

“Do coherent risk measures identify assets risk profiles similarly? Evidence from international futures markets”

AUTHORS	Sharif Mozumder M. Humayun Kabir Michael Dempsey  https://orcid.org/0000-0002-9059-0416
ARTICLE INFO	Sharif Mozumder, M. Humayun Kabir and Michael Dempsey (2017). Do coherent risk measures identify assets risk profiles similarly? Evidence from international futures markets. <i>Investment Management and Financial Innovations</i> , 14(3), 361-380. doi: 10.21511/imfi.14(3-2).2017.07
DOI	http://dx.doi.org/10.21511/imfi.14(3-2).2017.07
RELEASED ON	Tuesday, 28 November 2017
RECEIVED ON	Sunday, 20 August 2017
ACCEPTED ON	Wednesday, 04 October 2017
LICENSE	 This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
JOURNAL	"Investment Management and Financial Innovations"
ISSN PRINT	1810-4967
ISSN ONLINE	1812-9358
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”



NUMBER OF REFERENCES

38



NUMBER OF FIGURES

5



NUMBER OF TABLES

5

© The author(s) 2025. This publication is an open access article.



BUSINESS PERSPECTIVES



LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10, Sumy,
40022, Ukraine

www.businessperspectives.org

Received on: 20th of August, 2017

Accepted on: 4th of October, 2017

© Sharif Mozumder, M. Humayun Kabir, Michael Dempsey, 2017

Sharif Mozumder, Senior Lecturer,
Department of Mathematics,
University of Dhaka, Bangladesh.

M. Humayun Kabir, Senior Lecturer,
School of Economics and Finance,
Massey University, New Zealand.

Michael Dempsey, Professor, Faculty
of Finance and Banking, Ton Duc
Thang University, Ho Chi Minh City,
Vietnam.



This is an Open Access article, distributed under the terms of the [Creative Commons Attribution-NonCommercial 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/), which permits re-use, distribution, and reproduction, provided the materials aren't used for commercial purposes and the original work is properly cited.

Sharif Mozumder (Bangladesh), M. Humayun Kabir (New Zealand),
Michael Dempsey (Vietnam)

DO COHERENT RISK MEASURES IDENTIFY ASSETS RISK PROFILES SIMILARLY? EVIDENCE FROM INTERNATIONAL FUTURES MARKETS

Abstract

The authors consider Lévy processes with conditional distributions belonging to a generalized hyperbolic family and compare and contrast full density-based Lévy-expected shortfall (ES) risk measures and Lévy-spectral risk measures (SRM) with those of a traditional tail-based unconditional extreme value (EV) approach. Using the futures data of leading markets the authors find that ES and SRM often differ in recognizing the risk profiles of different assets. While EV (extreme value) is often found to be more consistent than Lévy models, Lévy measures often perform better than EV measures when compared with empirical values. This becomes increasingly apparent as investors become more risk averse.

Keywords

Lévy-Khintchine formula, extreme value, spectral risk measures, expected shortfall, generalized hyperbolic distributions

JEL Classification C52, G13

INTRODUCTION

In the last few decades, we have experienced an increased level of financialization and securitization, and convergence between banking, insurance and security markets. This trend has led to a significant increase in financial risk and unpredictability of extreme events resulting in large losses faced by individual and institutional investors. Under such circumstances, the determination of the level of risk, and the management of risk have become even more challenging. With the advent of Value-at-Risk (VaR), both academicians and practitioners have been trying to devise models to measure risk more effectively. The VaR as a popular risk measure has the advantage of simplicity, but comes with inherent weaknesses. Specifically, the VaR does not satisfy the subadditivity requirement, which is an important property for a coherence risk measure (Artzer et al., 1999). VaR fixes tail events corresponding to a given confidence level, and considers the conditional likelihood of tail events while ignoring the actual size of extreme catastrophic events. Thus, VaR gives a partial snapshot of potential losses and fails to take into account the actual size of extreme losses after the point of cut-off.

In order to overcome this weakness and to ensure the subadditivity (and hence coherence) requirement is met, an expected shortfall (ES) measure has been proposed. ES estimates the potential loss by averag-

ing all the possible losses in the tail of the distribution¹. However, the ES gives all tail losses an equal weight implying that the individual is risk neutral at the margin between better and worse tail outcomes (Grootveld & Hallerbach, 2004; Cotter & Dowd, 2006). On the other hand, the spectral risk measure (SRM) proposed by Acerbi (2002, 2004) is independent of any particular extreme event and requires that catastrophic tail events and usual non-tail events have different weights and that the weight of the catastrophic tail events are allowed to vary according to how averse an investor is towards the risk. In contrast, extreme value (EV) models use only the data remaining in the tail of the distribution after the cut-off point. The SRM provides investors with the flexibility to choose their individual degree of aversion to risk; but this flexibility comes at a greater computational cost². A number of studies examine the SRM in the context of portfolio optimization, and the relation between risk aversion and the SRM (Giannopoulos & Tunaru, 2005; Inui & Kijima, 2005; Adam et al., 2008; Chen & Wang, 2008; Deng et al., 2009; Wächter & Mazzoni, 2013). Sorwar and Dowd (2010) estimate ES and SRM for various option positions. Cotter and Dowd (2006) use tail density based extreme value ES and SRM risk measures and compare the precision of the estimates of these risk measures. They find the ES standard errors are higher than the VaR for S&P 500, FTSE100, DAX, and Hang Seng except Nikkei 225 futures contracts. However, the ES has higher coefficient of variations and narrower confidence intervals suggesting that they are more precisely estimated. On the other hand, the SRM has significantly wider confidence intervals than the VaR and ES.

In this paper, we focus on estimating coherent risk measures ES and SRM for both Lévy and EV models. A Lévy process is characterized by stationary independent increments and the distribution is infinitely divisible so as to represent skewness and excess kurtosis in the data. Moreover, Lévy models use the entire data or full density to estimate the model parameters in contrast to an EV model that uses only the tail density of the distribution.

