J. Int. Financ. Markets Inst. Money 46 (2017) 30–53

Contents lists available at ScienceDirect

Journal of International Financial Markets, Institutions & Money

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

Value-at-Risk under Lévy GARCH models: Evidence from global stock markets

Skander Slim *, Yosra Koubaa, Ahmed BenSaïda

LaREMFiQ – IHEC, Sousse University, B.P. 40 Route de la ceinture, Sahloul III, Sousse 4054, Tunisia

ARTICLE INFO

Article history:
Received 20 October 2015
Accepted 30 August 2016
Available online 7 September 2016

JEL classification:
C22
G17

Keywords:
Value-at-Risk
Risk management
Lévy distributions
GARCH model
Asymmetry
Long memory

ABSTRACT

The aim of this paper is to reconsider the evidence on the forecasting ability of GARCH-type models in estimating the Value-at-Risk (VaR) of global stock market indices with improved return distribution. The performance of twenty-one VaR models that are generated by a combination of three conditional volatility specifications including GARCH, GJR and FIGARCH and seven distributional assumptions for return innovations is investigated. We implement stringent backtesting during crisis and post-crisis periods for developed, emerging and frontier markets. Results show that the skewed-*t* along with heavy-tailed Lévy distributions considerably improve the forecasts of one-day-ahead VaR for long and short trading positions during crisis period, regardless of the volatility model. However, we find no evidence that a given volatility specification outperforms the others across markets. The relevant models show evidence of long memory in developed markets and conditional asymmetry in frontier markets; whereas the standard GARCH is found to be the best suited specification for estimating VaR forecasts in emerging markets. The inclusion of high volatility period in the estimation sample highlights the predictability of VaR during post-crisis period, where even the normal distribution rivals the more sophisticated ones in terms of statistical accuracy and regulatory capital allocation.

1. Introduction

The Value-at-Risk (VaR) concept has been established as an industry standard measure of market risk. It provides financial institutions with information on the expected worst loss over a target horizon at a given confidence level. Despite its importance and simplicity, there is no universally accepted method to compute the VaR of a particular portfolio, while different models may lead to significantly different risk measures (see Kuester et al., 2006; McMillan and Kambouroudis, 2009, among others), a main concern in the estimation of market risk with the VaR method is the choice of the appropriate model, *e.g.*, a misspecified model may turn out to be costly for the risk manager, and leads to inaccurate risk estimation. Moreover, the extreme losses experienced by financial institutions during the recent global financial crisis, triggered by the U.S. subprime mortgage debacle of 2007–2008, have raised questions about the reliability of the implemented risk models. These questions have a direct bearing on the debate amongst financial industry, regulators and academicians over probabilistic market models for VaR forecasting, which are capable to properly account for extreme events and increased volatility during financial market downturns.

* Corresponding author.
E-mail addresses: slim_skander@yahoo.fr (S. Slim), yosra.koubaa@yahoo.fr (Y. Koubaa), ahmedbensaida@yahoo.com (A. BenSaïda).

http://dx.doi.org/10.1016/j.intfin.2016.08.008
1042-4431/

In obtaining accurate VaR measures, the prediction of future market volatility is of paramount importance, particularly in view of its time-varying *nature* as well as some prominent stylized facts of stock returns (Cont, 2001). Indeed, there is ample empirical evidence that spells of small amplitude for the price variations alternate with spells of large amplitude; one calls this feature volatility clustering. Numerous econometric models have been suggested to capture the volatility clustering effect, the most widely used one is the GARCH model (Bollerslev, 1986). The normal distribution arising from the Brownian motion assumption as a benchmark process for describing return innovations have dominated former GARCH-based VaR models drawing criticisms that such distributional assumption may not sufficiently capture the frequency of extreme shocks to asset prices, as well as the amplitude of these shocks, and usually leads to risk underestimation. The introduction of more flexible distributions, allowing for skewness and heavy tails into return modeling, exonerates GARCH-type models from such criticisms (Bao et al., 2007; BenSaïda, 2015).

Recently, enhanced conditional volatility models with Lévy distributions have emerged with a view towards improved tail modeling. Examples include the normal inverse Gaussian (Forsberg and Bollerslev, 2002; Venter and de Jongh, 2002; Broda and Paolella, 2009; Wilhelmsson, 2009), the multivariate generalized hyperbolic distribution (McNeil et al., 2005) and used in a GARCH context by Paolella and Polak (2015a,b), the Meixner distribution (Grigoletto and Provasi, 2008), α-stable and tempered stable distributions (Mittnik and Paolella, 2003; Broda et al., 2013; Kim et al., 2008, 2011). The current study pursues the use of Lévy distributions as statistical tools for advanced risk modeling.

More concretely, our objective is to examine the suitability of univariate GARCH-type models in modeling conditional volatility and VaR under different assumptions of error distribution for global stock market indices.[1] Within the class of conditional volatility models, we employ the standard GARCH model, GJR (Glosten et al., 1993) and FIGARCH (Baillie et al., 1996). Four infinitely divisible distributions arising from popular Lévy processes are considered, namely, the Variance Gamma (Madan et al., 1998), the CGMY (Carr et al., 2002), the normal inverse Gaussian (Barndorff-Nielsen, 1997), and the Meixner distribution (Schoutens, 2001).

The forecasting performance of these models is discussed and compared to the benchmark normal GARCH model, fat-tailed time series models using the Student's t and the skewed-t distribution of Hansen (1994). We also consider two higly competitive models, namely, the mixed normal GARCH and the fast Asymmetric Power ARCH model driven by noncentral t innovations, recently introduced by Krause and Paolella (2014). We empirically test the accuracy of VaR estimates during high volatility period for the MSCI World, MSCI emerging markets and MSCI frontier markets indices by means of an extensive backtesting exercise that includes the frequency-based test (Kupiec, 1995), independence tests (Christoffersen, 1998; Engle and Manganelli, 2004), duration-based tests (Candelon et al., 2011) and a test that jointly accounts for the frequency and the magnitude of VaR exceedances, recently introduced by Colletaz et al. (2013). Such backtesting strategy encompasses almost all risk model validation techniques in the literature, and provides a meaningful framework to evaluate the accuracy of VaR models. Furthermore, we evaluate the economic importance of our results by computing daily capital requirements under the Basel II Accord (Basel Committee on Banking Supervision, 2006).

This paper distinguishes itself from the relevant literature in several ways. (1) It allows for direct comparisons to be made on the performance of a broad set of GARCH-type specifications combined with non-normal Lévy distributions for gauging and managing market risk. Although, the option pricing literature establishes Lévy processes as quite suitable in capturing the behavior of returns innovation; other evaluation metrics and discussions on their applicability as risk management tools are needed. (2) Our empirical assessment of the performance of competing VaR models not only allows to test whether VaR forecasts are misspecified overall, but also identifies how they are misspecified by looking into which individual hypothesis is rejected. The hypotheses induced by the employed backtesting methods have different economic implications. Financial institutions focus on the frequency test since the empirical number of violations (*i.e.*, when the realized losses exceed the VaR values) is used to determine their market risk minimum capital requirements. Even in practice, market participants and regulators do care about the number of VaR violations, they often have a more stringent requirement on the magnitude of their losses beyond VaR, as such quantity is not given by any VaR model and represents the main conceptual deficiency of VaR. The duration independence test considers whether the VaR reflects the time-varying nature of risk. In particular, clustered violations may signal that the VaR model reacts too slowly to changes in market conditions, and hence may induce solvency issues.

Further, to the best of our knowledge, no published article examines all of these practically important issues in the backtesting of VaR estimates. (3) Until recently, a substantial literature of empirical applications narrowly focuses on the computation of the VaR on the left tail of the distribution which corresponds to the long position risk. The short position risk, which is theoretically unlimited, has been neglected and only a few papers have attempted to jointly account for both long and short position risk in equity markets (Giot and Laurent, 2003; So and Yu, 2006; Tang and Shieh, 2006; Diamandis et al., 2011). In our empirical analysis, the performance of VaR models is also examined in terms of the critical issue of model consistency between long and short position risk.

The main contributions of this study are the following. First, our results corroborate the calls for the use of more realistic assumptions in financial modeling. In model estimation, it is shown that allowing for leptokurtic and skewed return distributions significantly improves the fit of conditional volatility models. Second, even though, non-normal Lévy distributions do

[1] Although, multivariate GARCH models for VaR prediction are beyond the scope of the present paper, Santos et al. (2013) show that they outperform their univariate counterparts at least on an out-of-sample basis for large and diversified portfolios.

provide more accurate VaR estimates than those generated by both normal and Student's t distributions, conditional volatility specification bears on the accuracy of VaR as none of the three GARCH-type models outperform the others across markets. The empirical evidence favors non-normal Lévy-based FIGARCH models in developed markets, whereas the GJR and the standard GARCH models are preferred for the frontier and emerging markets, respectively. Third, this paper offers important implications regarding the recent global financial crisis with respect to the estimation of VaR for developed, emerging and frontier markets. In general, the forecasting performance of the VaR models deteriorates during crises periods, predominantly in the case of developed markets. For emerging and frontier markets, nevertheless, the forecasting performance is less affected. Fourth, the inclusion of the crisis period in the estimation sample, significantly improves the backtesting performance of the competing VaR models during the post-crisis period. Even though, normal time series models rival the more flexible non-normal Lévy models. However, our findings provide compelling evidence for the inadequacy of the Student's t distribution, as it tends to produce overly conservative risk measures during both high and low volatility periods.

The remainder of the paper is organized as follows. Section 2 outlines the basic concept of VaR and presents the employed time series models. Section 3 presents the alternative non-normal Lévy distributions. Section 4 describes the data and discusses the results of the empirical investigation. Section 5 summarizes the main findings of the paper.

2. VaR and volatility specifications

The VaR concept has emerged as the most prominent measure of downside market risk. It places an upper bound on losses at a given confidence level over a given forecast horizon. Thus, assuming that the VaR model is correct, realized losses will exceed the VaR threshold with only a small target probability α, typically chosen between 1% and 5%. More specifically, conditional on the information until time $t-h$, the VaR (for long position) on time t of one unit of investment is the α-quantile of the conditional return distribution, that is:

$$\mathrm{VaR}_t = q_\alpha(r_t|\mathcal{F}_{t-h}) = \inf\{x \in \mathbb{R}|P(r_t \leqslant x|\mathcal{F}_{t-h}) \geqslant \alpha\}, \tag{1}$$

where q_α denotes the quantile function, r_t is the index return in period t, and $\mathcal{F}_{t-1}$ designates the information available at date $t-1$. When the expected returns, r_t, are assumed to follow a location-scale distribution, they are regarded as a function of an innovation process, ϵ_t, conditional on a given $\mathcal{F}_{t-1}$-measurable parameter a_t. Specifically:

$$\begin{cases} r_t = a_t + \epsilon_t, \\ \epsilon_t = \sigma_t z_t. \end{cases} \tag{2}$$

where a_t and σ_t are the location and scale components, respectively. z_t is an independent and identically distributed (*i.i.d.*) random variable that follows a zero-mean, unit-variance probability density function f_z. Without loss of generality, we assume a constant conditional location $a_t \equiv a_0$.

In our study, we consider the modeling and calculation of VaR for portfolio managers who have taken either a long position or a short position. In the former case the risk comes from a drop in the price of the asset, while in the later case the trader will incur a loss when the asset price increases. To fully take into account risk exposure arising from both long and short trading positions, one needs an adequate modeling of the left and the right tails of the returns distribution. The one-day-ahead VaR forecast for long and short trading positions (*i.e.*, $\mathrm{VaR}_{L,t}$ and $\mathrm{VaR}_{S,t}$) at time t can be written, respectively, as follows:

$$\begin{cases} \mathrm{VaR}_{L,t} = \hat{a}_0 + q_\alpha(z)\hat{\sigma}_t, \\ \mathrm{VaR}_{S,t} = \hat{a}_0 + q_{1-\alpha}(z)\hat{\sigma}_t. \end{cases} \tag{3}$$

where $q_\alpha(z)$ and $q_{1-\alpha}(z)$ denote the left and right quantile implied by f_z at the level of significance α and $(1-\alpha)$, respectively. $\hat{a}_0$ and $\hat{\sigma}_t$ refer to the estimated/forecasted conditional location and scale of the index returns, respectively. While the location parameter is of little relevance for short holding periods, the key issue in VaR modeling is the specification of the conditional scale. In light of the stylized facts of financial returns such as volatility clustering, asymmetry and long memory (Cont, 2001), the GARCH-type models have become largely popular for describing the time-varying conditional volatility of financial returns. In the following, we present the standard GARCH model, the fractionally integrated GARCH (FIGARCH) model and the GJR model.

2.1. GARCH model

Developed by Bollerslev (1986), the conditional variance in the GARCH(1,1) specification is represented by:

$$\sigma_t^2 = \omega + \beta_1\epsilon_{t-1}^2 + \beta_2\sigma_{t-1}^2. \tag{4}$$

The conditional variance process is positive and stationary if the following conditions hold: $\omega > 0,\ \beta_1 > 0,\ \beta_2 > 0$ and $\beta_1 + \beta_2 < 1$. The GARCH(1,1) model can only handle short memory in the volatility process since its autocorrelation function decays rapidly with an exponential rate of $\beta_1 + \beta_2$.

2.2. FIGARCH model

Developed by Baillie et al. (1996), the FIGARCH model allows for fractional order of integration to parsimoniously capture long memory in volatility as frequently seen in practice, as well as to distinguish between the long memory and short memory in the conditional variance. Formally, the FIGARCH(p, d, q) implies the following infinite ARCH representation:

$$\sigma_t^2 = \frac{\omega}{\beta(1)} + \left[1 - \frac{\phi(L)}{\beta(L)}(1-L)^d\right]\epsilon_t^2 = \frac{\omega}{\beta(1)} + \lambda(L)\epsilon_t^2, \tag{5}$$

where $\phi(L) = 1 - \sum_{i=1}^{q}\phi_i L^i$ and $\beta(L) = 1 - \sum_{i=1}^{p}\beta_i L^i$ are lag polynomials with roots assumed to lie outside the unit circle, L is the lag operator and $\lambda(L) = \lambda_1 L + \lambda_2 L + \ldots$.[2] To guarantee the non-negativity of the conditional variance as surely for all t, all the coefficients in the infinite ARCH representation need to be non-negative.[3] For the FIGARCH $(1,d,1)$ model, conditional volatility is non-negative provided that $\omega > 0, \lambda_1 = d + \phi - \beta \geqslant 0$, and $\phi \leqslant (1-d)/2$ for some $0 < \beta < 1$. The long memory parameter, $0 < d < 1$, indicates that shocks to the conditional variance dissipate slowly following a hyperbolic rate of decay. The FIGARCH model nests the GARCH(1,1), when $d = 0$, as well as the IGARCH(1,1), when $d = 1$, whose variance process is no longer mean-reverting.

