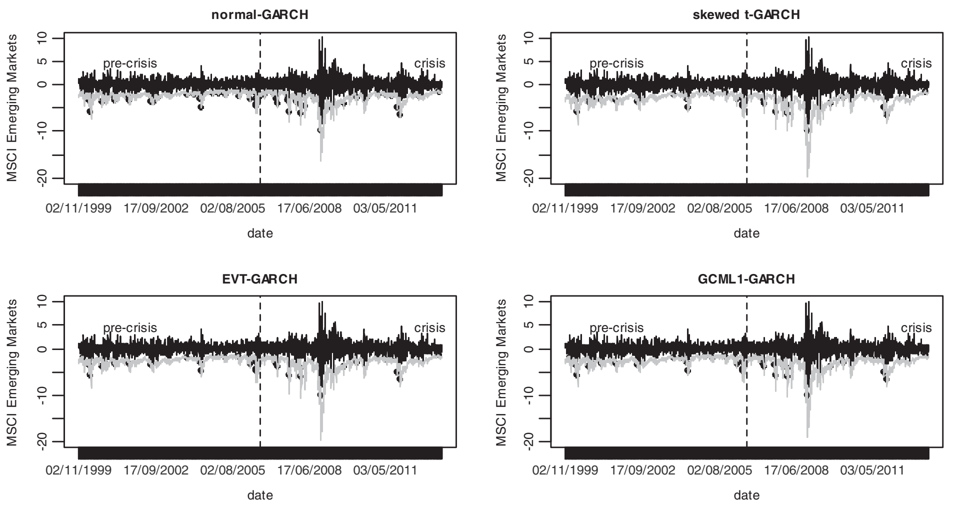

Fig. 6. VaR at 99% in the backtesting period under different specifications. This figure represents the returns of the MSCI EM index compared to its forecasted 1-day VaR at 99% confidence under different specifications.

Finally, Fig. 8 illustrates the relative performance of the GC–GARCH model against its more traditionally used competitors: the normal–GARCH and Student's t–GARCH, in terms of the magnitude of the ex-post VaR forecasted errors. These quantities are extremely important for true investment portfolios, as noted in Berkowitz and O'Brien (2002), and thus we make this comparison for the DJCSEM HF. For this purpose, we compute the excess loss at time t and for model G as

$$EL_t^G = x_t - VaR_t^G, \tag{15}$$

where x_t is the daily loss (negative return turned to a positive value, i.e. $x_t = -r_t$ and VaR_t^G is the forecasted 1-day, 99% VaR computed at time t by assuming disturbances distributed according to the pdf G. Then the ‘savings’ (S_t) incurred by fund managers using GC–GARCH (GC) compared with those using a normal-GARCH (n) are given by

$$S_t^{n-GC} = EL_t^n - EL_t^{GC}. \tag{16}$$

Note that these ‘savings’ are the excess losses not provisioned by the institution if they implement a normal–GARCH instead of a GC–GARCH model. These magnitudes are plotted in the left picture of Fig. 8, and the savings of the Student's t–GARCH and the GC–GARCH comparison are depicted in the picture on the right. In both cases, the average daily savings of using the GC–GARCH is 0.2681% (0.1614%) compared to the normal–GARCH (Student's t–GARCH). These figures represent an important improvement for the efficient allocation of regulatory capital in favor the GC–GARCH.

4. Conclusions

The recent financial crises have caused high volatility in financial markets and big losses for many investors. To quantify the potential losses and comply with regulatory capital requirements, financial institutions implement VaR methodologies. However, traditional VaR measures, based on the normal distribution, have been criticized because of their inability to adequately capture market risk, particularly at high confidence levels. For this purpose, the use of alternative thick-tailed and skewed distributions has been proposed, but there is still no consensus about the most appropriate methodology for VaR forecasting. In this paper, we investigate this issue by comparing the relative performance of alternative specifications for the innovations of