Formally, $ES(\alpha)$ is restricted to the tail at the extreme end of the density distribution (with the confidence level, α value is as high as 0.95 or even 0.99): $ES(\alpha)$ is the value of the integral

$$\frac{1}{1-\alpha} \cdot \int_{\alpha}^1 VaR(u) \cdot du.$$

However, SRM with an exponential risk aversion function is an integral of the form

$$\int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}} VaR(u) \cdot du$$

(where R is the Arrow-Pratt coefficient of absolute risk aversion), which is not only restricted to such a tail, but also embraces the data outside the tail. Cotter and Dowd (2006) focus only on EV-ES and EV-SRM to determine the clearing house's margin requirements.

We investigate the performance of full density based Lévy-SRM and Lévy-ES risk measures, and compare the results to the tail density-based EV-SRM and EV-ES. To the best of our knowledge, this is the first paper to apply full density-based Lévy-SRM and Lévy-ES risk measures to international futures markets.

We discuss the computational challenges that arise in the implementation of Lévy models for estimating the SRM. We then conduct detailed empirical analysis with international futures markets³ to determine

1 Recently, Acharya et al. (2010, 2012) have introduced marginal expected shortfall (MES) as a measure of losses faced by a firm in the tail of the aggregate sector's loss distribution, as well as a systemic expected shortfall (SES), which increases with the firm's leverage and with its expected loss in the tail of the system's loss distribution. For more detail on the estimation, see Brownlees and Engle (2012).

2 The computational issues are discussed by Kevin and Cotter (2006) in the context of an extreme value (EV) approach. Specifically, the authors evaluate the integrals associated with the calculation of VaR, ES and SRM.

3 We consider futures data on the S&P 500, FTSE100, DAX, Hang Seng and Nikkei 225 indexes. The discussion on data is presented in section 2.

whether coverage based coherent risk measure ES and risk aversion based coherent risk measure SRM provide similar risk scenarios.

The Lévy approach, though mathematically elegant, comes with a major drawback, in that with few exceptions, there are no closed form formulae for risk measures. As such, even a relatively straightforward VaR estimation is difficult to implement. As the risk measures ES and SRM are compounded versions of VaR, their implementation is even more difficult. Our approach follows a procedure of fixing the tail as applied in EV calibration followed by calculation of ES and SRM. We then use the Lévy models from generalized hyperbolic class with calibration based on the entire data and calculate ES and SRM for these models.

The paper is structured as follows. Section 1 briefly describes the Lévy and EV frameworks. Section 2 provides the initial data analysis. In section 3 we discuss conceptual matters regarding estimation and bootstrapping of ES and SRM. In Section 4, we discuss goodness of fit under Lévy and EV models. Section 5 presents the analysis of estimates of ES and Lévy-SRM models. Section 6 describes the empirical findings. Last section concludes the paper.

1. CHARACTERIZATION IN LÉVY FRAMEWORK

The characteristic function of a stochastically continuous process starting at zero and with stationary independent increments can be written as

$$\begin{aligned} \Phi_{X_t}(s) &= E \cdot [e^{isX_t}] = \\ &= \exp \left\{ t \cdot \left[ias - \frac{s^2 \cdot b^2}{2} \right] + \int_{\mathfrak{R} \setminus \{0\}} [e^{isx} - 1 - isxI_{\{-1,1\}}(x)] \nu(dx) \right\}, \end{aligned} \tag{1}$$

for $s \in \mathfrak{R}$, $t \geq 0$ and with constants $a \in \mathfrak{R}$, $b \in \mathfrak{R}^+$, where ν is a measure defined on $\mathfrak{R} \setminus \{0\}$ that satisfies:

$$\int_{\mathfrak{R} \setminus \{0\}} (|x|^2 \wedge 1) \cdot \nu(dx) < \infty. \tag{2}$$

Equation (1) is the Lévy-Khinchine representation of a Lévy process⁴.

Given the transition density of a process on $[t_1; t_2]$, the characteristic function (1) of the conditional distribution of the process at t_2 , given the information available at t_1 , can be obtained by a Fourier transform. However, the transition density itself is sufficient for the estimation of the model and calculation of the risk measures. Availability

of closed form transition density ensures that the underlying distribution is closed under convolution. However, in general, the infinite divisibility of Lévy processes can be used to obtain the conditional characteristics of the process on an interval of length t :

$$\Phi_{X_t}(s) = [\Phi_{X_1}(s)]^t. \tag{3}$$

Thus, an inverse Fourier transform on any time scale can be used to numerically obtain the transition density from the characteristic function (1) with the Lévy measure of the process. The numerical transition densities can then be used to estimate the risk measure under different model assumptions.

In this paper, our interest is limited to those members of the generalized hyperbolic (GH) family of Lévy processes that have been widely used in financial modeling. The Lévy process has extensively been used in option analysis (German, 2002; Fajardo, 2015; Fajardo & Mordecki, 2006, 2014; Fuse & Meucci, 2008; Wong & Guan, 2011). Kim et al. (2008) use exponential Lévy market models and construct a new GARCH model that incorporates volatility clustering, fat tails, and skewness to estimate the parameters for the S&P 500 index and examine the out-of-sample forecasting performance for the GARCH models for the S&P 500 option prices. A GH family of Lévy processes was originally introduced to model grain-size distribution of wind-blown sands (Barndorff, 1977).