2.3. GJR model

In financial markets, it is often the case that downward movements in the market are followed by higher volatility than upward movements of the same magnitude (Engle and Ng, 1993). This asymmetry can be modeled using the GJR model of Glosten et al. (1993), where the impact of ϵ_{t-1}^2 depends on the sign of the shock, that is:

$$\sigma_t^2 = \omega + \beta_1\epsilon_{t-1}^2 + \beta_2\sigma_{t-1}^2 + \beta_3\epsilon_{t-1}^2 I_{t-1}, \tag{6}$$

where I_{t-1} is equal to unity if $\epsilon_{t-1} < 0$ and zero otherwise. The conditional volatility is positive when parameters satisfy $\omega > 0,\ \beta_1 + \beta_3 > 0$, and $\beta_2 > 0$. The process is covariance stationary if $\beta_1 + \beta_2 + \frac{1}{2}\beta_3 < 1$. The impact of shocks on conditional variance is asymmetric if β_3 is significantly different from zero.[4]

As we focus on VaR performance due to the distributional hypotheses, a number of distributional assumptions for the standardized innovation distribution f_z are combined with the above-mentioned volatility specifications. In the empirical analysis below, distributional assumptions are the normal, the Student's t, the skewed-t of Hansen (1994)[5] and four non-normal Lévy distributions.

2.4. Alternative models

In this section we briefly present two alternative models that have proven to provide outstanding prediction performance. The first one is the mixed normal (MN) GARCH model which can also accommodate long range dependence in volatility (see Haas and Paolella, 2012, and references therein). We focus on the *diagonal* MN-GARCH specification which is found to be favored in empirical applications and allows a clear-cut interpretation of the component-specific volatility processes (see Haas et al., 2004a,b; Alexander and Lazar, 2006; Haas et al., 2013, for discussion). The second one is the Asymmetric Power ARCH (APARCH) model (Ding et al., 1993) with noncentral t (NCT) innovations.[6]

2.4.1. Normal mixture GARCH

The *diagonal* mixed normal MN(K)-GARCH(1,1) model has K conditional variance components representing different market conditions. The error term has a conditional normal mixture density with zero mean as a weighted average of K normal density functions with different means and variances. That is:

$$\begin{aligned}
&\epsilon_t|\mathcal{F}_{t-1} \sim MN(p_1,\ldots,p_K;\mu_1,\ldots,\mu_K;\sigma_{1t}^2,\ldots,\sigma_{Kt}^2), \sum_{i=1}^{K}p_i = 1, \sum_{i=1}^{K}p_i\mu_i = 0,\\
&\sigma_{it}^2 = \omega_i + \beta_{1i}\epsilon_{t-1}^2 + \beta_{2i}\sigma_{it-1}^2 \quad \text{for} \quad i = 1,\ldots,K.
\end{aligned} \tag{7}$$

[2] Note that in practice, the infinite number of lags are truncated at 1000, which is large enough to examine the long memory process. The infinite ARCH terms are given by $\lambda_1 = d + \phi - \beta, \lambda_j = \beta\lambda_{j-1} + (f_j - \phi)$ for all $j \geqslant 2$ with $f_j = (j-1-d)/j$ and $g_j = f_j g_{j-1} = \prod_{i=1}^{j} f_j$ for $j = 1, 2, \ldots$ (Baillie et al., 1996; Conrad and Haag, 2006).

[3] Various different sets of sufficient parameter constraints for the conditional variance to be strictly positive are discussed in the literature (see Baillie et al., 1996; Bollerslev and Mikkelsen, 1996; Conrad and Haag, 2006, for more details).

[4] Negative shocks have more impact on the volatility than positive shocks of equal magnitude when $\beta_3 > 0$.

[5] Alternative (but very similar) asymmetric Student's t distributions have been proposed in several comparative VaR prediction studies including Giot and Laurent (2003, 2004) and Diamandis et al. (2011) for the skewed-t as introduced in Lambert and Laurent (2001), Mittnik and Paolella (2000) and Kuester et al. (2006) for applications of the three-parameter generalized asymmetric t (or equivalently, t_3 distribution) proposed in Paolella (1997). They have shown promise in VaR applications. For the sake of comparison, we have also considered the t_3 distribution in combination with GARCH-type models. Qualitatively similar results hold for the Hansen (1994)'s skewed-t distribution which is preferred on the grounds of parsimony.

[6] The authors are very grateful to an anonymous reviewer for suggesting to include these models into the "horse race".

where $0 < p_i < 1$ are the mixing weights. Models characterized by non-zero mean component densities in the mixture are called asymmetric MN(K)-GARCH models and these are particularly relevant when the returns densities are expected to be skewed and heavy-tailed. Additional details for the mixed GARCH-type models and the extensions thereof can be found in Haas and Paolella (2012) and Haas et al. (2013). In the empirical application, we use a two-component MN(2)-GARCH(1,1) model (henceforth MN-GARCH) which is found to be preferred over its higher dimensional counterparts according to standard model selection criteria.

2.4.2. NCT-APARCH

Krause and Paolella (2014) propose an APARCH process driven by NCT innovations. The location-centered returns are given by

$$r_t - a_0 = (z_t - \mathbb{E}(z_t))\sigma_t, \quad z_t \sim NCT(\nu, \gamma), \tag{8}$$

and the scale parameter evolves following the APARCH model:

$$\sigma_t^\delta = \omega + \beta_1(|\epsilon_{t-1}| - \kappa\epsilon_{t-1})^\delta + \beta_2\sigma_{t-1}^\delta, \tag{9}$$

where $\nu \in \mathbb{R}^+$ denotes the degree of freedom and $\gamma \in \mathbb{R}$ is the noncentrality parameter, which controls the asymmetry of the distribution.

To speed up the calculation of the NCT-APARCH model, the authors employ a fixed-parameter APARCH model by choosing a typical set of values associated with daily financial returns data. That is $\delta = 2, \ \omega = 0.04, \ \beta_1 = 0.05, \ \kappa = 0.4$ and $\beta_2 = 0.90$. Then the two NCT parameters are obtained from a pre-computed table, using a function of the sample quantiles. In addition to being much faster than the maximum likelihood estimation method (MLE), Krause and Paolella (2014) show that this method actually outperforms highly competitive models.[7]

3. Modeling return innovations with Lévy distributions

A Lévy process is a continuous time stochastic process with stationary independent increments, analogous to *i.i.d.* innovations in a discrete-time setting. The simplest continuous-time stochastic process is the Brownian motion which generates normally distributed innovations. The introduction of Lévy processes into financial modeling provides a flexible framework to replace the normal distribution by more sophisticated infinitely divisible ones. More specifically, while the Brownian motion component in a Lévy process generates a normal distribution, non-normal distributions can be generated via the appropriate specification of the Lévy density for a Lévy jump process, which determines the arrival rate of jumps of all possible sizes. A Lévy process $X = \{X_t, t \geqslant 0\}$ is closely related to an infinitely divisible distribution F. The Lévy-Khintchine formula gives the general representation for the characteristic function of F as follows:

$$\phi_F(u) = \exp\left(iu\mu - \frac{1}{2}u^2\sigma^2 + \int_{\Re_0}(e^{iux} - 1 - iux1_{|x|<1})k(x)dx\right), \tag{10}$$

where $\mu \in \mathbb{R}$ is the drift rate, $\sigma^2 \geqslant 0$ denotes the diffusion component and the Lévy density $k(x)$ describes the arrival rate for jumps of size x. The Lévy triplet $(\mu, \sigma^2, k(x))$ fully characterizes the distribution of the underlying process. A number of Lévy models proposed in the financial literature simply correspond to different choices for the Lévy density which is responsible for the richness of the class of Lévy processes. The classical example of such a process is the compound Poisson jump process with normal jump sizes underlying the jump-diffusion model of Merton (1976).

The rationale usually given for describing asset returns by jump-diffusion models is that the diffusion component captures frequent small moves, while Poisson-like jumps capture rare large moves. Given the ability of infinite-activity jump processes to capture both frequent small moves and rare large moves (Carr et al., 2002), we consider in this paper distributions that arise from infinite-activity pure jump processes.[8] Thus, the diffusion component, σ^2, is set to zero. In particular, we consider the following infinitely divisible distribution: normal inverse Gaussian, Meixner, Variance Gamma and CGMY. All of these distributions exhibit semi-heavy tails, a property that makes them appealing candidates for modeling index return innovations (see Schoutens, 2003; Cont and Tankov, 2004, among others).

3.1. Normal inverse Gaussian distribution

Barndorff-Nielsen (1997), Jensen and Lunde (2001), and Wilhelmsson (2009) have suggested the normal inverse Gaussian (NIG) distribution for the errors in GARCH and stochastic volatility models. The NIG distribution was first used for modeling speculative returns in Barndorff-Nielsen (1995, 1997). The density function of a NIG distributed random variable z is given by:

[7] The code for the fast NCT-APARCH method was kindly provided by Krause and Paolella (2014).

[8] When the Lévy density $k(x)$ integrates to infinity in the neighborhood of zero, the process is of infinite-activity, and generates an infinite number of jumps within any finite interval.

$$f_{NIG}(z;\mu,\alpha,\beta,\delta)=\frac{\alpha\delta K_1\left(\alpha\sqrt{\delta^2+(z-\mu)^2}\right)}{\pi\sqrt{\delta^2+(z-\mu)^2}}e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(z-\mu)}, \tag{11}$$

with $\mu\in\mathbb{R},\ 0\leqslant|\beta|<\alpha$ and $\delta>0$. $K_1(.)$ denotes the modified Bessel function of the third kind and index one. α controls the kurtosis of the distribution and β the asymmetry. The parameters μ and δ denote the location and scale of the distribution, respectively. The attractive features of the NIG distribution include the ability to fit leptokurtic and skewed data combined with nice analytical properties. One appealing property of the NIG distribution is that it is closed under convolution for fixed values of α and β. Therefore, if, say the daily returns are NIG distributed, then the h-day returns will also be NIG distributed.

3.2. Meixner distribution

The Meixner (MXN) distribution was introduced by Schoutens (2001, 2003) in the context of stochastic volatility models, and is a subclass of the generalized z-distributions (Grigelionis, 2001). A random variable z is said to follow a MXN distribution having parameters $\mu,\ \alpha,\ \beta,\ \delta$, with $\alpha>0,\ -\pi<\beta<\pi,\ \delta>0$ and $\mu\in\mathbb{R}$, if its density function is:

$$f_{MXN}(z;\mu,\alpha,\beta,\delta)=\frac{\left[2\cos\left(\frac{\beta}{2}\right)\right]^{2\delta}}{2\alpha\pi\Gamma(2\delta)}\left|\Gamma\left(\delta+\frac{i(z-\mu)}{\alpha}\right)\right|^2 e^{\frac{\beta(z-\mu)}{\alpha}}, \tag{12}$$

where $i=\sqrt{-1}$ and $\Gamma(.)$ is the gamma function. The location and scale of the distribution are decided by μ and α, while β and δ control the shape of the distribution.

3.3. Variance Gamma distribution

The symmetric Variance Gamma (VG) distribution was introduced to the financial literature by Madan and Seneta (1990) and further extended to incorporate skewness by Madan et al. (1998). A closed form for the density function having parameters $\mu,\ \sigma,\ v,\ \theta$ with $\sigma,\ v>0$ and $\theta,\mu\in\mathbb{R}$, is given by:

$$f_{VG}(z;\mu,\sigma,v,\theta)=\frac{2}{v^{1/v}\sqrt{2\pi}\sigma\Gamma\left(\frac{1}{v}\right)}\left[\frac{(z-\mu)^2}{\theta^2+2\frac{\sigma^2}{v}}\right]^{\frac{1}{2v}-\frac{1}{4}} e^{\theta\frac{z-\mu}{\sigma^2}}\times K_{\frac{1}{v}-\frac{1}{2}}\left(\frac{|z-\mu|}{\sigma^2}\sqrt{\theta^2+2\frac{\sigma^2}{v}}\right), \tag{13}$$

Table 1
Parameter restrictions for standardized distributions.

$\mathrm{NIG}(z;\alpha,\beta)$	$\mu=-\delta\beta/\sqrt{\alpha^2-\beta^2}$
	$\delta=\left(\alpha^2-\beta^2\right)^{3/2}/\alpha^2$
	$\mathrm{Skew}=3\beta/\left(\alpha^2-\beta^2\right)$
	$\mathrm{Kurt}=3\frac{\alpha^2+4\beta^2}{\left(\alpha^2-\beta^2\right)^2}$
$\mathrm{MXN}(z;\beta,\delta)$	$\mu=-\alpha\delta\tan(\beta/2)$
	$\alpha=\sqrt{\frac{\cos(\beta)+1}{\delta}}$
	$\mathrm{Skew}=\sin(\beta/2)\sqrt{2/\delta}$
	$\mathrm{Kurt}=\frac{(2-\cos(\beta))}{\delta}$
$\mathrm{VG}(z;v,\theta)$	$\mu=-\theta$
	$\sigma=\sqrt{1-v\theta^2}$
	$\mathrm{Skew}=\theta v\left(3-v\theta^2\right)$
	$\mathrm{Kurt}=3v\left(1+2v\theta^2-v^2\theta^4\right)$
$\mathrm{CGMY}(z;G,M,Y)$	$\mu=C\Gamma(1-Y)\left(M^{Y-1}-G^{Y-1}\right)$
	$C=\left(\Gamma(2-Y)\left(M^{Y-2}-G^{Y-2}\right)\right)^{-1}$
	$\mathrm{Skew}=\frac{(2-Y)(M^{Y-3}-G^{Y-3})}{(M^{Y-2}+G^{Y-2})}$
	$\mathrm{Kurt}=\frac{(2-Y)(3-Y)(M^{Y-4}+G^{Y-4})}{(M^{Y-2}+G^{Y-2})}$

This table presents the restrictions on the parameters to ensure that the Lévy distributions exhibit zero-mean and unit-variance (see BenSaïda and Slim, 2016). The table also gives a measure of the corresponding skewness and excess kurtosis (actual kurtosis minus 3) under these restrictions.

where $K_{\frac{1}{\nu}-\frac{1}{2}}$ is the modified Bessel function of the second kind with real index $\frac{1}{\nu}-\frac{1}{2}$. The VG distribution is symmetric around the location parameter μ if $\theta = 0$ whereas negative values of θ result in negative skewness. Also, the parameter ν primarily controls the kurtosis.

3.4. CGMY distribution

The CGMY distribution belongs to the class of tempered stable distributions and has been introduced to the finance literature by Carr et al. (2002). This distribution is also known as truncated Lévy flight (Koponen, 1995), the KoBoL distribution (Boyarchenko and Levendorskii, 2000) and dampened power law (Wu, 2006). The CGMY distribution does not have a closed form expression for the density function. Instead, the distribution is defined by the following characteristic function:

$$\phi_{CGMY}(u;\mu,C,G,M,Y) = \exp\left\{iu\mu + C\,\Gamma(-Y)\times\left((M-iu)^Y - M^Y + (G+iu)^Y - G^Y\right)\right\}, \tag{14}$$

with $C > 0,\ G \geqslant 0,\ M \geqslant 0,\ Y < 2$ and $\mu \in \mathbb{R}$. The scale parameter C characterizes the overall activity of the process while the parameters G and M control the rate of exponential decay on the left and right of the Lévy density, respectively. The difference in the two parameters generates asymmetry in the tails of the distribution. With small values of G and M there is a high probability of large jumps which yields high kurtosis. The Lévy process induced from the CGMY distribution generalizes the VG process by introducing the "fine" structure parameter Y which allows to recover a wider range of jump behaviors. For negative values of Y, the process is of finite activity, as a compound Poisson process, whereas positive values yield infinite-activity processes, like the NIG, the MXN and the VG processes.

To be suitable for GARCH modeling, the mean and variance of the unconditional distribution must equal 0 and 1, respectively. Table 1 gives the parameter restrictions in order to obtain standardized Lévy distributions.