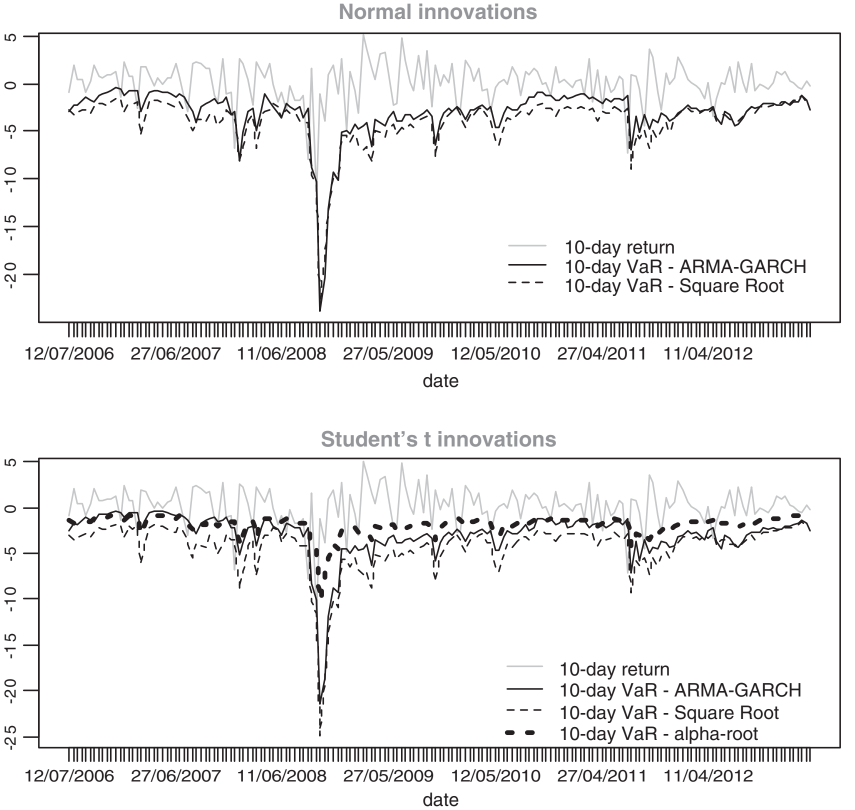

Fig. 7. Multiperiod VaR at 99% in the backtesting for DJCSEM HF. These figures represent the multi-returns (10-days) of the DJCSEM HF index and their forecasted 99% VaR under both normal and Student's t innovations and computed through the ARMA–GARCH model and the square-root and α-root (tail index) approximations. For the Student-t distribution $\alpha = 1/\nu$, where ν stands for the degree of freedom.

an ARMA–GARCH model for the returns of different stock indices (mainly of EM) and a hedge fund on EM. We consider three parametric models (normal, Student's t and skewed-t), the EVT–POT approach and a semi-nonparametric model (GC). We argue that the model ranking depends on the period under analysis, and thus we compare the sensitivity of VaR measures to the increase in volatility by studying VaR measures before and after the recent subprime and sovereign debt crisis. We compute 1-day 99% VaR but also provide an

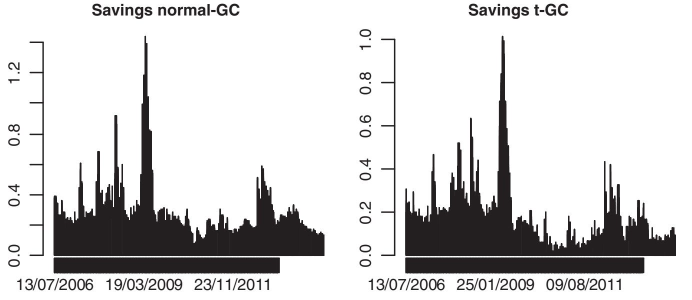

Fig. 8. Savings along the backtesting for the DJCSEM HF. This figure represents the differences (savings) between the VaR obtained assuming normal and GC innovations (left) and Student's t and GC disturbances (right).

example of the 10-day 99% and of the savings incurred by the hedge fund managers by assuming GC innovations instead of the widely used normal or Student's t.

At least eight main conclusions may be drawn from our study: (i) the normal underestimates risk even in low volatility scenarios; (ii) Student's t seems to be adequate for capturing VaR only for "relatively calm" periods and its skewed counterpart seems to be a better model for high volatility periods, although it tends to provide conservative risk measures; (iii) both EVT and GC are accurate methods for computing VaR at high confidence levels (99%); (iv) the larger GC expansions (GC–ML2) do not necessarily improve VaR measures compared to the simpler two-parameter (skewness and excess kurtosis) formulations; (v) the MM estimation method of GC densities seems to provide straightforward and accurate risk measures; (vi) the VaR of EM indices during the crises seems to be more predictable; (vii) multi-period VaR measures obtained through traditional squared-root rule are misspecified; and (viii) assuming GC innovations represents important savings in terms of the efficiency management of regulatory capital.