4 The theory of Lévy processes can be found in Bertoin (1996), Sato (1999), and Kyprianou (2006), amongst others.

Later Eberlein and Prause (2002) and Prause (1999) studied the whole family of GH distributions as a tool to model log-returns of financial assets. Some of its subclasses were separately studied in a financial context. Eberlein and Keller (1995), Bingham and Kiesel (2001) studied the hyperbolic distribution (HYP) and Barndorff (1995) applied the normal inverse Gaussian (NIG) to financial data. Eberlein and Hammerstein (2002) provide a complete and useful overview of limiting cases for this rich family of processes. We focus on a subclass of Lévy processes – variance gamma (VG), normal inverse Gaussian (NIG), hyperbolic distribution (HYP) and generalized hyperbolic (GH)⁵. Restricting ourselves to the subclasses of variance gamma (VG), normal inverse Gaussian (NIG), hyperbolic (HYP) and generalized hyperbolic (GH) allows us to obtain either the transition densities across time for processes closed under convolution, or at least the densities at time $t = 1$, for those which are not closed under convolution. Furthermore, in our empirical section, authors use daily return data for the indices under consideration and maintain a time scale in days, so that $t = 1$ in equation (3), which ensures that we are not required to use any inversion to obtain the transition densities numerically even when the underlying distribution is not closed under convolution.

Let, $X_1 = \log(S_{t+1} / S_t)$ for any non-negative integer t and characterized by the Lévy-Khintchine formula in equation (1). For our models, the equivalent processes are given more effectively by their densities.

For variance gamma:

$$f_{X_1}^{vg}(x) = \frac{2e^{\frac{\theta x}{\sigma^2}}}{\sigma\sqrt{2\pi\nu}^{\frac{1}{\nu}}} \cdot \left(\frac{x}{\sqrt{\frac{2\sigma^2}{\nu} + \theta^2}} \right)^{\frac{1}{\nu} - \frac{1}{2}} \times \times K_{\frac{1}{\nu} - \frac{1}{2}} \left(\frac{x \cdot \sqrt{\frac{2\sigma^2}{\nu} + \theta^2}}{\sigma^2} \right) \quad (4)$$

Here K_I is the modified Bessel function of the third kind with index I .

For normal inverse Gaussian:

$$f_{X_1}^{nig}(x) = \frac{\alpha \cdot \delta}{\pi} \cdot e^{\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta x\right)} \times \times \frac{K_1 \cdot \left(\alpha\sqrt{\delta^2 + x^2}\right)}{\sqrt{\delta^2 + x^2}} \quad (5)$$

For hyperbolic:

$$f_{X_1}^{hyp}(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2 \cdot \delta \alpha K_1 \left(\delta\sqrt{\alpha^2 - \beta^2}\right)} \cdot e^{\left(-\alpha\sqrt{\delta^2 + x^2} + \beta x\right)} \quad (6)$$

and for generalized hyperbolic (GH):

$$f_{X_1}^{gh}(x) = \frac{\left(\alpha^2 - \beta^2\right)^{\frac{\nu}{2}}}{\sqrt{2\pi}\alpha^{\nu - \frac{1}{2}} \cdot \delta^\nu \cdot K_\nu \left(\delta\sqrt{\alpha^2 - \beta^2}\right)} \times \times \left(\delta^2 + x^2\right)^{\frac{\nu - 1}{2}} \cdot K_{\nu - \frac{1}{2}} \cdot \left(\alpha\sqrt{\beta^2 + x^2}\right) \cdot e^{\beta x} \quad (7)$$

The availability of closed form densities makes it easier to obtain the standard errors of each parameter through Fisher's information matrix.

The competing approach to Lévy idea for this paper is the extreme value (EV) model, which incorporates only extreme returns in calibration. As explained in Dowd (2005), and subsequently as applied in Cotter and Dowd (2006), perhaps the most elegant approach to such objectives is to utilize the peaks-over-threshold (POT). The essence of the POT approach lies in the fact that as the threshold u becomes larger, the distribution of exceedances converge to a two parameter generalized Pareto (GP) distribution:

$$GP_{\xi, \beta}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} & \text{if } \xi \geq 0, \\ 1 - \exp\left(-\frac{x}{\beta}\right) & \text{if } \xi < 0, \end{cases} \quad (8)$$

5 Variance gamma (VG), normal inverse Gaussian (NIG) and hyperbolic distribution (HYP) are restricted versions of a GH model. So, though GH contains the other models, authors are interested to see the effect of full flexibility and selected restrictions for this family of processes. For this reason, we keep the GH model itself alongside its restricted versions.

where

$$x \in \begin{cases} [0, \infty) & \text{if } \xi \geq 0, \\ \left[0, \frac{-\beta}{\xi}\right] & \text{if } \xi < 0. \end{cases} \quad (9)$$

The parameters ξ and $\beta > 0$ are, respectively, shape and scale parameters, contingent upon the threshold u .

2. INITIAL DATA ANALYSIS

Our analysis is based on future contracts return data. More specifically, we study the returns based on the heavily traded S&P 500, FTSE100, DAX, Hang Seng and Nikkei 225 indices for the period from January 1, 1991 to December 31, 2003 collected from Datastream. The data refer to futures contracts that expire in the following trading months and rollover from one expiring contract to the next one at the start of each trading month. When dealing with bank holidays, Data stream considers padding the dataset and takes the bank holidays end-of-day price to be the previous trading day's end-of-day price. Thus, we have the same number of daily returns for all contracts (3,392). Our selection of data and sample period is intentional so that we can compare the results with Cotter and Dowd (2006) who use same data and sample period for EV-ES and EV-SRM models.

Figures 1 to 5 (see Appendix) show the extreme patterns in each futures return. The cut-off point

in each extreme tail plot is selected according to the extreme value theory and is discussed in Cotter and Dowd (2006). Clearly, the extents of extremity in return corresponding to various indexes are different and the visual goodness-of-fit of various models, both tail-based and full density-based, are also clearly distinct.

In Table 1, we replicate the unconditional maximum likelihood estimates of futures indexes of the generalized Pareto (GP) distribution of Cotter and Dowd (2006), which provides a good fit to the data both for long and short positions. The tail indices are positive except for the Nikkei 225 and the estimated scale parameters fluctuate around 1. Table 1 also provides assumed thresholds u , the associated number of exceedances (N_u) and the observed exceedance probabilities (Prob). While the numbers and probabilities of exceedances change, the assumed thresholds are in the stable tail-index regions based on the tail index plots.