4. Empirical results

4.1. Data and preliminary analysis

The data for this study consists of three global stock market indices, including MSCI Emerging Markets (EM), MSCI Frontier Markets (FM) and MSCI World designed to measure the equity market performance of developed markets. All the data are obtained from Thomson Reuters Eikon.[9] For the MSCI World and MSCI EM indices, the sample comprises 20 years of daily data from January 3, 1995 to March 10, 2015, and covers the period from May 31, 2002 to March 10, 2015, for the MSCI FM index (approximately 13 years). The continuously compounded daily returns are calculated as the logarithmic difference of daily closing prices multiplied by 100. To investigate the ability of the VaR models into measuring the risk with sufficient accuracy in different volatility scenarios, we choose two forecast sub-periods. The first forecast period starts on August 1, 2007, as suggested by Covitz et al. (2013) to include the sub-prime financial crisis. The second forecast period is referred to as the post-crisis period from May 15, 2012 to March 10, 2015.

Table 2 summarizes the main statistics of return series for the whole sample period and the two forecast periods. Descriptive statistics show that the mean return for all indices is positive during the whole sample period as well as in the post-crisis period, but 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 to the post-crisis period, but also that EM are more volatile in both forecast periods. As indicated by the coefficients of skewness and kurtosis, each of the return series presents left-skewed and leptokurtic distribution. The FM index, however, exhibits positive skewness in the post crisis periods. The Jarque-Bera (J-B) normality test strongly rejects the hypothesis of normality for all of the three periods. Moreover, the Ljung-Box $Q^2(20)$ statistics for the squared returns are statistically significant, providing evidence of strong ARCH effects. These salient features of daily returns for all groups of stock markets justify the use of VaR based on skewed and leptokurtic distributions.

4.2. Estimates of GARCH-type models

The empirical analysis focuses on the performance of VaR models for country stock portfolios that are generated by a combination of three conditional volatility structures (GARCH, FICARCH and GJR) and seven distributions for return innovations. Their forecasting ability is also compared to that gleaned from two highly competitive models (*i.e.*, MN-GARCH and NCT-APARCH). Parameters of the volatility models with normal, Student's t, skewed-t, NIG, MXN and VG innovations are estimated straightforwardly using the MLE. The MN-GARCH model is estimated via the extended augmented likelihood estimator (Broda et al., 2013). For the case of CGMY innovations, fitting the parameters is far away from being trivial. The implementation of the MLE raises the issue of numerical stability since the density function is not known in closed form and has to be calculated numerically using the fast Fourier Transform (FFT). Therefore, the parameters are estimated using the following two-step procedure. First, we estimate the parameters of the volatility models with Student's t distributed innovations by MLE and extract the residuals using the estimated parameters. Second, we fit the standardized CGMY distribution to the

[9] The list of country indices constituents as of June 2014, is provided in Appendix A.

Table 2
Descriptive statistics of daily returns.

	World index	EM index	FM index
Panel A. Estimation period (whole sample)			
Mean	0.019	0.013	0.025
S.D.	0.983	1.201	0.811
Skewness	−0.353*	−0.531*	−1.480*
Excess kurtosis	10.808*	10.776*	18.613*
Max	9.097	10.073	4.633
Min	−7.325	−9.994	−9.143
J-B	13482.85*	13481.20*	35058.51*
$Q^2(20)$	9128.51*	7504.03*	1536.59*
Panel B. Forecast period (crisis period, August 1, 2007–May 14, 2012), 1250 observations			
Mean	−0.020	−0.015	−0.057
S.D.	1.440	1.697	1.005
Skewness	−0.315*	−0.366*	−1.895*
Excess kurtosis	8.179*	8.735*	16.845*
Max	9.097	10.073	4.437
Min	−7.325	−9.994	−9.143
J-B	1417.49*	1741.07*	10731.74*
$Q^2(20)$	2048.94*	1959.77*	620.31*
Panel C. Forecast period (post-crisis period, May 15, 2012–March 10, 2015), 730 observations			
Mean	0.052	0.008	0.035
S.D.	0.653	0.794	0.522
Skewness	−0.222**	−0.006	0.271*
Excess kurtosis	5.435*	4.913*	12.142*
Max	2.938	3.324	4.278
Min	−3.520	−4.127	−2.432
J-B	186.11*	111.29*	2550.94*
$Q^2(20)$	44.23*	115.40*	174.12*

* Significance at 5% confidence level.
** Significance at 1% confidence level.

extracted residuals using MLE.[10] For the assessment of the goodness-of-fit, we use the Kolmogorov-Smirnov (KS) test and also calculate the Anderson-Darling (AD) statistic, which provides better results to evaluate the tail fit.[11]

The parameter estimates are reported in Tables 3–5, together with the results of the goodness-of-fit tests applied to standardized residuals. The values of the GARCH parameters are statistically significant and $\beta_1 + \beta_2$ is close to one for all series, reflecting the persistence and clustering in volatility. The results for the GJR models show that the negative return shocks have significant and greater impact on conditional volatility than the positive return shocks of equal magnitude in view of the significance of the positive sign of the β_3 coefficient. When the long memory property of volatility is effectively accounted for within the FIGARCH framework, the value of the GARCH coefficient decreases substantially. The long memory parameter (d), ranging from 0.3808 to 0.6544, is significant regardless of what type of distribution is used. More interestingly, the estimates of d also provide useful information regarding the measurement of length of the long memory in conditional volatility (Davidson, 2004). Except for the case of the normal distribution, the long memory parameter estimates for the FM index are markedly lower than the corresponding values for both World and EM indices. This result suggests that shocks in FM tend to last longer than in the remaining markets. A possible explanation for this may be that smaller markets, which are characterized by less liquidity, are less efficient according to the efficient market hypothesis, thus exhibiting higher long-lasting persistence.

The fitted MN-GARCH model is in agreement with the interpretation of the two-component mixture as distinct "regimes" representing bull and bear market conditions. For instance, the estimates for the World index imply a variance level (as measured by the unconditional component variance) equal to 0.4863 in the first variance process, and 1.4597 in the second. Hence, the first component features a high mean (0.1018) with low volatility (bull market), and the second a low mean (−0.1512) with high variance (bear market). The model is useful not only to capture different levels of variance but also unconditional skewness and kurtosis.

The shape and skew parameters of the residuals distribution are significant for most of the considered models. The skew parameter is negative in almost all cases which is consistent with the fact that an asymmetric distribution for the error term (unconditional asymmetry) exists. Besides, based on the goodness-of-fit tests, we conclude that for all of the three market indices, the normal distribution is overwhelmingly rejected at 5% significance level. The MN distribution, however, appears as an appealing candidate for modeling return innovations. Regarding the KS test, the Lévy distributions are not rejected at

[10] This is a general method that can be used in a FFT-based approximation of the probability density and has proven its efficiency (Bianchi et al., 2010; Scherer et al., 2012).

[11] As tabulated values of the AD critical values are only available for a few specific distributions (see Stephens, 1976), we use simulation techniques to approximate the distribution of the test statistic under the null hypothesis, based on 10,000 simulated samples and obtain the associated p-values of the test (for further details, see Alexander, 2008, p. 129).

Table 3
Estimation results of GARCH-type models for MSCI World index.

Model	Innovation					a_0	ω	β_1	β_2		KS	AD
GARCH	Normal					0.0495	0.0075	0.0799	0.9124		0.0330	13.3922
						(0.0000)	(0.0006)	(0.0000)	(0.0000)		(0.0000)	(0.0005)
	Student's t	ν	7.8430			0.0624	0.0067	0.0817	0.9129		0.0180	6.1127
			(0.0000)			(0.0000)	(0.0006)	(0.0000)	(0.0000)		(0.0642)	(0.0025)
	Skewed-t	λ	−0.1190	ν	8.3956	0.0468	0.0064	0.0796	0.9144		0.0192	3.2094
			(0.0000)		(0.0000)	(0.0000)	(0.0003)	(0.0000)	(0.0000)		(0.0403)	(0.0250)
	NIG	α	1.7101	β	−0.3274	0.0464	0.0065	0.0799	0.9137		0.0176	3.0894
			(0.0000)		(0.0000)	(0.0000)	(0.0002)	(0.0000)	(0.0000)		(0.0764)	(0.0213)
	MXN	β	−0.4558	δ	0.9390	0.0462	0.0066	0.0798	0.9135		0.0168	2.8865
			(0.0000)		(0.0000)	(0.0000)	(0.0006)	(0.0000)	(0.0000)		(0.0995)	(0.0275)
	VG	ν	0.3472	θ	−0.3013	0.0461	0.0067	0.0799	0.9132		0.0170	2.7225
			(0.0000)		(0.0000)	(0.0000)	(0.0196)	(0.0000)	(0.0000)		(0.0938)	(0.0463)
	CGMY	G	1.7118	M	2.2846	Y	0.3858				0.0194	4.5413
			(0.0000)		(0.0000)		(0.0000)				(0.0372)	(0.0213)
Model	Innovation					a_0	ω	β_1	β_2	β_3	KS	AD
GJR	Normal					0.0292	0.0094	0.0043	0.9259	0.1097	0.0330	11.4013
						(0.0034)	(0.0000)	(0.4821)	(0.0000)	(0.0000)	(0.0000)	(0.0005)
	Student's t	ν	8.5742			0.0472	0.0088	−0.0002	0.9235	0.1237	0.0145	3.6247
			(0.0000)			(0.0000)	(0.0001)	(0.9745)	(0.0000)	(0.0000)	(0.2143)	(0.0150)
	Skewed-t	λ	−0.1276	ν	9.1277	0.0313	0.0090	0.0007	0.9246	0.1218	0.0119	0.5149
			(0.0000)		(0.0000)	(0.0015)	(0.0000)	(0.9043)	(0.0000)	(0.0000)	(0.4450)	(0.7088)
	NIG	α	1.8365	β	−0.3826	0.0310	0.0091	0.0012	0.9241	0.1214	0.0104	0.4241
			(0.0000)		(0.0000)	(0.0683)	(0.0000)	(0.9346)	(0.0000)	(0.0000)	(0.6177)	(0.8238)
	MXN	β	−0.4997	δ	1.0674	0.0308	0.0092	0.0014	0.9238	0.1209	0.0097	0.3428
			(0.0000)		(0.0000)	(0.0011)	(0.0000)	(0.7782)	(0.0000)	(0.0000)	(0.7067)	(0.9113)
	VG	ν	0.3061	θ	−0.3556	0.0307	0.0093	0.0016	0.9237	0.1203	0.0088	0.2949
			(0.0000)		(0.0000)	(0.0027)	(0.0029)	(0.9070)	(0.0000)	(0.0000)	(0.8091)	(0.9325)
	CGMY	G	2.0319	M	2.7129	Y	0.1890				0.0130	0.9790
			(0.0000)		(0.0000)		(0.0000)				(0.3286)	(0.5288)
Model	Innovation					a_0	ω	ϕ	β_1	d	KS	AD
FIGARCH	Normal					0.0493	0.0201	0.0925	0.5112	0.4632	0.0335	13.3935
						(0.0000)	(0.0079)	(0.1333)	(0.0000)	(0.0000)	(0.0000)	(0.0005)
	Student's t	ν	7.8503			0.0624	0.0179	0.1003	0.5660	0.5101	0.0185	6.3841
			(0.0000)			(0.0000)	(0.0013)	(0.0190)	(0.0000)	(0.0000)	(0.0534)	(0.0025)
	Skewed-t	λ	−0.1225	ν	8.3506	0.0468	0.0160	0.0951	0.5682	0.5097	0.0208	3.4721
			(0.0000)		(0.0000)	(0.0000)	(0.0030)	(0.0255)	(0.0000)	(0.0000)	(0.0207)	(0.0188)
	NIG	α	1.7158	β	−0.3406	0.0463	0.0160	0.0933	0.5626	0.5061	0.0190	3.3503
			(0.0000)		(0.0000)	(0.0000)	(0.0017)	(0.0232)	(0.0000)	(0.0000)	(0.0440)	(0.0288)
	MXN	β	−0.4722	δ	0.9381	0.0460	0.0160	0.0931	0.5608	0.5041	0.0181	3.1265
			(0.0000)		(0.0000)	(0.0000)	(0.0039)	(0.0424)	(0.0000)	(0.0000)	(0.0625)	(0.0225)
	VG	ν	0.3478	θ	−0.3110	0.0459	0.0161	0.0934	0.5593	0.5023	0.0165	2.9357
			(0.0000)		(0.0000)	(0.0000)	(0.0317)	(0.1792)	(0.0000)	(0.0000)	(0.1121)	(0.0325)
	CGMY	G	1.8181	M	2.3808	Y	0.3070				0.0209	4.5936
			(0.0000)		(0.0000)		(0.0032)				(0.0197)	(0.0200)
MN-GARCH	p	μ_1	μ_2	a_0	ω_1	β_{11}	β_{21}	ω_2	β_{12}	β_{22}	KS	AD
	0.5977	0.1018	−0.1512	0.0445	0.0027	0.0463	0.9094	0.0109	0.1280	0.9142	0.0145	2.1104
	(0.0000)	(0.0000)		(0.0000)	(0.0401)	(0.0000)	(0.0000)	(0.0061)	(0.0000)	(0.0000)	(0.2192)	(0.0870)

This table reports the estimation results of GARCH-type models for MSCI World index returns. We consider the MN-GARCH model and three volatility specifications: GARCH, GJR and FIGARCH with either normal, Student's t, Skewed-t, NIG, MXN, VG or CGMY disturbances. $\nu > 2$ denotes the degree of freedom and $-1 < \lambda < 1$ is the asymmetry parameter for the skewed-t distribution. p-values for the t-test, calculated with heteroskedasticity-robust standard errors, are reported in parentheses. KS and AD are the Kolmogorov-Smirnov and the Anderson-Darling test statistics applied to standardized residuals, respectively. The p-values associated to these goodness-of-fit statistics tests are presented in parentheses.

1% significance level with the exception of the CGMY distribution under the FIGARCH model for the case of World index and also when combined with the GJR specification for EM returns. The AD test provides compelling evidence that amongst the non-normal distributions, the four Lévy distributions are fairly suitable for describing tail behavior of stock returns for all three markets. The Student's t distribution, however, is rejected at 1% significance level under both GARCH and FIGARCH models for World index returns, in which case, the estimated tail index is too low, *i.e.*, the tails are thinner than suggested by the Student's t.

4.3. VaR performance

To test the validity of the different distributional assumptions on index returns (*i.e.*, normal, Student's t, skewed-t, NCT, MN and non-normal Lévy), the predicted one-day-ahead 1% VaR for both long and short positions is compared to the actual

Table 4
Estimation results of GARCH-type models for MSCI EM index.