Result (i) is standard in the literature (e.g. see McNeil et al., 2005, p. 46-7 and 58, for a VaR performance comparison of the normal and Student's t at different confidence levels). Result (ii) emphasizes the poor VaR forecasting performance of Student's t in highly volatile scenarios and the better performance of skewed-t in this case (although likely at the expense of overly conservative VaR measures). This finding might be explained by the fact that positive skewed distributions capture the left tail of the empirical distribution in bear market periods but not in bull market periods, where the distribution exhibits negative skewness. Result (iii) is a consequence of the EVT (POT) approach, developed to capture extreme events, and the fact that this methodology is very sensitive to the threshold selection. This evidence is consistent with other EVT studies, e.g. Rachev et al. (2010). The accurate performance of the GC density lies in the asymptotic properties of the Hermite polynomial expansion and the flexibility of its formulation, which is capable of capturing not only leptokurtosis and skewness but also other features such as the not strictly decreasing pattern of the distribution tails. Conclusion (iv) seems to contradict our former assessment but it supports the well-known fact that a good in-sample fit does not guarantee a good out-of-sample performance (see e.g. Hansen, 2009) and thus simpler models usually provide better forecasting outcomes. Result (v) is in line with Del Brio and Perote (2012) and implies that accurate VaR forecasts according to the GC specification can be straightforwardly obtained by implementing MM techniques. Conclusion (vi) features more stability of EM during the sovereign debt crisis and result (vii) is also standard in the literature (Daníelsson and Zigrand, 2006). Finally, the last assessment (viii) highlights the savings in terms of regulatory capital provisions from using more accurate risk measures.

All these results highlight the fact that the optimal VaR model depends not only on the assumed confidence level (risk aversion) but also on the observed market conditions (volatility). Therefore risk forecasting methodologies should accommodate the scenario in which forecasts are computed. Only by combining different methods or by using very flexible techniques can the regulatory capital and the provisions of financial institutions be accurately estimated. For this reason and according to our findings, we recommend implementation of the GC density to forecast VaR.

Appendix A. Dataset description.

Index	Description
MSCI Europe	The MSCI Europe Index is a free float-adjusted market capitalization weighted index that is designed to measure the equity market performance of the developed markets in Europe. The MSCI Europe Index consists of the following 16 developed market country indices: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.
MSCI USA	The MSCI USA Index is a free float-adjusted market capitalization index that is designed to measure large and mid cap US equity market performance. The MSCI USA Index is member of the MSCI Global Equity Indices and represents the US equity portion of the global benchmark MSCI ACWI Index.

(continued on next page)

Appendix (*continued*)

Index	Description
MSCI Emerging Markets (EM)	The MSCI EM Markets Index is a free float-adjusted market capitalization index that is designed to measure the equity market performance of EM. The MSCI EM Index consists of the following 21 emerging market country indices: Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Korea, Malaysia, Mexico, Morocco, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, and Turkey.
MSCI EM Latin America	The MSCI EM Latin America Index is a free float-adjusted market capitalization weighted index that is designed to measure the equity market performance of EM in Latin America. The MSCI EM Latin America Index consists of the following 5 EM country indexes: Brazil, Chile, Colombia, Mexico, and Peru.
MSCI EM Europe	The MSCI EM Europe Index is a free float-adjusted market capitalization weighted index that is designed to measure the equity market performance of EM in Europe. The MSCI EM Europe Index consists of EM indexes of the following 6 countries: Czech Republic, Greece, Hungary, Poland, Russia, and Poland.
MSCI EM Asia	The MSCI EM Asia Index is a free float-adjusted market capitalization weighted index that is designed to measure the equity market performance of EM in Asia. The MSCI EM Asia Index consists of the following 8 EM country indexes: China, India, Indonesia, Korea, Malaysia, Philippines, Taiwan, and Thailand.
DJCSEM HF	The Dow Jones Credit Suisse EM Index is an asset-weighted hedge fund index derived from the TASS database of more than 5000 funds. The strategy involves equity or fixed income investing in EM around the world. Because many markets do not allow short selling, nor offer viable futures or other derivative products with which to hedge, EM investing often employs a long-only strategy. Credit Suisse EM Index is updated on the 15th or the next business day for the previous month.

Source: Datastream, Thomson Financial, Bloomberg and msci.com.