The tail-based calibration provides significantly different estimates for long and short positions for the extreme value GP model, as it makes use of only tail observations. For tail based EV when the left skewed density become right skewed, the tail observations for long and short positions could be significantly different in numbers and hence significantly affect the estimates. As a result, we find tail asymmetry of long and short positions under an extreme value model in Table 1. Evidently, the same cut-off

Table 1. Generalized Pareto distribution (GPD) parameter estimates for futures indexes

Index	Position	Threshold (u)	Prob	N_u	Scale (β)	Tail (ξ)
S&P 500	Long	2	0.040	130	0.604 (0.079)	0.182 (0.099)
	Short	2	0.035	118	0.759 (0.130)	0.127 (0.146)
FTSE100	Long	1.5	0.077	250	0.707 (0.074)	0.097 (0.084)
	Short	1.5	0.085	276	0.727 (0.065)	0.022 (0.067)
DAX	Long	2	0.072	235	1.190 (0.099)	0.012 (0.052)
	Short	2	0.072	237	1.000 (0.097)	0.048 (0.072)
Hang Seng	Long	2	0.111	353	1.184 (0.096)	0.127 (0.062)
	Short	2	0.116	367	1.148 (0.086)	0.143 (0.055)
Nikkei 225	Long	2	0.088	277	0.891 (0.074)	-0.012 (0.058)
	Short	2	0.081	255	1.045 (0.085)	-0.068 (0.052)

Notes: Unconditional maximum likelihood estimates of the GPD parameters for long and short futures positions are based on daily % returns. Sample size $n = 3392$, threshold is u , probability of an observation in excess of u is Prob, number of exceedances in excess of u is N_u , estimated tale parameter is ξ and estimated scale parameter is β . Estimated standard errors of the parameters are in brackets. Thresholds are chosen as the approximate points where QQ plots change slope.

Table 2. Conditional maximum likelihood estimates for futures indexes

Panel A: variance gamma (VA)						
Index	u	Prob	σ	θ	ν	
S&P 500	2	0.03	1.09 (0.02)	-0.04 (0.02)	0.83 (0.06)	
FTSE100	1.5	0.08	1.16 (0.02)	-0.02 (0.02)	0.47 (0.05)	
DAX	2	0.08	1.51 (0.03)	-0.03 (0.03)	0.74 (0.05)	
Hang Seng	2	0.11	1.91 (0.03)	-0.05 (0.03)	0.81(0.05)	
Nikkei 225	2	0.09	1.53 (0.02)	0.03 (0.03)	0.40 (0.05)	
Panel B: normal inverse Gaussian (NIG)						
	u	Prob	A	β	δ	
S&P 500	2	0.03	0.74 (0.05)	-0.03 (0.02)	0.93 (0.04)	
FTSE100	1.5	0.08	1.04 (0.08)	-0.02 (0.02)	1.43 (0.09)	
DAX	2	0.07	0.59 (0.04)	-0.01 (0.01)	1.37 (0.07)	
Hang Seng	2	0.11	0.43 (0.03)	-0.01 (0.01)	1.62 (0.08)	
Nikkei 225	2	0.09	0.91 (0.08)	0.01 (0.01)	2.14 (0.16)	
Panel C: hyperbolic (HYP)						
	u	Prob	A	β	δ	
S&P 500	2	0.03	1.34 (0.04)	-0.03 (0.02)	0.24 (0.06)	
FTSE100	1.5	0.08	1.48 (0.06)	-0.02 (0.02)	0.88 (0.12)	
DAX	2	0.08	1.00 (0.03)	-0.01 (0.01)	0.49 (0.10)	
Hang Seng	2	0.11	0.47 (0.11)	-0.01 (0.01)	0.47 (0.11)	
Nikkei 225	2	0.09	1.19 (0.06)	0.01 (0.01)	1.41 (0.20)	
Panel C: generalized hyperbolic (GH)						
	u	Prob	A	β	δ	μ
S&P 500	2	0.03	0.937 (0.18)	-0.03 (0.02)	0.71 (0.20)	0.01 (0.46)
FTSE100	1.5	0.08	0.539 (0.36)	-0.02 (0.02)	1.92 (0.30)	-1.96 (0.85)
DAX	2	0.07	0.629 (0.16)	-0.01 (0.01)	1.29 (0.31)	-0.35 (0.55)
Hang Seng	2	0.11	0.347 (0.13)	-0.01 (0.01)	1.87 (0.39)	-0.85 (0.54)
Nikkei 225	2	0.09	1.031 (0.32)	0.01 (0.01)	1.83 (0.81)	0.15 (1.65)

Notes: Estimates of the VG, NIG, HYP and GH parameters for long futures positions are based on daily % returns. Sample size $n = 3392$, threshold is u , probability of an observation in excess of u is Prob (see Schoutens, 2003). Estimated standard errors of the parameters are in brackets. Thresholds are chosen as the approximate points where QQ plots change slope.

point maintains a different number of left-alone observations on the tails for long and short positions, and this causes the parameter estimates to differ for different positions under EV models.

In contrast, Lévy based calibration makes use of the complete data of short and long positions causing the densities to be reflected along the y -axis, and long and short positions just alter the sign of the parameter characterizing the skewness of the model. Thus, while a particular model, for example, VG with a long position gives a left skewed density, the shape of the density remains the same but becomes right skewed for a short position. Long and short positions thereby correspond to a sign change of the skewness characterizing parameter. Only the return vector, which is used in estimation for a long position gets multiplied by (-1) before being used in estimation

corresponding to a short position. The net effect is that the estimated density gets reflected along the y -axis. As a result, only the skewness controlling parameter has the sign change corresponding to long and short positions⁶. Table 2 presents the conditional maximum likelihood estimate of parameters for all five indexes for the four separate Lévy models. We report only the estimates of long positions for the brevity in the table. The skewness parameter is θ for VG, β for NIG, HYP and GH models. For short positions, this parameter is negative.