Model	Innovation					a_0	ω	β_1	β_2		KS	AD
GARCH	Normal					0.0615	0.0150	0.1077	0.8849		0.0329	10.1302
						(0.0000)	(0.0015)	(0.0000)	(0.0000)		0.0000	(0.0005)
	Student's t	ν	7.4143			0.0684	0.0104	0.0982	0.8987		0.0193	3.5466
			(0.0000)			(0.0000)	(0.0022)	(0.0000)	(0.0000)		0.0387	(0.0313)
	Skewed-t	λ	−0.1367	ν	7.7834	0.0399	0.0088	0.0879	0.9087		0.0145	1.4103
			(0.0000)		(0.0000)	(0.0015)	(0.0030)	(0.0000)	(0.0000)		0.2171	(0.1925)
	NIG	α	1.6583	β	−0.3557	0.0391	0.0089	0.0870	0.9091		0.0119	1.3304
			(0.0000)		(0.0000)	(0.0027)	(0.0063)	(0.0000)	(0.0000)		(0.4403)	(0.2413)
	MXN	β	−0.5145	δ	0.8891	0.0392	0.0091	0.0876	0.9083		0.0124	1.2721
			(0.0000)		(0.0000)	(0.0005)	(0.0015)	(0.0000)	(0.0000)		(0.3865)	(0.2263)
	VG	ν	0.3523	θ	−0.3381	0.0395	0.0093	0.0882	0.9072		0.0136	1.2519
			(0.0000)		(0.0000)	(0.0015)	(0.0200)	(0.0000)	(0.0000)		(0.2794)	(0.2588)
	CGMY	G	1.5367	M	2.1407	Y	0.5078				0.0187	3.7753
			(0.0000)		(0.0000)		(0.0000)				(0.0492)	(0.0525)
Model	Innovation					a_0	ω	β_1	β_2	β_3	KS	AD
GJR	Normal					0.0455	0.0198	0.0448	0.8900	0.0950	0.0298	10.8645
						(0.0004)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0005)
	Student's t	ν	8.1615			0.0583	0.0144	0.0463	0.9005	0.0818	0.0159	3.3480
			(0.0000)			(0.0000)	(0.0010)	(0.0000)	(0.0000)	(0.0000)	(0.1399)	(0.1265)
	Skewed-t	λ	−0.1216	ν	8.3384	0.0352	0.0121	0.0459	0.9111	0.0661	0.0094	0.5008
			(0.0000)		(0.0000)	(0.0014)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.7438)	(0.7300)
	NIG	α	1.7140	β	−0.3248	0.0356	0.0122	0.0460	0.9113	0.0648	0.0075	0.5011
			(0.0000)		(0.0000)	(0.0049)	(0.0019)	(0.0000)	(0.0000)	(0.0000)	(0.9273)	(0.7175)
	MXN	β	−0.4578	δ	0.9672	0.0355	0.0125	0.0461	0.9107	0.0652	0.0075	0.4407
			(0.0000)		(0.0000)	(0.0037)	(0.0005)	(0.0000)	(0.0000)	(0.0000)	(0.9242)	(0.8013)
	VG	ν	0.3263	θ	−0.3155	0.0356	0.0128	0.0460	0.9098	0.0662	0.0080	0.4137
			(0.0000)		(0.0000)	(0.0128)	(0.0280)	(0.0002)	(0.0000)	(0.0019)	(0.8898)	(0.8413)
	CGMY	G	2.0573	M	2.6233	Y	0.1580				0.0080	0.3186
			(0.0000)		(0.0000)		(0.0000)				(0.8857)	(0.9413)
Model	Innovation					a_0	ω	ϕ	β_1	d	KS	AD
FIGARCH	Normal					0.0646	0.0202	0.2394	0.6228	0.5212	0.0341	11.6279
						(0.0000)	(0.0232)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0005)
	Student's t	ν	7.7612			0.0719	0.0157	0.2394	0.6386	0.5211	0.0162	4.0250
			(0.0000)			(0.0000)	(0.0314)	(0.0000)	(0.0000)	(0.0000)	(0.1137)	(0.0125)
	Skewed-t	λ	−0.1330	ν	8.1399	0.0454	0.0131	0.2459	0.6420	0.5081	0.0152	1.9540
			(0.0000)		(0.0000)	(0.0001)	(0.0285)	(0.0000)	(0.0000)	(0.0000)	(0.1763)	(0.0988)
	NIG	α	1.7018	β	−0.3561	0.0447	0.0130	0.2473	0.6406	0.5055	0.0136	1.9215
			(0.0000)		(0.0000)	(0.0006)	(0.0028)	(0.0000)	(0.0000)	(0.0000)	(0.2862)	(0.0975)
	MXN	β	−0.5044	δ	0.9334	0.0446	0.0130	0.2475	0.6391	0.5049	0.0144	1.8406
			(0.0000)		(0.0000)	(0.0003)	(0.0315)	(0.0000)	(0.0000)	(0.0000)	(0.2239)	(0.0950)
	VG	ν	0.3377	θ	−0.3409	0.0447	0.0129	0.2481	0.6371	0.5038	0.0151	1.7971
			(0.0000)		(0.0000)	(0.0002)	(0.0101)	(0.0001)	(0.0000)	(0.0001)	(0.1794)	(0.0938)
	CGMY	G	1.5996	M	2.2230	Y	0.4942				0.0195	4.3731
			(0.0000)		(0.0000)		(0.0000)				(0.0370)	(0.0338)
MN-GARCH	p	μ_1	μ_2	a_0	ω_1	β_{11}	β_{21}	ω_2	β_{12}	β_{22}	KS	AD
	0.6770	0.0963	−0.2018	0.0436	0.0002	0.0388	0.9373	0.0661	0.2742	0.8170	0.0140	2.0364
	(0.0000)	(0.0000)		(0.0002)	(0.8344)	(0.0000)	(0.0000)	(0.0067)	(0.0000)	(0.0000)	(0.2559)	(0.1100)

This table reports the estimation results of GARCH-type models for MSCI EM index returns. We consider the MN-GARCH model and three volatility specifications: GARCH, GJR and FIGARCH with either normal, Student's t, Skewed-t, NIG, MXN, VG or CGMY disturbances. $\nu > 2$ denotes the degree of freedom and $-1 < \lambda < 1$ is the asymmetry parameter for the skewed-t distribution. p-values for the t-test, calculated with heteroskedasticity-robust standard errors, are reported in parentheses. KS and AD are the Kolmogorov-Smirnov and the Anderson-Darling test statistics applied to standardized residuals, respectively. The p-values associated to these goodness-of-fit statistics tests are presented in parentheses.

return. The model performance is tested on different volatility scenarios (crisis and post-crisis periods) by running various backtesting procedures with a key emphasis on the following four desirable properties: frequency, independence, duration and consistency of violations magnitude. The backtesting includes likelihood ratio (LR) tests for unconditional coverage (UC), multivariate UC (MUC, or equivalently, magnitude test) and conditional coverage (CC) as well as dynamic quantile (DQ) tests and duration-based (J) tests.[12] We use a rolling window that includes n previous days to derive recursive VaR forecasts. The rolling window technique updates the estimation sample regularly by incorporating new information reflected in each sample of the return series. Therefore, it can be argued that using a large history of data allows to implicitly take into account structural changes, such as mean and volatility shifts or changes in the distributional properties of the examined markets. For risk assessment during the crisis period, we consider a time window of about 10 years of historical data for World and EM, whereas an

[12] The specification of the tests is described in Appendix B.

Table 5
Estimation results of GARCH-type models for MSCI FM index.

Model	Innovation					a_0	ω	β_1	β_2		KS	AD
GARCH	Normal					0.0679	0.0039	0.0641	0.9326		0.0690	38.1605
						(0.0000)	(0.0176)	(0.0000)	(0.0000)		(0.0000)	0.0005
	Student's *t*	v	3.1705			0.0669	0.0094	0.0687	0.9256		0.0219	2.6006
			(0.0000)			(0.0000)	(0.0064)	(0.0001)	(0.0000)		(0.0807)	(0.0600)
	Skewed-*t*	λ	−0.0284	v	3.2025	0.0571	0.0092	0.0664	0.9270		0.0233	2.0423
			(0.0012)		(0.0000)	(0.0000)	(0.0036)	(0.0000)	(0.0000)		(0.0525)	(0.0888)
	NIG	α	0.7648	β	−0.0563	0.0524	0.0075	0.0558	0.9298		0.0199	1.1434
			(0.0000)		(0.0101)	(0.0000)	(0.0011)	(0.0000)	(0.0000)		(0.1420)	(0.2913)
	MXN	β	−0.1525	δ	0.2389	0.0528	0.0072	0.0550	0.9305		0.0196	1.0158
			(0.0586)		(0.0000)	(0.0000)	(0.0015)	(0.0000)	(0.0000)		(0.1531)	(0.3350)
	VG	v	0.9857	θ	−0.0047	0.0565	0.0065	0.0569	0.9291		0.0140	1.0470
			(0.0000)		(0.0000)	(0.0000)	(0.0128)	(0.0122)	(0.0000)		(0.5294)	(0.3238)
	CGMY	G	0.6050	M	0.6439	Y	0.8195				0.0174	1.8294
			(0.0000)		(0.0000)		(0.0000)				(0.2605)	(0.2613)
Model	Innovation					a_0	ω	β_1	β_2	β_3	KS	AD
GJR	Normal					0.0634	0.0050	0.0497	0.9289	0.0286	0.0668	37.3323
						(0.0000)	(0.0132)	(0.0000)	(0.0000)	(0.0324)	(0.0000)	(0.0005)
	Student's *t*	v	3.2014			0.0665	0.0125	0.0495	0.9134	0.0485	0.0209	1.6413
			(0.0000)			(0.0000)	(0.0138)	(0.0000)	(0.0000)	(0.0064)	(0.1085)	(0.1625)
	Skewed-*t*	λ	−0.0212	v	3.2199	0.0606	0.0121	0.0495	0.9153	0.0445	0.0218	1.8996
			(0.4866)		(0.0000)	(0.0000)	(0.0216)	(0.0000)	(0.0000)	(0.0023)	(0.0820)	(0.1025)
	NIG	α	0.7639	β	−0.0379	0.0564	0.0096	0.0421	0.9208	0.0343	0.0193	1.0495
			(0.0000)		(0.0933)	(0.0000)	(0.0056)	(0.0000)	(0.0000)	(0.0580)	(0.1654)	(0.3150)
	MXN	β	−0.1021	δ	0.2398	0.0565	0.0091	0.0412	0.9223	0.0333	0.0186	0.9066
			(0.3041)		(0.0000)	(0.0000)	(0.0054)	(0.0000)	(0.0000)	(0.0075)	(0.1973)	(0.4400)
	VG	v	0.9872	θ	0.0105	0.0595	0.0083	0.0413	0.9208	0.0368	0.0120	0.8552
			(0.0000)		(0.0000)	(0.0000)	(0.0005)	(0.0027)	(0.0000)	(0.0237)	(0.7198)	(0.4350)
	CGMY	G	0.9290	M	0.9554	Y	0.5269				0.0116	0.6344
			(0.0000)		(0.0000)		(0.0000)				(0.7590)	(0.7450)
Model	Innovation					a_0	ω	ϕ	β_1	d	KS	AD
FIGARCH	Normal					0.0700	0.0082	0.1728	0.7820	0.6544	0.0730	37.4387
						(0.0000)	(0.0433)	(0.0377)	(0.0000)	(0.0058)	(0.0000)	(0.0005)
	Student's *t*	v	3.2493			0.0670	0.0295	0.2840	0.6275	0.4319	0.0218	2.5953
			(0.0000)			(0.0000)	(0.0128)	(0.0000)	(0.0000)	(0.0000)	(0.0824)	(0.0388)
	Skewed-*t*	λ	−0.0316	v	3.2605	0.0583	0.0292	0.2863	0.6265	0.4275	0.0230	2.0684
			(0.0280)		(0.0000)	(0.0000)	(0.0114)	(0.0000)	(0.0000)	(0.0000)	(0.0572)	(0.0825)
	NIG	α	0.7481	β	−0.0493	0.0533	0.0183	0.3096	0.6096	0.3808	0.0192	1.1250
			(0.0000)		(0.0626)	(0.0000)	(0.0320)	(0.0000)	(0.0000)	(0.0003)	(0.1675)	(0.3050)
	MXN	β	−0.1348	δ	0.2306	0.0536	0.0163	0.3072	0.6148	0.3857	0.0187	1.0494
			(0.0453)		(0.0000)	(0.0000)	(0.0402)	(0.0000)	(0.0000)	(0.0003)	(0.1940)	(0.3513)
	VG	v	1.0017	θ	0.0031	0.0570	0.0109	0.2860	0.6425	0.4281	0.0154	1.2273
			(0.0000)		(0.0000)	(0.0000)	(0.1580)	(0.0013)	(0.0001)	(0.0139)	(0.4062)	(0.2650)
	CGMY	G	0.6104	M	0.6499	Y	0.8355				0.0177	1.6089
			(0.0000)		(0.0000)		(0.0000)				(0.2471)	(0.2925)
MN-GARCH	p	μ_1	μ_2	a_0	ω_1	β_{11}	β_{21}	ω_2	β_{12}	β_{22}	KS	AD
	0.7183	0.0238	−0.0607	0.0580	0.0023	0.0111	0.9545	0.0051	0.1654	0.9378	0.0210	1.9050
	(0.0000)	(0.0407)		(0.0000)	0.0462	(0.0137)	(0.0000)	(0.1742)	(0.0000)	(0.0000)	(0.1050)	(0.1045)

This table reports the estimation results of GARCH-type models for MSCI FM index returns. We consider the MN-GARCH model and three volatility specifications: GARCH, GJR and FIGARCH with either normal, Student's *t*, Skewed-*t*, NIG, MXN, VG or CGMY disturbances. $v > 2$ denotes the degree of freedom and $-1 < \lambda < 1$ is the asymmetry parameter for the skewed-*t* distribution. *p*-values for the *t*-test, calculated with heteroskedasticity-robust standard errors, are reported in parentheses. KS and AD are the Kolmogorov-Smirnov and the Anderson-Darling test statistics applied to standardized residuals, respectively. The *p*-values associated to these goodness-of-fit statistics tests are presented in parentheses.

historical window of about 5 years is used for FM, due the relatively shorter data sample. To conduct our forecasting analysis in the post-crisis period, we employ the crisis period (August 2007–May 2012) as the estimation sample. We also use a "stability window" of 10 days to update the model parameters.[13] Tables 6–8 summarize the results for the investigated VaR models applied to the three global indices.[14]

[13] In order to explore how often we need to update model parameters to obtain reliable VaR forecasts, a number of different "stability window" with lengths of 1, 5, 50, 100 and 250 days are also employed. All of results are too numerous to report in this paper. From our experiments, we note that daily updates, apart from being time-consuming, do not necessarily improve the forecasting performance whereas a much longer "stability window" deteriorates substantially the model ability to adjust for market moves.

[14] To save space, the results for the independence tests are not reported but are available upon request.

4.3.1. World index

For long trading positions, the results in Table 6 show that both the standard GARCH and the GJR models are strongly rejected for the different distributional assumptions, except for the Student's t distribution. The GARCH-Student's t model passes all the UC tests and provides satisfactory performance regarding the MUC, the Christoffersen (1998)'s CC test and the duration-based J test. Not surprisingly, the NCT-APARCH and the MN-GARCH models perform well, yet the latter fails the DQ_{CC} test. Under the FIGARCH specification, only the normal distribution has a very poor forecasting performance. The six non-normal models cannot be rejected at standard significance levels. The FIGARCH-Student's t is the best performing model with the highest p-values. Turning our attention to the short trading position, we conclude that, under all distributional assumptions, the GARCH model appears to be remarkably accurate, providing almost equal statistical sufficiency. The CGMY exhibits the best performance while the Student's t yields a much lower violation rate (0.4%) than expected and hence is less preferred regarding the UC tests. Under the GJR specification, the best performing models are the GJR-normal, GJR-VG, GJR-skewed-t and GJR-CGMY followed by the GJR-Student's t according to the magnitude test.