References

Bali, T.B., Theodossiou, P., 2008. Risk measurement performance of alternative distribution functions. J. Risk Insur. 75, 411–437.
Berkowitz, J., O'Brien, J., 2002. How accurate are value-at-risk models at commercial banks? J. Financ. 57, 1093–1111.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. J. Econ. 31, 307–327.
Christoffersen, P., 2003. Elements of Financial Risk Management. Academic Press, San Diego, CA.
Cont, R., 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quant. Finan. 1, 223–236.
Daníelsson, J., Zigrand, J.-P., 2006. On time-scaling of risk and the square-root-of-time rule. J. Bank. Finan. 30, 2701–2713.
Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds. J. R. Stat. Soc. Ser. B 52, 393–442.
de Vries, C.G., 2000. Second order diversification effects. In: Chan, W.S., Li, W.K., Tong, H. (Eds.), Statistics and Finance: An Interface. Imperial College Press, Singapore, pp. 153–165.
Del Brio, E.B., Perote, J., 2012. Gram–Charlier densities: maximum likelihood versus the method of moments. Insur. Math. Econ. 51, 531–537.
Diebold, F.X., Hickman, A., Inoue, A., Schuermann, T., 1997. Converting 1-day volatility into h-day volatility: Scaling by $\sqrt{h}$ is worse than you think. Working Paper. Wharton Financial Institutions Center.
Dowd, K., Blake, D., 2006. After VaR: the theory, estimation, and insurance applications of quantile-based risk measures. J. Risk Insur. 73, 193–229.
Edgeworth, F.Y., 1907. On the representation of statistical frequency by a series. J. R. Stat. Soc. Ser. A 70, 102–106.
Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom. Econometrica 50, 987–1007.
Fernández, C., Steel, M.F.J., 1998. On Bayesian modeling of fat tails and skewness. J. Am. Stat. Assoc. 93, 359–371.
Gallant, A.R., Nychka, D.W., 1987. Semi-nonparametric maximum likelihood estimation. Econometrica 55, 363–390.
Giot, P., Laurent, S., 2003. Value-at-risk for long and short trading positions. J. Appl. Econ. 18, 641–663.
Hansen, B.E., 1994. Autoregressive conditional density estimation. Int. Econ. Rev. 35, 705–730.
Hansen, P.R., 2009. In-sample fit and out-of-sample fit: their joint distribution and its implications for model selection. Working Paper Stanford University.
Harmantzis, F.C., Miao, L., Chien, Y., 2006. Empirical study of value-at-risk and expected shortfall models with heavy tails. J. Risk Finan. 7, 117–135.
Huang, A.Y.-H., Tseng, T.-W., 2009. Forecast of value at risk for equity indices: an analysis from developed and emerging markets. J. Risk Finan. 10, 393–409.
Jarrow, R., Rudd, A., 1982. Approximate option valuation for arbitrary stochastic processes. J. Financ. Econ. 10, 347–369.
Jondeau, E., Rockinger, M., 2001. Gram–Charlier densities. J. Econ. Dyn. Control. 25, 1457–1483.
Jorion, P., 2006. Value at Risk: The New Benchmark for Managing Financial Risk, 3rd ed. McGraw-Hill.
Kendall, M., Stuart, A., 1977. The Advanced Theory of Statistics. Griffin&Co, London.
Kolb, R.E., 2010. Lessons from the Financial Crisis: Causes, Consequences, and Our Economic Future. Robert E. Kolb Series, Wiley.

Kuester, K., Mittnik, S., Paolella, M.S., 2006. Value-at-risk prediction: a comparison of alternative strategies. J. Financ. Econom. 4, 53–89.
León, A., Mencía, J., Sentana, E., 2009. Parametric properties of semi-nonparametric distributions. J. Bus. Econ. Stat. 27, 176–192.
Lin, C.-H., Shen, S.-S., 2006. Can the Student-t distribution provide accurate value at risk? J. Risk Finan. 7, 292–300.
Mandelbrot, B., 1963. The variation of certain speculative prices. J. Bus. 36, 394–419.
Mauleón, I., 2003. Financial densities in emerging markets: an application of the multivariate ES density. Emerg. Mark. Rev. 4, 197–223.
Mauleón, I., Perote, J., 2000. Testing densities with financial data: an empirical comparison of the Edgeworth–Sargan density to the Student's t. Eur.J. Financ. 6, 225–239.
McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. J. Empir. Financ. 7, 271–300.
McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Pricenton.
Mishkin, F.S., 2011. Over the cliff: from the subprime to the global financial crisis. J. Econ. Perspect. 25, 49–70.
Rachev, Z., Racheva-Iotova, B., Stoyanov, S., 2010. Capturing fat tails. Risk Mag. 2010, 72–77 (May).
Sharpe, W., 1992. Asset allocation: management style and performance measurement. J. Portf. Manag. 18, 7–19.
Shiller, R.J., 2008. The Subprime Solution: How Today's Global Financial Crisis Happened, and What to Do about It. Princeton University Press.
Tsay, R., 2010. Analysis of Financial Time Series, 2nd ed. John Wiley & Sons.
J.-N.Wang, J.-N., J.-H.Ye, J.-H., N.Y.-P.Cheng, N.Y.-P., 2011. How accurate is the square-root-of-time rule in scaling tail risk: a global studyJ. Bank. Finan. 35, 1158–1169.
Zumbach, G., 2006. Backtesting risk methodologies from one day to one year. J. of Risk 9, 55–91.