While the tail masses for observations in excess of thresholds are observed to be different for EV and Lévy models, different Lévy models that correspond to the same threshold, the tail masses exhibit further differences. These differences possibly explain how different Lévy models feed informa-

⁶ In relation to all the models, it can be said that under turbulent market conditions, investment with a short position is riskier than investment with a long position.

tion from observations outside the tails in fitting the tails. As a result, the corresponding quantiles of extreme-value and Lévy models do not lie along a vertical line. To illustrate the differences between the two approaches, we use the same number of tail observations and compare the QQ plot of EV with each of the Lévy models separately as illustrated in Figures 1 to 5.

At the very extreme tail, there is a clear evidence of deviation between EV and Lévy quantiles, and this deviation is smaller for EV in most cases. Specifically, EV provides a better fit for S&P 500, FTSE100, DAX, Hang Seng. Furthermore, for these indices, we see that NIG and GH provide a better fit on the tail than is the case for VG and HYP models. However, for the Nikkei 225, we observe the opposite feature. A close look reveals that while S&P 500, FTSE100, DAX, and Hang Seng indices show a greater fall in price than for a rise, Nikkei 225 shows a greater rise in price than a fall during the sample period. To visualize tail fits of the models, we present the generalized Pareto EV tail with each of our considered Lévy models separately.

Figures 1 to 5 show the tails for S&P 500, FTSE100, DAX, Hang Seng and Nikkei 225, respectively. We obtain the EV quantiles in excess of thresholds and then obtain the corresponding quantiles from the Lévy models. In other words, we do not fix the tail mass, but instead fix the thresholds. The consequence is that some of the Lévy quantiles closed to EV thresholds are, in fact, somewhat less than the thresholds in magnitude. This means that the Lévy tails are slightly fatter than the EV tails. This, in turn, explains the difference in tail masses covered by EV and Lévy models as reported in Tables 1 and 2.

3. ESTIMATION OF RISK MEASURES: METHODOLOGY AND PERFORMANCE

Apart from few specific cases, VaR in general is obtained as the solution of the quantile-integral equation:

$$\int_{x_{\min}}^{VaR} f(u) \cdot du - \alpha = 0, \tag{10}$$

where α is the coverage level.

As in the GP model, the significantly high α^{th} quantile, which is also VaR at a high confidence level α , is given by:

$$VaR(\alpha) = u + \frac{\beta}{\xi} \left\{ \left(\frac{n}{N_u} \cdot \alpha \right)^{-\xi} - 1 \right\} \tag{11}$$

and the expected shortfall (ES) with a coverage to the level of α is:

$$ES^{gp}(\alpha) = \frac{VaR(\alpha)}{1 - \xi} + \frac{\beta - \xi u}{1 - \xi}. \tag{12}$$

In equation (11), n is the total number of observations and N_u is the number of observations which exceeds the threshold u . Expected shortfall (ES) is estimated using the following equation:

$$ES(\alpha) = \frac{1}{1 - \alpha} \int_{\alpha}^1 VaR(u) \cdot du \tag{13}$$

$$ES(\alpha) = \frac{1}{1 - \alpha} \int_{\alpha}^1 \left\{ S \left| \int_{x_{\min}}^s f(x) \cdot dx = u \right. \right\} \cdot du. \tag{14}$$

Variance gamma ES model can then be obtained from the equation:

$$ES^{vg}(\alpha) = \frac{1}{1 - \alpha} \left\{ \int_{\alpha}^1 S \left[\int_{x_{\min}}^s \frac{2e^{\frac{\theta x}{\sigma^2}} \left(\frac{x}{\sqrt{\frac{2\sigma^2}{v} + \theta^2}} \right)^{\frac{1}{v} - \frac{1}{2}}}{\sigma \sqrt{2\pi v}^{\frac{1}{v}}} \cdot K_{\frac{1}{v} - \frac{1}{2}} \left(\frac{x \sqrt{\frac{2\sigma^2}{v} + \theta^2}}{\sigma^2} \right)} \cdot dx = u \right] \cdot du \right\}. \tag{15}$$

The ES formula for other Lévy models can also be obtained by considering other densities from section 1.

Spectral risk measures, however, do not rely on any particular confidence level. Instead, given a parameter characterizing the degree of investors' risk aversion, they consider the entire spectrum of losses. For our benchmark EV model, the closed form VaR formula provides a relatively simpler expression for SRM:

$$M_{\phi}^{sp}(R) = \int_0^1 \frac{Re^{-R(1-\alpha)}}{1-e^{-R}} \times \left[u + \frac{\beta}{\xi} \left\{ \left(\frac{n}{N_u} \cdot \alpha \right)^{-\xi} - 1 \right\} \right] \cdot dx \quad (16)$$

In the case of Lévy models, however, computation of SRM is very time consuming in regard to the closed form VaR measure:

$$M_{\phi}(R) = \int_0^1 \frac{Re^{-R(1-\alpha)}}{1-e^{-R}} \cdot VaR(u) \cdot du \quad (17)$$

The variance gamma SRM model can, then, be obtained from the equation:

The SRM estimates of other Lévy models can be obtained by considering respective densities in equation (18). The ϕ symbolizes that the SRM is calculated using the exponential risk aversion function

$$\phi(R) = \frac{Re^{R(1-p)}}{1-e^{-R}}$$

$$M_{\phi}^{vg}(R) = \int_0^1 \frac{Re^{-R(1-u)}}{1-e^{-R}} \left\{ \int_{\alpha}^1 S \int_{x_{\min}}^S \frac{2e^{\frac{\theta x}{\sigma^2}} \left(\frac{x}{\sqrt{\frac{2\sigma^2}{v} + \theta^2}} \right)^{\frac{1}{v} \frac{1}{2}}}{\sigma \sqrt{2\pi v}^{\frac{1}{v}}} \cdot K_{\frac{1}{v} \frac{1}{2}} \left(\frac{x \sqrt{\frac{2\sigma^2}{v} + \theta^2}}{\sigma^2} \right) \cdot dx = u \right\} \cdot du \quad (18)$$