For these models, the hypothesis of correct CC cannot be rejected at 5% level, albeit UC tests show lower degree of VaR accuracy. The discrepancy between UC and CC tests indicates that VaR violations are not clustered, and that today's VaR violation is not significantly affected by the previous day's VaR violation. Besides, when the long memory is accounted for in the FIGARCH model, the Lévy distributions along with the skewed-t and the normal distributions, not only provide quite accurate VaR forecasts, but actually outperform highly competitive models while the Student's t distribution exhibits again the worst performance due to overestimation of short position risk.

4.3.2. EM index

The results in Table 7 show that non-normal VaR models perform accurately regardless of the volatility specification considered for long trading position. The EM index seems to be affected to a much lesser extent by the crisis period, as evidenced by the lower violation rate incurred by the investigated models. The GARCH-Student's t model, however, tends to overpredict risk as it involves overly conservative risk measures. The highest p-values from the four backtesting procedures are attributed to the non-normal GARCH models, thereby indicating the usefulness of the standard GARCH model in modeling long position risk in emerging markets. In addition, the GARCH-skewed-t, the GARCH-MXN the and the GARCH-VG models exhibit the best performance according to the magnitude test. The Lévy-based FIGARCH models provide equivalent VaR accuracy with the FIGARCH-skewed-t model. Although, the GJR and FIGARCH models with Student's t innovations are not rejected according to the LR and J tests, the picture is completely different when the DQ test is applied. Indeed, they are rejected by the most restrictive DQ_{CC} test. These rejections merely arise from violations of the independence test given that the DQ_{UC} test tends to validate the considered risk models. For short trading position, only the Student's t-based models are rejected once again due to the extremely low violation rate.

4.3.3. FM index

Table 8 provides evidence from the application of the backtests for the FM index. For long trading position, the normal distribution is strongly rejected regardless of the volatility model. The Student's t distribution is considerably outperformed by the skewed-t, the MN and the Lévy distributions. However, the FIGARCH-VG model is rejected by the magnitude test at 5% level. Overall, the Lévy distributions appear to be quite suitable for long position risk on FM markets during the crisis period, and the performance ranking for all of the distributional hypotheses is fairly stable across volatility specifications. However, the performance ranking is altered for short trading position with the normal distribution providing the best performance. The non-normal FIGARCH models are rejected. Regarding the combined results of the four backtests, only the skewed-t, the NIG, the MXN and the VG distributions show acceptable performance under the GARCH and GJR specifications, along with the MN-GARCH and the NCT-APARCH model. The Student's t VaR models have no violations; hence, we cannot perform the tests. Furthermore, the notable superiority of the normal distribution mainly arises from the extremely low empirical violation rate induced by non-normal distributions. This means that non-normal VaR models may overpredict future short position risk, or equivalently, the capital charges required would be unnecessary large.[15]

More importantly, the results highlight an apparent asymmetry in the performance of VaR models between long and short positions. This is mainly due to the skewed pattern of VaR violations induced by overreaction to bad news as most of the VaR violations are found in the left quantile during the global financial crisis. The overreaction of investors is not surprising during this period and makes the selection of accurate VaR models for both long and short positions challenging. Overall, the combined long and short VaR accuracy results across all of the models lead us to conclude the following:

- For the World index, the VaR estimates provided by the Lévy-based FIGARCH models are the most accurate with the CGMY distribution providing slightly better performance.
- For the EM index, the most accurate risk estimates are obtained by the GARCH model combined with either the MN, the skewed-t or the Lévy distributions.
- For the FM index, both the NIG and the MXN distributions provide the best performance when combined with the GARCH and the GJR specifications.

[15] See the discussion in Section 4.5.

Table 6
VaR backtesting results for MSCI World index.

Model	Innovation	Long trading position								Short trading position							
		N/Ns	LR_{MUC}	LR_{UC}	LR_{CC}	DQ_{UC}	DQ_{CC}	J_{UC}	J_{CC}	N/Ns	LR_{MUC}	LR_{UC}	LR_{CC}	DQ_{UC}	DQ_{CC}	J_{UC}	J_{CC}
GARCH	Normal	36/17	0.000	0.000	0.000	0.000	0.000	0.001	0.008	17/4	0.390	0.224	0.378	0.200	0.297	0.180	0.509
	Student's *t*	20/3	0.094	0.050	0.105	0.063	0.000	0.067	0.131	5/2	0.055	0.015	0.052	0.034	0.474	0.008	0.062
	Skewed-*t*	26/3	0.001	0.001	0.002	0.000	0.000	0.006	0.022	17/4	0.390	0.224	0.378	0.200	0.297	0.194	0.503
	NIG	26/3	0.001	0.001	0.002	0.000	0.000	0.006	0.023	17/4	0.390	0.224	0.378	0.200	0.297	0.177	0.516
	MXN	26/4	0.002	0.001	0.002	0.000	0.000	0.005	0.023	17/4	0.390	0.224	0.378	0.200	0.297	0.171	0.503
	VG	26/4	0.002	0.001	0.002	0.000	0.000	0.005	0.022	17/4	0.390	0.224	0.378	0.200	0.297	0.186	0.511
	CGMY	25/3	0.003	0.002	0.004	0.001	0.000	0.015	0.028	16/4	0.509	0.339	0.514	0.321	0.329	0.274	0.493
GJR	Normal	34/14	0.000	0.000	0.000	0.000	0.000	0.001	0.009	17/4	0.390	0.224	0.378	0.200	0.297	0.179	0.565
	Student's *t*	24/5	0.011	0.004	0.009	0.003	0.000	0.016	0.049	5/2	0.055	0.015	0.052	0.034	0.474	0.012	0.051
	Skewed-*t*	27/5	0.001	0.000	0.001	0.000	0.000	0.006	0.022	21/4	0.067	0.027	0.061	0.014	0.085	0.041	0.175
	NIG	26/5	0.002	0.001	0.002	0.001	0.000	0.007	0.028	24/4	0.010	0.004	0.009	0.001	0.012	0.016	0.059
	MXN	26/5	0.002	0.001	0.002	0.001	0.000	0.006	0.028	22/4	0.037	0.015	0.034	0.006	0.049	0.022	0.126
	VG	26/6	0.002	0.001	0.002	0.001	0.000	0.007	0.027	21/4	0.067	0.027	0.061	0.014	0.085	0.042	0.176
	CGMY	29/8	0.000	0.000	0.000	0.000	0.000	0.005	0.015	21/5	0.063	0.027	0.061	0.014	0.085	0.044	0.177
FIGARCH	Normal	30/12	0.000	0.000	0.000	0.000	0.000	0.005	0.013	12/3	0.942	0.888	0.881	0.892	0.993	0.960	0.931
	Student's *t*	13/1	0.418	0.885	0.863	0.882	0.267	0.763	0.753	5/0	0.025	0.015	0.052	0.034	0.474	0.007	0.049
	Skewed-*t*	18/2	0.160	0.142	0.261	0.115	0.237	0.129	0.431	13/1	0.418	0.885	0.863	0.873	0.987	0.710	0.967
	NIG	18/2	0.160	0.142	0.261	0.115	0.237	0.126	0.434	13/3	0.942	0.885	0.863	0.873	0.987	0.712	0.972
	MXN	18/2	0.160	0.142	0.261	0.115	0.237	0.122	0.434	13/3	0.942	0.885	0.863	0.873	0.987	0.698	0.970
	VG	18/2	0.160	0.142	0.261	0.115	0.237	0.130	0.432	13/3	0.942	0.885	0.863	0.873	0.987	0.736	0.971
	CGMY	17/2	0.257	0.224	0.378	0.199	0.284	0.168	0.595	12/3	0.942	0.888	0.881	0.892	0.993	0.956	0.912
MN-GARCH		19/6	0.098	0.086	0.170	0.092	0.001	0.082	0.063	19/3	0.160	0.086	0.170	0.048	0.373	0.070	0.093
NCT-APARCH		16/2	0.348	0.342	0.517	0.323	0.316	0.303	0.535	7/2	0.329	0.088	0.224	0.113	0.773	0.087	0.117

This table displays backtesting results of competing VaR models for MSCI World index during the crisis period from August 1, 2007 to May 14, 2012 (1250 trading days). Entries report the *p*-values associated to the magnitude test (LR_{MUC}), unconditional and conditional coverage likelihood ratio tests (LR_{UC}, LR_{CC}), dynamic quantile tests (DQ_{UC}, DQ_{CC}) and duration-based tests (J_{UC}, J_{CC}). The number of violations and *super* violations are denoted by N and Ns, respectively.

Table 7
VaR backtesting results for MSCI EM index.

Model	Innovation	Long trading position								Short trading position							
		N/Ns	LR_{MUC}	LR_{UC}	LR_{CC}	DQ_{UC}	DQ_{CC}	J_{UC}	J_{CC}	N/Ns	LR_{MUC}	LR_{UC}	LR_{CC}	DQ_{UC}	DQ_{CC}	J_{UC}	J_{CC}
GARCH	Normal	29/9	0.000	0.000	0.000	0.000	0.000	0.005	0.026	10/3	0.676	0.463	0.705	0.470	0.980	0.645	0.809
	Student's *t*	6/1	0.164	0.040	0.118	0.155	0.000	0.043	0.023	3/1	0.010	0.001	0.005	0.009	0.202	0.001	0.007
	Skewed-*t*	13/3	0.942	0.885	0.863	0.871	0.230	0.720	0.987	10/3	0.676	0.463	0.705	0.470	0.980	0.654	0.561
	NIG	11/3	0.847	0.665	0.826	0.715	0.116	0.877	0.816	11/4	0.485	0.665	0.826	0.667	0.991	0.847	0.224
	MXN	12/3	0.942	0.888	0.881	0.916	0.174	0.900	0.967	11/3	0.847	0.665	0.826	0.667	0.991	0.850	0.225
	VG	12/3	0.942	0.888	0.881	0.916	0.174	0.907	0.971	11/3	0.847	0.665	0.826	0.667	0.991	0.843	0.222
	CGMY	11/3	0.847	0.665	0.826	0.715	0.116	0.882	0.687	8/3	0.282	0.172	0.373	0.194	0.883	0.232	0.352
GJR	Normal	27/9	0.000	0.000	0.001	0.000	0.000	0.006	0.022	13/5	0.338	0.885	0.863	0.873	0.987	0.701	0.743
	Student's *t*	9/2	0.671	0.296	0.543	0.393	0.030	0.473	0.421	4/1	0.032	0.005	0.019	0.018	0.324	0.003	0.022
	Skewed-*t*	15/3	0.712	0.489	0.656	0.475	0.294	0.387	0.878	12/2	0.936	0.888	0.881	0.892	0.993	0.913	0.806
	NIG	14/3	0.856	0.674	0.781	0.663	0.273	0.494	0.908	12/2	0.936	0.888	0.881	0.892	0.993	0.911	0.803
	MXN	14/3	0.856	0.674	0.781	0.663	0.273	0.492	0.909	12/2	0.936	0.888	0.881	0.892	0.993	0.904	0.802
	VG	14/4	0.674	0.674	0.781	0.663	0.273	0.497	0.906	12/3	0.942	0.888	0.881	0.892	0.993	0.914	0.810
	CGMY	15/4	0.612	0.489	0.656	0.475	0.294	0.375	0.879	12/3	0.942	0.888	0.881	0.892	0.993	0.947	0.756
FIGARCH	Normal	23/8	0.006	0.015	0.034	0.008	0.030	0.019	0.134	6/3	0.056	0.040	0.118	0.064	0.632	0.037	0.051
	Student's *t*	8/1	0.394	0.172	0.373	0.257	0.022	0.235	0.281	2/0	0.001	0.000	0.001	0.006	0.114	0.000	0.001
	Skewed-*t*	10/2	0.840	0.463	0.705	0.516	0.099	0.627	0.630	7/1	0.274	0.089	0.226	0.114	0.775	0.085	0.115
	NIG	10/2	0.840	0.463	0.705	0.516	0.099	0.644	0.642	7/1	0.274	0.089	0.226	0.114	0.775	0.110	0.124
	MXN	10/2	0.840	0.463	0.705	0.516	0.099	0.638	0.639	7/1	0.274	0.089	0.226	0.114	0.775	0.109	0.127
	VG	10/2	0.840	0.463	0.705	0.516	0.099	0.643	0.633	7/1	0.274	0.089	0.226	0.114	0.775	0.095	0.119
	CGMY	11/2	0.935	0.665	0.826	0.696	0.160	0.855	0.575	6/1	0.164	0.040	0.118	0.064	0.632	0.037	0.051
MN-GARCH		11/5	0.189	0.665	0.826	0.715	0.116	0.882	0.824	13/4	0.675	0.885	0.863	0.873	0.987	0.682	0.482
NCT-APARCH		12/1	0.479	0.884	0.881	0.677	0.206	0.894	0.490	6/2	0.173	0.040	0.117	0.063	0.630	0.031	0.038

This table displays backtesting results of competing VaR models for MSCI EM index during the crisis period from August 1, 2007 to May 14, 2012 (1250 trading days). Entries report the *p*-values associated to the magnitude test (LR_{MUC}), unconditional and conditional coverage likelihood ratio tests (LR_{UC}, LR_{CC}), dynamic quantile tests (DQ_{UC}, DQ_{CC}) and duration-based tests (J_{UC}, J_{CC}). The number of violations and *super* violations are denoted by N and Ns, respectively.

Table 8
VaR backtesting results for MSCI FM index.