The parametric bootstrap is applied to obtain the standard errors (SE) and confidence intervals (CI) of each risk measure. However, as we are dealing with Lévy models, which have no closed form expressions for risk measures, it is infeasible to use bootstraps with a large number of resampling. For each resample we draw the same number of uniform (0,1) random numbers as sample size and after sorting them in ascending order, we find the relevant quantile corresponding to the coverage level. This quantile is then used as the bootstrap coverage level, corresponding to which we obtain the VaR and ES as given by equations (10) and (14). Since, for a given bootstrap coverage level, this VaR equation needs to be resolved numerically, the corresponding ES equation takes a long time to find a converging value. This is because any numerical scheme applied to obtain the ES searches the converging value by evaluating the integrand 'vector byvector' and for each element of a vector the VaR needs to be obtained as a solution of the quantile integral equation (10). Bootstrapped VaR and ES vary because of the variation in the bootstrapped confidence level. Thus, since SRM does not depend on any particular confidence level, in order to obtain bootstrapped estimates of SRM, we need to randomize the whole spectrum. We thus need to approximate the integral in (20) by slicing the spectrum (due to the fact that randomization of the integral often fails to converge under numerical schemes, such as the one used by the MATLAB function 'quadl'). Thus, authors restrict ourselves to bootstrapping with 100 resampling⁷.

7 Even with 100 resampling, we find that a machine sophisticated configuration takes several hours to provide the SE and CI for ES from a Lévy model, corresponding to a confidence level. The same is true for SRM with each particular choice of risk aversion parameter. However, as Cotter (2006) reports, SE and CI with 5000 resampling for an EV model – which has closed form expressions both for VaR and ES and where closed form VaR allows SRM to be calculated equally well in seconds – determines that the difference between 100 and 5000 resampling is not significant between VaR and ES. However, in case of SRM, the difference is enormous. This is because, in addition to considering a small number of resampling, we evaluate the integral in SRM by considering only 100 slices. This makes the estimation performance of SRM comparable only between Lévy models.

Table 3. Anderson-Darling and left-truncated Anderson-Darling goodness-of-fit tests

Index	Model	AD-stat	1% CV	5% CV	10% CV	p-value
S&P 500	EV	0.25	2.95*	1.87*	1.49*	0.98
	VG	1.65	2.89*	1.73*	1.22	0.06
	NIG	1.50	3.00*	1.69*	1.18	0.06
	HYP	1.53	2.90*	1.72*	1.21	0.07
	GH	1.46	2.99*	1.71*	1.19	0.07
FTSE100	EV	0.74	3.10*	1.81*	1.32*	0.78
	VG	0.65	2.97*	1.88*	1.37*	0.25
	NIG	0.35	3.03*	1.86*	1.35*	0.35
	HYP	0.47	2.99*	1.87*	1.37*	0.31
	GH	0.24	3.03*	1.84*	1.35*	0.40
DAC	EV	1.01	4.38*	2.44*	1.62*	0.65
	VG	1.41	3.97*	2.41*	1.72*	0.14
	NIG	1.09	4.12*	2.36*	1.67*	0.18
	HYP	1.24	4.00*	2.40*	1.71*	0.16
	GH	1.09	4.12*	2.37*	1.67*	0.18
Hang Seng	EV	1.32	5.30*	2.97*	2.10*	0.68
	VG	0.71	5.06*	3.04*	2.15*	0.30
	NIG	0.50	5.29*	2.98*	2.08*	0.35
	HYP	0.57	5.08*	3.03*	2.14*	0.33
	GH	0.54	5.29*	2.95*	2.07*	0.34
Nikkei 225	EV	0.80	3.89*	2.47*	1.86*	0.73
	VG	0.72	3.99*	2.54*	1.88*	0.30
	NIG	0.63	4.00*	2.52*	1.86*	0.32
	HYP	0.65	3.98*	2.53*	1.87*	0.32
	GH	0.64	3.99*	2.52*	1.87*	0.32

Note: Anderson-Darling and left-truncated Anderson-Darling tests for Lévy and EV models, respectively. The *p*-value for left-truncated Anderson-Darling test is obtained by bootstrapping with 1000 resampling. (*) implies that the model survives the test to the corresponding significance level.

4. GOODNESS-OF-FIT TESTS

The Anderson-Darling (AD) test is particularly suitable in assessing the performance of tail based risk management models for goodness-of-fit⁸. Anderson and Darling (1952, 1954) proposed a weighing rule in a distance based on Kolmogorov-Smirnov test, which puts more emphasis on the tail observations. Anna, Rachev, and Fabozzi (2005) provide a formula for an AD test statistic when observations are only available at the extreme tail and the distribution of the complete sample is unknown. It is referred to as the Anderson-Darling test for left-truncated data. This scenario applies precisely to an EV model. For Lévy models that require complete distributions with closed form densities, the *p*-values of AD-test can be obtained analytically either by Monte Carlo simulation or bootstrapping. We carry out bootstrapping with 1,000 resampling to obtain the *p*-value. The test statistics are denoted as AD and AD_{ev}^2 for Lévy and EV models, respectively:

$$AD^2 = -N - \sum_{i=1}^N \frac{2i-1}{N} \times \left[\log(F(x_i)) + \log(1-F(x_{N+1-i})) \right], \tag{19}$$

$$AD_{ev}^2 = -n + 2n \cdot \log(1-F(u)) - \frac{1}{n} \sum_{i=1}^n (1+2(n-i)) \cdot \log(1-F(x_j)) + \frac{1}{n} \sum_{i=1}^n (1-2i) \cdot \log(F(x_j)-F(u)), \tag{20}$$

where *u* is the level of truncation and x_j is the j^{th} observed value of the order statistic $X_1 \leq X_2 \leq \dots \leq X_n$ and *n* is the total number of observations available on the tail.

We report the goodness-of-fit statistics in Table 3 using an AD and AD_{ev} test statistic. Though an AD test utilizes the entire dataset, the test puts more weight on tail observations. Thus, both AD and AD_{ev} are more informative regarding tail fit while hardly providing any information regarding the density fit far from the tail. Table 3 shows that both EV and full density Lévy models perform well on the tail, with EV appearing more reliable when tail fit alone is concerned. However, the fact that AD and AD_{ev} provide hardly any information regarding the fit far from the tails explains why models with better tail-fits fail to give any information regarding the quantification of risk measures that use quantiles far from tails for SRM models. This is demonstrated by the results in Table 4.

8 Other GOF tests such as a Chi-square test are not comparable for Lévy models on complete data or on an EV model on left truncated incomplete data.

Table 4. Estimates of ES and Lévy spectral risk measures for futures position

Index	Risk measure	ES			SRM		
		$\alpha = 0.99$ ($R = 20$)	$\alpha = 0.995$ ($R = 100$)	$\alpha = 0.999$ ($R = 200$)	$\alpha = 0.99$ ($R = 20$)	$\alpha = 0.995$ ($R = 100$)	$\alpha = 0.999$ ($R = 200$)
S&P 500	Empirical	3.902	4.613	7.246	2.265	3.685	4.476
	GP	3.862 (0.162) [0.935 1.070]	4.559 (0.237) [0.905 1.093]	6.558 (0.802) [0.804 1.175]	2.251 (0.153) [0.898 1.097]	3.288 (0.508) [0.781 1.259]	3.703 (0.773) [0.690 1.292]
	VG	3.604(0.121) [0.948 1.052]	4.099(0.160) [0.943 1.073]	5.243(0.313) [0.904 1.111]	2.274(0.952) [0.365 2.007]	3.484(3.256) [0.063 2.819]	4.001(5.080) [0.005 3.333]
	NIG	3.920(0.155) [0.939 1.061]	4.569(0.214) [0.932 1.087]	6.154(0.445) [0.885 1.136]	2.353(1.056) [0.409 1.802]	3.818(3.699) [0.057 2.832]	4.500(5.899) [0.004 3.590]
	HYP	3.634(0.122) [0.954 1.064]	4.141(0.164) [0.942 1.073]	5.318(0.321) [0.903 1.113]	2.282(0.973) [0.431 1.731]	3.518(3.297) [0.005 3.358]	4.049(5.155) [0.005 3.358]
	GH	3.843(0.142) [0.949 1.071]	4.440(0.195) [0.936 1.082]	5.868(0.395) [0.892 1.126]	2.338(1.031) [0.416 1.764]	3.729(3.567) [0.059 2.831]	4.358(5.644) [0.004 3.525]
FTSE 100	Empirical	4.018	4.532	5.590	2.338	3.739	4.377
	GP	4.009(0.163) [0.940 1.071]	4.689(0.259) [0.929 1.089]	6.459(0.666) [0.849 1.176]	2.234(0.147) [0.889 1.113]	3.415(0.551) [0.766 1.232]	3.848(0.808) [0.699 1.323]
	VG	3.617(0.109) [0.953 1.052]	4.062(0.143) [0.925 1.053]	5.078(0.326) [0.904 1.102]	2.319(0.803) [0.529 1.685]	3.423(3.147) [0.067 2.738]	3.881(4.864) [0.005 3.217]
	NIG	3.782(0.127) [0.948 1.058]	4.309(0.172) [0.916 1.060]	5.565(0.414) [0.889 1.118]	2.352(0.832) [0.521 1.701]	3.574(3.349) [0.063 2.824]	4.112(5.245) [0.005 3.403]
	HYP	3.679(0.115) [0.951 1.054]	4.153(0.153) [0.922 1.056]	5.249(0.359) [0.893 1.119]	2.331(0.814) [0.525 1.694]	3.478(2.924) [0.052 2.567]	3.964(4.998) [0.005 3.279]
	GH	3.865(0.140) [0.944 1.062]	4.462(0.197) [0.908 1.067]	5.966(0.523) [0.866 1.156]	2.363(0.846) [0.518 1.708]	3.658(3.217) [0.049 2.616]	4.262(5.539) [0.004 3.597]
DAX	Empirical	5.511	6.330	9.738	3.185	5.219	6.276
	GP	5.563(0.212) [0.939 1.063]	6.421(0.276) [0.919 1.076]	8.441(0.691) [0.858 1.116]	3.035(0.218) [0.889 1.103]	4.761(0.684) [0.787 1.249]	5.336(1.036) [0.692 1.255]
	VG	4.922(0.160) [0.949 1.056]	5.579(0.212) [0.919 1.057]	7.096(0.411) [0.907 1.108]	3.072(1.292) [0.438 1.723]	4.668(3.866) [0.104 2.434]	5.345(5.831) [0.012 3.374]
	NIG	5.328(0.204) [0.941 1.066]	6.179(0.279) [0.905 1.069]	8.248(0.581) [0.888 1.132]	3.162(1.398) [0.417 1.776]	5.055(4.388) [0.096 2.481]	5.925(6.665) [0.011 3.447]
	HYP	4.994(0.167) [0.948 1.058]	5.680(0.222) [0.917 1.059]	7.273(0.433) [0.904 1.111]	3.088(1.310) [0.434 1.727]	4.733(3.947) [0.103 2.441]	5.440(5.962) [0.012 3.388]
	GH	5.300(0.199) [0.942 1.065]	6.133(0.272) [0.907 1.067]	8.141(0.560) [0.890 1.129]	3.158(1.390) [0.418 1.766]	5.027(4.341) [0.097 2.476]	5.879(6.591) [0.011 3.441]

Table 4 (cont.). Estimates of ES and Lévy spectral risk measures for futures position