Model	Innovation	Long trading position								Short trading position							
		N/Ns	LR_{MUC}	LR_{UC}	LR_{CC}	DQ_{UC}	DQ_{CC}	J_{UC}	J_{CC}	N/Ns	LR_{MUC}	LR_{UC}	LR_{CC}	DQ_{UC}	DQ_{CC}	J_{UC}	J_{CC}
GARCH	Normal	29/16	0.000	0.000	0.000	0.000	0.000	0.005	0.010	12/5	0.270	0.888	0.881	0.892	0.993	0.996	0.958
	Student's *t*	5/0	0.025	0.015	0.052	0.034	0.474	0.011	0.012	0/0							
	Skewed-*t*	15/3	0.712	0.489	0.307	0.479	0.326	0.334	0.670	6/1	0.164	0.040	0.118	0.064	0.632	0.037	0.012
	NIG	15/5	0.375	0.489	0.307	0.479	0.326	0.330	0.664	7/1	0.274	0.089	0.226	0.114	0.775	0.114	0.279
	MXN	15/5	0.375	0.489	0.307	0.479	0.326	0.331	0.669	7/1	0.274	0.089	0.226	0.114	0.775	0.114	0.281
	VG	16/7	0.065	0.339	0.276	0.320	0.318	0.288	0.632	6/2	0.139	0.040	0.118	0.064	0.632	0.041	0.160
	CGMY	12/4	0.607	0.888	0.258	0.900	0.213	0.964	0.659	4/1	0.032	0.005	0.019	0.018	0.324	0.006	0.011
GJR	Normal	27/16	0.000	0.000	0.001	0.000	0.000	0.005	0.018	11/5	0.189	0.665	0.826	0.667	0.991	0.779	0.998
	Student's *t*	4/0	0.009	0.005	0.019	0.018	0.324	0.003	0.002	0/0							
	Skewed-*t*	13/2	0.851	0.885	0.299	0.882	0.269	0.692	0.699	6/0	0.060	0.040	0.118	0.064	0.632	0.030	0.163
	NIG	14/5	0.375	0.674	0.317	0.670	0.309	0.542	0.635	7/3	0.140	0.089	0.226	0.114	0.775	0.106	0.283
	MXN	15/5	0.375	0.489	0.307	0.479	0.326	0.319	0.666	7/3	0.140	0.089	0.226	0.114	0.775	0.114	0.281
	VG	17/7	0.065	0.224	0.230	0.200	0.287	0.168	0.571	7/3	0.140	0.089	0.226	0.114	0.775	0.110	0.274
	CGMY	12/6	0.086	0.888	0.881	0.892	0.993	0.998	0.719	6/1	0.164	0.040	0.118	0.064	0.632	0.036	0.124
FIGARCH	Normal	28/17	0.000	0.000	0.000	0.000	0.000	0.006	0.016	9/6	0.010	0.296	0.543	0.311	0.948	0.407	0.821
	Student's *t*	5/0	0.025	0.015	0.052	0.034	0.474	0.007	0.006	0/0							
	Skewed-*t*	16/5	0.340	0.339	0.276	0.320	0.318	0.276	0.642	4/0	0.009	0.005	0.019	0.018	0.324	0.006	0.009
	NIG	14/5	0.375	0.674	0.317	0.670	0.309	0.537	0.638	4/0	0.009	0.005	0.019	0.018	0.324	0.005	0.009
	MXN	14/5	0.375	0.674	0.317	0.670	0.309	0.525	0.639	4/0	0.009	0.005	0.019	0.018	0.324	0.004	0.010
	VG	18/8	0.022	0.142	0.179	0.137	0.085	0.124	0.424	4/0	0.009	0.005	0.019	0.018	0.324	0.005	0.011
	CGMY	11/5	0.189	0.665	0.201	0.694	0.150	0.836	0.485	3/0	0.002	0.001	0.005	0.009	0.202	0.000	0.003
MN-GARCH		15/6	0.170	0.489	0.307	0.479	0.326	0.374	0.633	7/1	0.274	0.089	0.226	0.114	0.775	0.110	0.287
NCT-APARCH		15/6	0.170	0.489	0.307	0.479	0.326	0.365	0.439	4/2	0.016	0.005	0.019	0.018	0.324	0.004	0.010

This table displays backtesting results of competing VaR models for MSCI FM index during the crisis period from August 1, 2007 to May 14, 2012 (1250 trading days). Entries report the *p*-values associated to the magnitude test (LR_{MUC}), unconditional and conditional coverage likelihood ratio tests (LR_{UC}, LR_{CC}), dynamic quantile tests (DQ_{UC}, DQ_{CC}) and duration-based tests (J_{UC}, J_{CC}). The number of violations and *super* violations are denoted by N and Ns, respectively.

Table 9
VaR forecasting performance in the post-crisis period.

Model	Innovation	Long trading position						Short trading position					
		World		EM		FM		World		EM		FM	
		N/Ns	Accuracy	N/Ns	Accuracy	N/Ns	Accuracy	N/Ns	Accuracy	N/Ns	Accuracy	N/Ns	Accuracy
GARCH	Normal	12/5	X	8/2	X	13/6	LR_{UC}, LR_{CC}	7/2	X	3/2	LR, DQ	9/4	LR, *J*
	Student's *t*	2/0		3/1	X	3/0	X	1/0		2/0		3/1	X
	Skewed-*t*	7/1	X	3/1	X	8/0	LR, *J*	7/0	X	7/3	X	7/3	LR, *J*
	NIG	5/1	X	3/1	X	6/1	LR, *J*	7/2	X	7/3	X	8/3	LR, *J*
	MXN	6/1	X	3/1	X	7/1	LR, *J*	7/2	X	7/3	X	8/3	LR, *J*
	VG	7/1	X	3/1	X	8/2	LR, *J*	7/2	X	6/3	X	8/3	LR, *J*
	CGMY	5/1	X	5/1	X	7/0	LR, *J*	7/0	X	10/2	X	6/3	LR, *J*
GJR	Normal	8/3	X	7/2	X	12/5	LR, *J*	5/2	X	3/2	LR, DQ	11/4	LR, *J*
	Student's *t*	2/0		3/1	X	3/0	X	2/0	3/1	LR, DQ	3/1	X	
	Skewed-*t*	4/1	X	3/1	X	7/0	LR, *J*	5/2	X	6/2	X	7/3	LR, *J*
	NIG	3/1	X	3/1	X	6/1	LR, *J*	5/2	X	6/2	X	8/3	LR, *J*
	MXN	3/1	X	3/1	X	6/2	LR, *J*	5/2	X	6/2	X	7/3	LR, *J*
	VG	4/1	X	3/1	X	6/2	LR, *J*	5/2	X	6/2	X	8/3	LR, *J*
	CGMY	4/1	X	3/1	X	7/1	X	5/2	X	4/2	X	6/3	LR, *J*
FIGARCH	Normal	15/4	DQ, *J*	9/4	X	13/5	LR_{UC}, LR_{CC}, J	8/2	X	5/2	X	7/5	LR_{UC}, LR_{CC}, DQ, *J*
	Student's *t*	3/1	X	4/2	X	2/0		1/0		3/1	LR, DQ	3/0	X
	Skewed-*t*	6/2	X	4/2	X	8/1	LR, *J*	7/0	X	13/3	LR, *J*	10/4	X
	NIG	5/2	X	4/2	X	8/1	LR, *J*	8/1	X	12/3	X	11/5	X
	MXN	5/2	X	4/2	X	8/1	LR, *J*	8/1	X	12/3	X	10/5	X
	VG	6/2	X	4/3	X	8/2	LR, *J*	8/1	X	12/3	X	10/5	X
	CGMY	6/2	X	4/2	X	7/1	LR, *J*	4/1	X	9/2	X	6/5	LR_{UC}, LR_{CC}, DQ, *J*
MN-GARCH		5/1	X	4/1	X	5/2	X	7/2	X	8/3	X	10/5	X
NCT-APARCH		3/1	X	4/1	X	2/0		3/0	LR, DQ	4/3	LR_{UC}, LR_{CC}, DQ, *J*	3/2	X

This table summarizes the accuracy of VaR models during the post-crisis period from May 15, 2012 to March 10, 2015 (730 trading days) with respect to the magnitude test, UC test of Kupiec (1995) as well as the hypothesis of correct CC for the likelihood ratio test of Christoffersen (1998), the dynamic quantile test of Engle and Manganelli (2004) and the duration test of Candelon et al. (2011). The number of violations and *super* violations are denoted by *N* and *Ns*, respectively. The symbol X is assigned to VaR models which pass all of the five tests for the respective market. The LR figure indicates VaR accuracy with respect to all of the likelihood ratio tests (LR_{MUC}, LR_{UC} and LR_{CC}) whereas DQ and *J* designate models that fulfill the hypothesis of correct CC for the dynamic quantile and duration tests, respectively. Empty cells denote cases for which we cannot perform accuracy tests given insufficient number of violations.

Table 10
Average daily capital requirements under the Basel II Accord.

Model	Innovation	Long trading position						Short trading position					
		World		EM		FM		World		EM		FM	
		Crisis	Post-crisis	Crisis	Post-crisis	Crisis	Post-crisis	Crisis	Post-crisis	Crisis	Post-crisis	Crisis	Post-crisis
GARCH	Normal		5.191		5.999		3.229	9.933	5.363	12.354	5.917	7.432	3.317
	Student's *t*	11.985			7.175		5.268						5.529
	Skewed-*t*		5.867	13.866	6.497	9.413	4.023	9.849	5.434	12.043	5.648	8.221	3.398
	NIG		5.963	14.120	6.576	9.557	4.115	9.801	5.429	11.917	5.685	8.061	3.384
	MXN		5.942	14.020	6.565	9.385	4.038	9.830	5.470	11.978	5.691	8.087	3.422
	VG		5.905	13.838	6.526	8.765	3.818	9.871	5.491	12.027	5.704	8.310	3.414
	CGMY		5.852	13.933	6.813	10.492	4.105	9.690	5.631	12.335	6.228		3.634
GJR	Normal		**5.102**		**5.791**		3.231	9.726	5.230	12.151	5.873	7.463	3.264
	Student's *t*			14.491	6.773		5.130				6.947		5.393
	Skewed-*t*		5.998	13.433	6.597	9.757	3.919	9.843	5.255	12.109	5.641	8.596	3.379
	NIG		6.120	13.645	6.644	9.684	4.017		5.245	12.018	5.698	8.263	3.367
	MXN		6.087	13.567	6.640	9.509	3.938		5.280	12.065	5.693	8.273	3.337
	VG		6.008	13.423	6.615	**8.662**	3.705	9.791	5.289	12.106	5.695	8.056	3.399
	CGMY		5.969	13.382	6.581	10.781	4.001	**9.617**	5.437	12.480	5.874	9.143	3.592
FIGARCH	Normal				5.809		**3.144**	10.140	**4.958**	12.307	5.704	**7.407**	**3.176**
	Student's *t*	11.491	6.247	15.011	6.558						6.862		4.965
	Skewed-*t*	**10.938**	5.433	13.839	6.279	9.5383.672		10.045	4.996	11.968	**5.424**		3.284
	NIG	11.081	5.526	14.110	6.386	10.158	3.824	9.996	5.008	**11.880**	5.434		3.272
	MXN	11.084	5.497	14.049	6.364	9.995	3.772	10.035	5.035	11.949	5.445		3.300
	VG	11.080	5.433	13.915	6.312	9.113	3.592	10.079	5.048	12.009	5.458		3.335
	CGMY	11.087	5.490	13.830	6.263	10.372	3.687	9.878	5.130	12.275	5.540		3.263
MN-GARCH		10.987	6.128	**13.173**	6.535	9.591	4.048	10.036	5.364	12.341	5.585	9.307	3.371
NCT-APARCH		11.065	5.734	13.905	6.281	9.861		10.781	5.753	12.901	6.781	9.648	5.077

This table reports the average daily capital requirements (ADCR) incurred by accurate VaR models for each evaluation sample (crisis and post-crisis). Bold numbers indicate the most appropriate models in terms of minimum ADCR.

4.4. Post-crisis period

A summary of backtesting results for the post-crisis period is presented in Table 9. The *LR* figure indicates VaR accuracy with respect to all of the likelihood ratio tests (LR_{MUC}, LR_{UC} and LR_{CC}) whereas DQ and *J* designate models that fulfill the hypothesis of correct CC for the dynamic quantile and duration tests, respectively. The symbol X is used in the table to denote VaR models which pass all of five tests for the respective market. Empty cells denote either rejection of UC due to overestimation of the realized VaR, or cases for which we cannot compute the CC tests because of an insufficient number of violations.

Contrary to the findings from the crisis period, the 1% loss quantile seems to be more predictable, during the post-crisis period, as statistical sufficiency is achieved effortlessly by a larger number of the investigated VaR models. Even the normal distribution rivals the more sophisticated and flexible distributions. In line with Bao et al. (2006) and Dimitrakopoulos et al. (2010), this result suggests that while market turmoil hurdles against making accurate risk forecasts, most of VaR models behave similarly well during tranquil periods.

In the case of World index, the performance of GARCH-type models improves when the estimation sample includes the crisis period for both long and short trading positions. Under each of the three volatility specifications, the Student's *t* distribution, however, exhibits the lowest violation rate, and most of the accuracy tests could not be conducted. Consistent with the results obtained from the crisis period, risk misspecification is mainly due to VaR overestimation for the Student's *t* time series models.

For the EM index, all of the VaR models perform accurately, as they cannot be rejected by almost all of the selected backtesting methods. Nevertheless, under the GARCH and the GJR specifications, the normal distribution, along with the GJR-Student's *t*, the FIGARCH-Student's *t* and the NCT-APARCH models fail to pass the *J* test for the short position risk. Similar results emerge for the FIGARCH-skewed-*t* model according to the DQ test.

For the FM index, the MN-GARCH model and the Student's *t* distribution in combination with both the GARCH and the GJR models provide the most reliable VaR forecasts, as evidenced by their statistical sufficiency according to all of the accuracy tests. Besides, as far as the DQ_{CC} test is concerned, the normal, the skewed-*t* and the non-normal Lévy distributions in combination with either the GARCH or the GJR model are rejected for both long and short positions, whereas the FIGARCH specification is only rejected for long position risk. Hence, violation clustering seems to be more pronounced on frontier markets. The normal as well as the non-normal Lévy-based models perform accurately, albeit some of them are rejected by the magnitude test, namely the GARCH-normal and the FIGARCH-normal for long position risk, along with the FIGARCH-CGMY and the FIGARCH-normal for short position risk. Even if these models are validated on the basis of the violation rate, the frequency of *super* violations is abnormally high which indicates that the magnitude of losses beyond VaR is too large.

4.5. VaR and Basel capital charges

One of the main implications of VaR is the computation of the market risk capital requirements as defined by the Basel II Accord (Basel Committee on Banking Supervision, 2006), and representing the daily minimum capital requirements which must be met by financial institutions to face market risk. Formally, the risk-capital charges on a daily basis are defined as follows:

$$\mathrm{DCR}_t = \mathrm{Max}\left\{(3+k)\frac{1}{60}\sum_{i=1}^{60}\mathrm{VaR}_{t-i};\mathrm{VaR}_{t-1}\right\}, \tag{15}$$

where DCR_t denotes the daily capital requirements at day t, VaR_{t-1} designates the VaR for day $t-1$ and k is a penalty factor depending on the number of VaR violations in the previous 250 trading days, as described in Appendix C.

Table 10 reports the average daily capital requirements (ADCR) implied by models that at least pass the LR_{UC} at 5% level over the two forecasting periods. Our two forecasting periods totalize 1250 and 730 observations, but the DCR are computed for only 1000 and 480 rolling windows, respectively, as they require the estimation of the number of violations over the previous 250 trading days. From a practical perspective, ADCR values are related to the cost of risk management. Specifically, the smaller ADCR, the more economically efficient is the VaR model; provided that it fulfills statistical sufficiency. Note that reporting ADCR values for inaccurate models with high violation rates may be misleading as they systematically generate lower capital requirements given that the implied VaR tends to underestimate the actual loss.

In line with our former finding that normal time series models perform adequately during the post-crisis period, the results in Table 10 show that they are less demanding, in terms of capital requirements. The sole fact supporting non-normal models is the lowest ADCR value implied by the FIGARCH-skewed-*t* model when it comes to provision against short position risk in emerging markets. However, during market downturn, the MN-GARCH and the FIGARCH-NIG are the best models for long and short positions, respectively, as they yield the lowest ADCR in emerging markets. In the case of frontier markets, the VG-GJR is the best performer with this regard for long position risk, whereas substantial savings in daily capital charges are incurred by the FIGARCH-normal for short position risk. Similar findings hold for the FIGARCH-skewed-*t* and the GJR-CGMY in developed markets for long and short positions, respectively.

The benefit of using daily capital requirements (DCR) is twofold. First, the DCR can be used as a further measure for model selection since it provides meaningful information about savings incurred by VaR models with equal accuracy. These savings, plotted in Fig. 1 as the differences between DCR implied by the FIGARCH-VG and both FIGARCH-NIG and FIGARCH-MXN

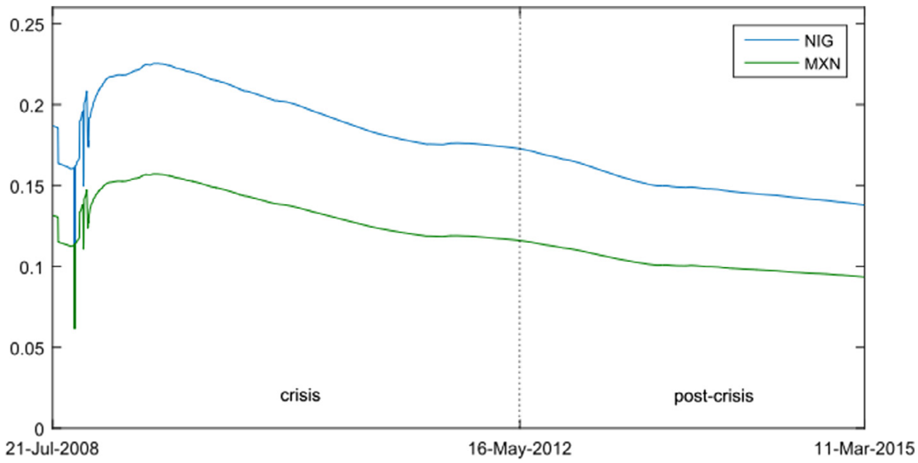

Fig. 1. The DCR differences between the FIGARCH-VG and FIGARCH-NIG models (blue line), and the FIGARCH-VG and FIGARCH-MXN models (green line) for long position risk in emerging markets. Dotted line divides the crisis and post-crisis period. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

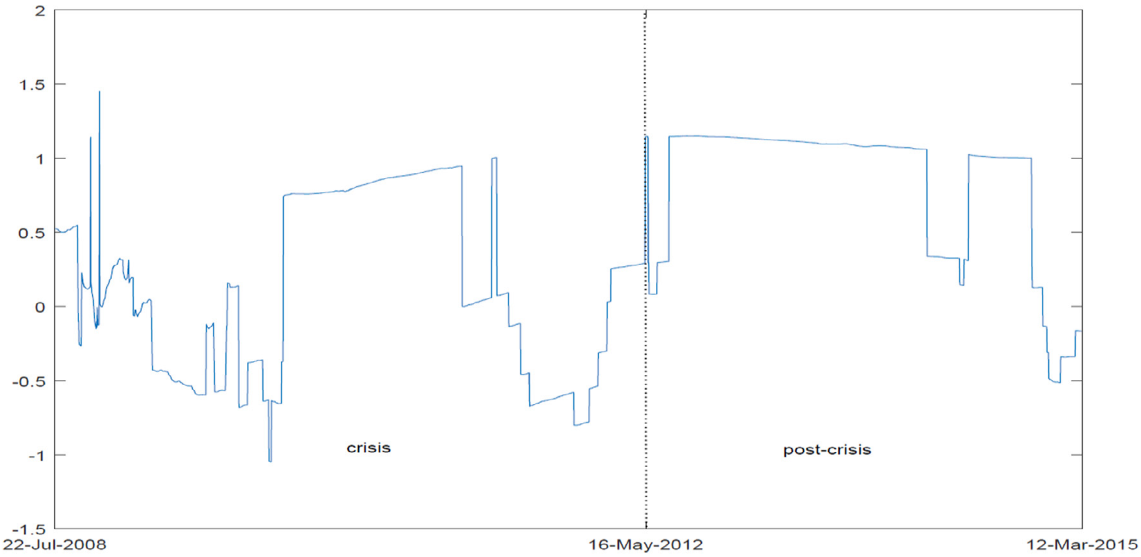

Fig. 2. The differences between DCR implied by the GJR-VG and the FIGARCH-normal models for long position risk in frontier markets. Dotted line divides the crisis and post-crisis period.

models for long position risk in emerging markets, represent an improvement for the efficient allocation of regulatory capital. Second, as the DCR is straightforwardly related to the cost of risk management, it can help internal model builders to decide on the trade-off between correctly specified VaR and smooth capital requirements. Fig. 2 illustrates the cost of risk management in terms of additional DCR that should be provisioned by financial institutions if they implement the GJR-VG (the most efficient model in terms of ADCR during the crisis period) instead of the FIGARCH-normal model (with the lowest ADCR in the post-crisis period) for long position VaR in the case of frontier markets. Even if this cost exhibits a somewhat unstable dynamics in the volatile period, its relatively high level after the full unfolding of the global financial crisis shows that the use of non-normal models would largely improve the stability of an investment in the case of a collapse in financial markets, at the expense of inefficient regulatory capital allocation during periods of calm.

5. Conclusion

Risk assessment is an important and complex task faced by market regulators and financial institutions, especially after the last subprime crisis. It is argued that since market data is endogenous to market behavior, statistical analysis made in times of stability does not provide much guidance in times of crisis. Consequently, we conduct our empirical investigation on the accuracy of parametric VaR models during stressed financial markets with stringent view of VaR models' performance in terms of frequency, independence, duration and magnitude of violations.

Specifically, we examine the forecasting power of three conditional volatility models and seven distributional assumptions of the error term in estimating the VaR of global stock market indices for both long and short trading positions. In order to investigate the effects of distribution assumptions, we compare the performance of the normal, the Student's *t*, the skewed-*t* and four non-normal Lévy distributions (NIG, MXN, VG and CGMY) under each volatility specification. If the performance ranking of the seven error distributions depends crucially on the specific conditional volatility model, then the choice of a given distribution should be contingent on which volatility specification we use. On the other hand, if the performance rankings are the same under each of the three volatility specifications, then we would conclude that in forecasting the VaR, the superiority of one distribution over the others, for modeling return innovations, is unconditional and robust to variations in the volatility specification.

A number of important findings emerge from this paper. First, for long position risk, the performance ranking of the underlying distributions is irrelevant to the specific volatility model. Risk misspecification in the case of the normal distribution is mainly due to underestimation of VaR, in contrast to the Student's *t* distribution where the majority of risk misspecifications are attributed to overestimation of VaR. Both skewed-*t* and Lévy distributions seem to perform fairly well across developed, emerging and frontier markets. However, the picture arising from short position is completely different lending strong support for the use of normal time series models. The accuracy of non-normal Lévy distributions strongly depends on the conditional volatility model.

Second, although the normal distribution might be useful for describing the risk in low volatility state, the results show that it might provide an unsatisfactory approximation in high volatility state. The outperformance of the non-normal Lévy distributions is clearer during the crisis period and their predictive ability looks very close to that of highly competitive models (MN-GARCH and NCT-APARCH).

Third, the performance of VaR models depends on the examined market. To some extent, our results support the Basel Accord's tolerance for financial institutions that invest in global markets, to build internal models to forecast the VaR; since none of the investigated volatility specifications could be established as a "universal" model in capturing the risk of both developed and developing market portfolios. The relevant models show evidence of long memory in developed markets, which suggests that the FIGARCH model is preferred to the GARCH and GJR models. The GJR and GARCH are the most relevant specifications in capturing risk in frontier and emerging markets, respectively. This implies that risk managers should favor models which cater for asymmetry when they examine frontier markets.

Fourth, special care should be taken when estimating VaR in crises periods, since VaR consistency across long and short trading positions is hardly achieved. The spectrum of models that accurately describe the joint behavior of both long and short positions is much reduced in the case of developed markets than that in emerging and frontier markets. This features more stability in developing markets during the sovereign debt crisis (Del Brio et al., 2014).

Fifth, statistical accuracy of GARCH-type models combined with the investigated skewed and leptokurtic distributions may exonerate financial institutions from using more complicated approaches such as extreme value theory (EVT), a method that is not without shortcomings (Diebold et al., 2000), which focuses only on the left tail, at least when including a sufficiently large estimation sample.

Finally, as far as regulatory capital allocation for equity exposure is concerned, there is no real incentive for financial institutions to use the more sophisticated non-normal models during periods of market calm. However, capital savings incurred by normal models may have significant repercussions when a market collapse occurs. The capital charges required for the risk manager may not be enough to cover losses arising from market downturns.

Appendix A. Dataset description

Index	Description
MSCI World	The MSCI World index is a free float-adjusted market capitalization weighted index that is designed to measure the equity market performance of developed markets. The MSCI World index consists of the following 23 developed market country indices: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, and the United States. With 1635 large and mid cap constituents, the index covers approximately 85% of the free float-adjusted market capitalization in each country
MSCI Emerging Markets (EM)	The MSCI EM index is a free float-adjusted market capitalization index that is designed to measure equity market performance of emerging markets. The MSCI EM Index consists of the following 23 emerging market country indices: Brazil, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, India, Indonesia, Korea, Malaysia, Mexico, Peru, Philippines, Poland, Qatar, Russia, South Africa, Taiwan, Thailand, Turkey and United Arab Emirates. The index covers over 800 securities represents approximately 13% of world market capitalization

(*continued on next page*)

Appendix A. Dataset Description (*continued*)

Index	Description
MSCI Frontier Markets (FM)	The MSCI FM index is a free float-adjusted market capitalization index that is designed to measure equity market performance of frontier markets. The MSCI FM index consists of the following 24 frontier market country indices: Argentina, Bahrain, Bangladesh, Bulgaria, Croatia, Estonia, Jordan, Kenya, Kuwait, Lebanon, Lithuania, Morocco, Kazakhstan, Mauritius, Nigeria, Oman, Pakistan, Romania, Serbia, Slovenia, Sri Lanka, Tunisia, Ukraine, and Vietnam

Source: Thomson Reuters and https://www.msci.com/.

Appendix B. Backtesting of the VaR models

A variety of backtesting methods have been proposed to gauge the accuracy of VaR estimates (see Campbell, 2007; Ziggel et al., 2014, for a review). Backtesting is a formal statistical framework that consists in verifying if actual trading losses are in line with model-generated VaR forecasts and relies on testing over VaR violations (also called the *hit*). A violation is said to occur when the realized trading loss exceeds the VaR forecast. We briefly present the backtesting methods used in our empirical assessment of VaR models regarding the following properties: frequency, independence, duration and magnitude of violations.

1. Frequency: The unconditional coverage (UC) test (Kupiec, 1995) is the industry standard mostly due to the fact that it is implicitly incorporated in the "traffic Light" system proposed by the Basel Committee on Banking Supervision (2006, 2009) which remains the reference backtest methodology for banking regulators. The test consists of examining if the realized coverage rate equals the theoretical coverage rate (α) of the VaR for a backtesting sample of T non-overlapping observations. This is equivalent to testing if the hit variable $I_t(\alpha)$, which takes values of 1 if the loss exceeds the reported VaR measure and 0 otherwise, follows a binomial distribution with parameter α. Under the UC hypothesis, the likelihood ratio (LR) test statistic follows a χ^2 distribution with one degree of freedom. That is:
$$LR_{UC}(\alpha) = -2\ln\left[(1-\alpha)^{T-N}\alpha^N\right] + 2\ln\left[\left(1-\frac{N}{T}\right)^{T-N}\left(\frac{N}{T}\right)^N\right] \sim \chi^2(1), \tag{B.1}$$
where N is the number of VaR violations.
2. Independence: An enhancement of the unconditional backtesting framework is achieved by additionally testing for the independence (IND) of the sequence of VaR violations yielding a combined test of conditional coverage (CC). For risk model assessment under the joint hypothesis of IND and correct UC, we use the LR test of Christoffersen (1998) and the regression-based test of Engle and Manganelli (2004). The Christoffersen (1998)'s LR test for independence against an explicit first-order Markov alternative is given by:
$$LR_{IND}(\alpha) = -2\ln\left[\left(1-\frac{N}{T}\right)^{T-N}\left(\frac{N}{T}\right)^N\right] + 2\ln\left[(1-\hat{\pi}_{01})^{n_{00}}\hat{\pi}_{01}^{n_{01}}(1-\hat{\pi}_{11})^{n_{10}}\hat{\pi}_{11}^{n_{11}}\right] \sim \chi^2(1), \tag{B.2}$$
where $n_{ij};\ i,j = 0,1$ is the number of times we have $I_t(\alpha) = j$ and $I_{t-1}(\alpha) = i$ with $\hat{\pi}_{01} = n_{01}/(n_{00}+n_{01})$ and $\hat{\pi}_{11} = n_{11}/(n_{10}+n_{11})$. The LR statistic for the CC test is then given by:
$$LR_{CC}(\alpha) = LR_{UC}(\alpha) + LR_{IND}(\alpha) \sim \chi^2(2). \tag{B.3}$$
In order to consider higher-order dependence in the violation series, Engle and Manganelli (2004) propose a linear regression model linking current violations to past violations and a set of chosen exogenous variables so as to test the conditional efficiency hypothesis. The test involves the estimation of the following regression:
$$Hit_t(\alpha) = \lambda + \sum_{k=1}^{K}\beta_k Hit_{t-k}(\alpha) + \sum_{k=1}^{K}\gamma_k x_{t-k}(\alpha) + u_t, \tag{B.4}$$
where $Hit_t(\alpha) = I_t(\alpha) - \alpha$ is the de-meaned indicator function at time t, u_t is an *i.i.d.* process and x_{t-k} are exogenous regressors from the information set $\mathcal{F}_{t-1}$. Testing for the null hypothesis of independence is equivalent to testing the joint nullity of the coefficients $\beta_k, \gamma_k,\ \forall k = 1,\ldots,K$ while the UC hypothesis is fulfilled if the intercept λ is null. The joint nullity test of all coefficients, including the intercept, thus corresponds to a conditional efficiency test. The test statistic associated with the CC test is given by:
$$DQ_{CC} = \frac{\hat{\Omega}'X'X\hat{\Omega}}{\alpha(1-\alpha)} \sim \chi^2(2K+1), \tag{B.5}$$

where X is the matrix of explanatory variables and $\hat{\Omega}$ denotes the OLS estimator of the parameters vector $\Omega = (\lambda, \beta_1, \ldots, \beta_K, \gamma_1, \ldots, \gamma_K)$ in the regression model (B.4). In the empirical application, we follow Engle and Manganelli (2004) by using a constant, four lagged *Hits* and the contemporaneous VaR estimates as regressors.

3. Duration: The intuition behind duration tests is that the clustering of violations will induce an excessive number of relatively short and long no-hit durations. Candelon et al. (2011) recently introduced a new duration-based test using orthonormal polynomials and the GMM test framework proposed by Bontemps and Meddahi (2005, 2012). Under the null hypothesis of CC, the test statistic is given by:

$$J_{CC}(p) = \left(\frac{1}{\sqrt{m}}\sum_{i=1}^{m} M(d_i;\alpha)\right)^T \left(\frac{1}{\sqrt{m}}\sum_{i=1}^{m} M(d_i;\alpha)\right) \sim \chi^2(p), \tag{B.6}$$

where $M(d_i;\alpha)$ is a $(p,1)$ vector whose components are the orthonormal polynomials $M_j(d_i;\alpha)$, for $j = 1\ldots,p$, associated with a geometric distribution with a success probability α; and d_i denotes the duration between two consecutive violations (see Candelon et al., 2011, for more details). By a suitable choice of the number of moment conditions, p, one may elaborate separate tests for the UC and IND assumptions, similar to the LR test of Christoffersen (1998). The test statistic for the UC hypothesis is given by $J_{UC} = J_{CC}(1)$, while under IND hypothesis, the test statistic, J_{IND}, is derived by evaluating $J_{CC}(p)$ at the estimated (actual) violation rate $\hat{\alpha}$. By means of Monte Carlo simulations, Candelon et al. (2011) show that this new test has relatively high-power properties compared to earlier duration-based backtesting procedures (*e.g.*, Christoffersen and Pelletier, 2004) making it more appealing for practitioners.

4. Violation magnitude: Since VaR is non-informative about losses beyond its value, a validation technique allowing risk managers to assess the validity of a risk model by accounting, in addition to the violation rate, for the magnitude of extreme losses is a stringent requirement. In this line, Colletaz et al. (2013) propose a backtesting methodology based on both the number and the severity of VaR violations. Their approach exploits the concept of *super* violation, which is defined as a loss greater than VaR(α'), with α' much smaller than the coverage rate α. If the rate of *super* violations is abnormally high, this means that the magnitude of losses with respect to VaR(α) is too large. This test, called the Risk Map, is based on three indicator variables: $J_{1,t} = I_t(\alpha) - I_t(\alpha')$, which takes values of 1 if the loss falls between the estimated $\text{VaR}_t(\alpha')$ and $\text{VaR}_t(\alpha)$ and 0 otherwise; $J_{2,t} = I_t(\alpha')$ for which a value of 1 is assigned if the loss exceeds $\text{VaR}_t(\alpha')$ and 0 otherwise; and $J_{0,t} = 1 - J_{1,t} - J_{2,t}$. Under the joint null hypothesis that both the number of VaR violations and *super* violations are accurate, the corresponding test statistic is a multivariate unconditional coverage (MUC) LR test. That is:

$$LR_{MUC}(\alpha,\alpha') = -2\ln\left[(1-\alpha)^{N_0}(\alpha-\alpha')^{N_1}(\alpha')^{N_2}\right] + 2\ln\left[\left(\frac{N_0}{T}\right)^{N_0}\left(\frac{N_1}{T}\right)^{N_1}\left(\frac{N_2}{T}\right)^{N_2}\right] \sim \chi^2(2), \tag{B.7}$$

where $N_i = \sum_{i=1}^{T} J_{i,t}$, for $i = 0,1,2$, denotes the count variable associated with each indicator variable. Following Colletaz et al. (2013), we set $\alpha = 1\%$ and $\alpha' = 0.2\%$ for VaR and *super* VaR calculation, respectively.[16]

Appendix C. The three-zone approach of Basel II Accord for penalty structure

Zone	Number of violations	Penalty factor k
Green	0–4	0.00
Yellow	5	0.40
	6	0.50
	7	0.65
	8	0.75
	9	0.85
Red	10+	1.00

The number of violations is given for a forecasting period of 250 days.
Source: Basel Committee on Banking Supervision, http://www.bis.org/bcbs/.

References

Alexander, C., 2008. Market Risk Analysis. Quantitative Methods in Finance, vol. I. John Wiley & Sons, West Sussex, England.
Alexander, C., Lazar, E., 2006. Normal mixture GARCH(1,1): applications to exchange rate modelling. J. Appl. Econom. 21, 307–336.
Baillie, R., Bollerslev, T., Mikkelsen, H., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. J. Econom. 74 (1), 3–30.
Bao, Y., Lee, T.-H., Saltoglu, B., 2006. Evaluating predictive performance of value-at-risk models in emerging markets: a reality check. J. Forecasting 25 (2), 101–128.

[16] Note that any low threshold level $\alpha' \ll \alpha$ for *super* VaR calculation could be chosen. For instance, α' may be set such that $VaR(\alpha')$ corresponds either to the stressed VaR or to the expected shortfall (Colletaz et al., 2013).

Bao, Y., Lee, T.-H., Saltoglu, B., 2007. Comparing density forecast models. J. Forecasting 26, 203–225.
Barndorff-Nielsen, O.E., 1995. Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. Tech. Rep. 300. Department of Theoretical Statistics, University of Aarhus.
Barndorff-Nielsen, O.E., 1997. Normal inverse Gaussian distributions and stochastic volatility modelling. Scand. J. Stat. 24 (1), 1–13.
Basel Committee on Banking Supervision, 2006. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Tech. Rep. Bank for International Settlements, Basel, Switzerland. <http://www.bis.org/publ/bcbs128.htm>.
Basel Committee on Banking Supervision, 2009. Revisions to the Basel II Market Risk Framework: Final version, Tech. Rep. Bank for International Settlements, Basel, Switzerland. <http://www.bis.org/publ/bcbs158.htm>.
BenSaïda, A., 2015. The frequency of regime switching in financial market volatility. J. Empirical Finance 32, 63–79.
BenSaïda, A., Slim, S., 2016. Highly flexible distributions to fit multiple frequency financial returns. Physica A: Stat. Mech. Appl. 442, 203–213. http://dx.doi.org/10.1016/j.physa.2015.09.021.
Bianchi, M.L., Rachev, S.T., Kim, Y.S., Fabozzi, F.J., 2010. Tempered stable distributions and processes in finance: numerical analysis. In: Corazza, M., Pizzi, C. (Eds.), Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer Milan, New York, pp. 33–42.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticty. J. Econom. 31 (3), 307–327.
Bollerslev, T., Mikkelsen, H.O., 1996. Modelling and pricing long memory in stock market volatility. J. Econom. 73 (1), 151–184.
Bontemps, C., Meddahi, N., 2005. Testing normality: a GMM approach. J. Econom. 142 (1), 149–186.
Bontemps, C., Meddahi, N., 2012. Testing distributional assumptions: a GMM approach. J. Appl. Econom. 27 (6), 978–1012.
Boyarchenko, S.I., Levendorskii, S.Z., 2000. Option pricing for truncated Lévy processes. Int. J. Theor. Appl. Finance 3 (3), 549–552.
Broda, S.A., Paolella, M.S., 2009. CHICAGO: a fast and accurate method for portfolio risk calculation. J. Financ. Econom. 7 (4), 412–436.
Broda, S.A., Haas, M., Krause, J., Paolella, M.S., Steude, S.C., 2013. Stable mixture GARCH models. J. Econom. 172 (2), 292–306.
Campbell, S., 2007. A review of backtesting and backtesting procedures. J. Risk 9 (2), 1–18.
Candelon, B., Colletaz, G., Hurlin, C., Tokpavi, S., 2011. Backtesting value-at-risk: a GMM duration-based test. J. Financ. Econom. 9 (2), 314–343.
Carr, P., Geman, H., Madan, D.B., Yor, M., 2002. The fine structure of asset returns: an empirical investigation. J. Bus. 75 (2), 305–332.
Christoffersen, P., 1998. Evaluating interval forecasts. Int. Econ. Rev. 39 (4), 841–862.
Christoffersen, P.F., Pelletier, D., 2004. Backtesting value-at-risk: a duration-based approach. J. Financ. Econom. 2 (1), 84–108.
Colletaz, G., Hurlin, C., Pérignon, C., 2013. The risk map: a new tool for validating risk models. J. Banking Finance 37 (10), 3843–3854.
Conrad, C., Haag, B.R., 2006. Inequality constraints in the fractionally integrated GARCH model. J. Financ. Econom. 4 (3), 413–449.
Cont, R., 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quant. Finance 1 (2), 223–236.
Cont, R., Tankov, P., 2004. Financial Modelling with Jump Processes, Chapman and Hall/CRC Financial Mathematics Series. Chapman and Hall/CRC, London.
Covitz, D., Liang, N., Suarez, G.A., 2013. The evolution of a financial crisis: collapse of the asset-backed commercial paper market. J. Finance 68 (3), 815–848.
Davidson, J., 2004. Moment and memory properties of linear conditional heteroscedasticity models, and a new model. J. Bus. Econ. Stat. 22 (1), 16–29.
Del Brio, E.B., Mora-Valencia, A., Perote, J., 2014. VaR performance during the subprime and sovereign debt crises: an application to emerging markets. Emerg. Markets Rev. 20, 23–41.
Diamandis, P.F., Drakos, A.A., Kouretas, G.P., Zarangas, L., 2011. Value-at-risk for long short trading positions: evidence from developed emerging equity markets. Int. Rev. Financ. Anal. 20 (3), 165–176.
Diebold, F.X., Schuermann, T., Stroughair, J., 2000. Pitfalls and opportunities in the use of extreme value theory in risk management. J. Risk Finance 1 (2), 30–35.
Dimitrakopoulos, D.N., Kavussanos, M.G., Spyrou, S.I., 2010. Value at risk models for volatile emerging markets equity portfolios. Quart. Rev. Econ. Finance 50 (4), 515–526.
Ding, Z., Granger, C., Engle, R., 1993. A long memory property of stock market returns and a new model. J. Empirical Finance 1 (1), 83–106.
Engle, R., Manganelli, S., 2004. CAViaR: conditional autoregressive value-at-risk by regression quantiles. J. Bus. Econ. Stat. 22 (4), 367–381.
Engle, R., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. J. Finance 48 (5), 1749–1778.
Forsberg, L., Bollerslev, T., 2002. Bridging the gap between the distribution of realized (ECU) volatility and ARCH modeling (of the Euro): the GARCH-NIG model. J. Appl. Econom. 17 (5), 535–548.
Giot, P., Laurent, S., 2003. Value-at-risk for long and short trading positions. J. Appl. Econom. 18 (6), 641–664.
Giot, P., Laurent, S., 2004. Modelling daily value-at-risk using realized volatility and ARCH type models. J. Empirical Finance 11 (3), 379–398.
Glosten, L.R., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48 (5), 1779–1801.
Grigelionis, N., 2001. Generalized it z-distributions and related stochastic processes. Lithuanian Math. J. 41 (3), 239–251.
Grigoletto, M., Provasi, C., 2008. Simulation and estimation of the Meixner distribution. Commun. Stat. Simul. Comput. 38 (1), 58–77.
Haas, M., Paolella, M.S., 2012. Mixture and regime-switching GARCH models. In: Bauwens, L., Hafner, C.M., Laurent, S. (Eds.), Handbook of Volatility Models and Their Applications, vol. 3. John Wiley & Sons, Inc., Hoboken, New Jersey.
Haas, M., Mittnik, S., Paolella, M.S., 2004a. A new approach to markov switching GARCH models. J. Financ. Econom. 2 (4), 493–530.
Haas, M., Mittnik, S., Paolella, M.S., 2004b. Mixed normal conditional heteroskedasticity. J. Financ. Econom. 2 (2), 211–250.
Haas, M., Krause, J., Paolella, M.S., Steude, S.C., 2013. Time-varying mixture GARCH models and asymmetric volatility. North Am. J. Econ. Finance 26, 602–623.
Hansen, B.E., 1994. Autoregressive conditional density estimation. Int. Econ. Rev. 35 (3), 705–730.
Jensen, M.B., Lunde, A., 2001. The NIG-S & ARCH model: a fat tailed, stochastic, and autoregressive conditional heteroscedastic volatility model. Econom. J. 4 (2), 319–342.
Kim, Y.S., Rachev, S.T., Bianchi, M.L., Fabozzi, F.J., 2008. Financial market models with Levy processes and time-varying volatility. J. Banking Finance 32 (7), 1363–1378.
Kim, Y.S., Rachev, S.T., Bianchi, M.L., Mitov, I., Fabozzi, F.J., 2011. Time series analysis for financial market meltdowns. J. Banking Finance 35 (8), 1879–1891.
Koponen, I., 1995. Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. Phys. Rev. E 52 (1), 1197–1199.
Krause, J., Paolella, M.S., 2014. A fast, accurate method for value-at-risk and expected shortfall. Econometrics 2 (2), 98–122.
Kuester, K., Mittnik, S., Paolella, M.S., 2006. Value-at-risk prediction: a comparison of alternative strategies. J. Financ. Econom. 4 (1), 53–89.
Kupiec, P.H., 1995. Techniques for verifying the accuracy of risk measurement models. J. Derivatives 3 (2), 73–84.
Lambert, P., Laurent, S., 2001. Modelling Financial Time Series Using GARCH-type Models and a Skewed Student Density. Université de Liège, Mimeo.
Madan, D., Seneta, E., 1990. The variance gamma (V.G.) model for share market returns. J. Bus. 63 (4), 511–524.
Madan, D., Carr, P., Chang, E., 1998. The variance gamma process and option pricing. Eur. Finance Rev. 2 (1), 79–105.
McMillan, D., Kambouroudis, D., 2009. Are riskmetrics forecasts good enough? Evidence from 31 stock markets. Int. Rev. Financ. Anal. 18 (3), 117–124.
McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques, and Tools. Princeton University Press, Princeton.
Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. J. Financ. Econ. 3 (1–2), 125–144.
Mittnik, S., Paolella, M.S., 2000. Conditional density and value-at-risk prediction of asian currency exchange rates. J. Forecasting 19 (4), 313–333.
Mittnik, S., Paolella, M.S., 2003. Prediction of financial downside risk with heavy tailed conditional distributions. In: Rachev, S.T. (Ed.), Handbook of Heavy Tailed Distributions in Finance. Elsevier Science, Amsterdam.
Paolella, M.S., 1997. Tail Estimation and Conditional Modeling of Heteroscedastic Time-Series. PhD Thesis. Institute of Statistics and Econometrics, Christian Albrechts University at Kiel.
Paolella, M.S., Polak, P., 2015a. COMFORT: a common market factor non-Gaussian returns model. J. Econom. 187 (2), 593–605.

Paolella, M.S., Polak, P., 2015b. Density and Risk Prediction with Non-Gaussian COMFORT Models, Working Paper. Swiss Finance Institute.
Santos, A.A.P., Nogales, F.J., Ruiz, E., 2013. Comparing univariate and multivariate models to forecast portfolio value-at-risk. J. Financ. Econom. 11 (2), 400–441.
Scherer, M., Rachev, S.T., Kim, Y.S., Fabozzi, F.J., 2012. Approximation of skewed and leptokurtic return distributions. Appl. Financ. Econ. 22 (16), 1305–1316.
Schoutens, W., 2001. The Meixner Process in Finance. Tech. Rep. 2001-002, Eurandom, Leuven, Belgium.
Schoutens, W., 2003. Lévy Processes in Finance: Pricing Financial Derivatives. John Wiley & Sons, West Sussex, England.
So, M.K.P., Yu, P.L.H., 2006. Empirical analysis of GARCH models in value at risk estimation. J. Int. Financ. Markets Inst. Money 16 (2), 180–197.
Stephens, M.A., 1976. Asymptotic results for goodness-of-fit statistics with unknown parameters. Ann. Stat. 4 (2), 357–369.
Tang, T.L., Shieh, S.J., 2006. Long memory in stock index futures markets: a value-at-risk approach. Physica A 366, 437–448.
Venter, J.H., de Jongh, P.J., 2002. Risk estimation using the normal inverse Gaussian distribution. J. Risk 4 (2), 1–24.
Wilhelmsson, A., 2009. Value at risk with time varying variance, skewness and kurtosis – the NIG-ACD model. Econom. J. 12 (1), 82–104.
Wu, L., 2006. Dampened power law: reconciling the tail behavior of financial asset returns. J. Bus. 79 (3), 1445–1473.
Ziggel, D., Berens, T., Wei, G.N., Wied, D., 2014. A new set of improved value-at-risk backtests. J. Banking Finance 48 (2), 29–41.