Index	Risk measure	ES			SRM		
		$\alpha = 0.99$ ($R = 20$)	$\alpha = 0.995$ ($R = 100$)	$\alpha = 0.999$ ($R = 200$)	$\alpha = 0.99$ ($R = 20$)	$\alpha = 0.995$ ($R = 100$)	$\alpha = 0.999$ ($R = 200$)
Hang Seng	Empirical	7.099	8.454	11.791	3.974	6.682	8.169
	GP	7.056(0.296) [0.932 1.079]	8.379(0.515) [0.916 1.114]	11.939(1.156) [0.818 1.154]	3.757(0.253) [0.899 1.110]	5.943(1.026) [0.749 1.256]	6.781(1.779) [0.685 1.346]
	VG	6.302(0.209) [0.948 1.057]	7.169(0.280) [0.918 1.059]	9.221(0.697) [0.889 1.120]	3.919(1.659) [0.435 1.727]	5.995(4.984) [0.103 2.438]	6.879(7.524) [[0.012 3.382]
	NIG	6.906(0.289) [0.942 1.066]	8.048(0.395) [0.922 1.073]	10.835(0.989) [0.846 1.133]	4.059(1.820) [0.409 1.800]	6.581(5.769) [0.094 2.491]	7.754(8.777) [0.010 3.459]
	HYP	6.359(0.214) [0.948 1.058]	7.243(0.286) [0.917 1.059]	9.294(0.666) [0.893 1.113]	3.937(1.678) [0.431 1.730]	6.064(5.067) [0.102 2.444]	6.978(7.658) [0.011 3.393]
	GH	6.997(0.306) [0.939 1.069]	8.215(0.424) [0.918 1.077]	11.266(1.102) [0.837 1.143]	4.068(1.848) [0.405 1.832]	6.678(5.957) [0.010 3.480]	7.927(9.069) [0.010 3.480]
Nikkei 225	Empirical	4.856	5.536	6.614	3.009	4.535	5.289
	GP	4.706(0.153) [0.951 1.060]	5.299(0.200) [0.938 1.055]	6.657(0.466) [0.892 1.101]	2.892(0.206) [0.876 1.095]	4.118(0.578) [0.772 1.243]	4.500(0.946) [0.693 1.339]
	VG	4.769(0.139) [0.955 1.050]	5.336(0.182) [0.928 1.051]	6.622(0.345) [0.916 1.097]	2.944(1.015) [0.532 1.679]	4.317(2.889) [0.049 2.576]	4.882(5.787) [0.003 3.370]
	NIG	4.926(0.157) [0.951 1.055]	5.574(0.210) [0.921 1.057]	7.098(0.500) [0.895 1.112]	2.963(1.246) [0.396 1.692]	4.427(4.112) [0.065 2.797]	5.058(5.983) [0.010 3.058]
	HYP	4.857(0.148) [0.953 1.052]	5.463(0.195) [0.948 1.066]	6.861(0.457) [0.896 1.117]	2.956(1.134) [0.486 1.665]	4.380(3.435) [0.045 2.382]	4.979(6.260) [0.005 3.255]
	GH	4.901(0.153) [0.952 1.054]	5.531(0.204) [0.922 1.056]	6.997(0.478) [0.898 1.108]	2.962(1.241) [0.397 1.689]	4.411(4.084) [0.066 2.779]	5.029(5.917) [0.010 3.032]

Note: Estimates in daily % return terms based on the parameter values shown in Tables 1 and 2, where α is the confidence interval used for GP, and R is the absolute risk aversion used for Lévy models. The holding period is 1 day. SE's are reported besides each estimate and 90% normalized (by means of bootstrapped estimates) confidence intervals are reported right below. The parameters used are those obtained through calibrations. Most appropriate ES estimate is reported in bold; and the most appropriate SRM estimate is in bold and italic.

5. ESTIMATES OF ES AND LÉVY SRM

In this section, authors attempt to discuss and compare the estimates of ES and SRM that are based on GP and four chosen Lévy models. The calibrated parameters for each are as in Table 4. The ES risk measure is based on a high confidence level and the SRM risk measure is based on a large risk aversion parameter. We consider confidence interval α to be 0.99, 0.995, and 0.999 for ES, and the coefficient of absolute risk measure R to be 20, 100, and 200. Thus, both the risk measures cater for high trading losses due to extreme events. Across all models, ES estimates are larger than SRM estimates. Furthermore, across all the models and indices, both ES and SRM behave in a similar manner. The estimates of risk measures increase as the confidence levels and the absolute risk coefficients get larger. For all indices, both the risk measures are approximately of the same order of magnitude across the empirical estimates.

We now focus on the precision of the risk estimates. For the GP model, SRM standard errors are often higher than those of ES (clearly at higher risk levels) for all indices. Thus, we could argue that for GP model, ES estimation is more precise than that of SRM. Further, we note that ES estimates have higher coefficients of variation (estimated risk measure divided by standard error) for GP model than SRM estimates. This further suggests that GP-SRM estimation is more precise than that of ES estimation.

In Table 4, bootstrapped estimates of the 90% confidence interval for both the risk measures are included. For GP models across all the indices, these confidence intervals are narrower for ES estimates than those for SRM estimates, suggesting that risk exposures of ES are estimated more precisely, in general, for the GP model across indexes than is the case for the SRM. Furthermore, we also find that the estimated confidence intervals are symmetric for low confidence levels. However, at higher risk levels, confidence levels are asymmetric with the right bound further away from the mean of the bootstrapped estimates. Cotter and Dowd (2006) also find that EV-ES standard errors are higher than the VaRs and that VaR is estimated more precisely than the ES. However, the EV-ES

has higher coefficients of variation than the VaRs suggesting that the ES is more precisely estimated. Similarly, the SRM estimates are less precise than VaR or ES estimates. We find that the ES estimates for four Lévy models across all the indices have narrower confidence intervals. Thus, the ES estimation, in general, based on Lévy models is more precise than that of GP.

Finally, the standardized confidence intervals for the SRM estimates are similar across different positions, but increase with the risk aversion parameters. Similar to ES, the confidence intervals for very high R value display asymmetry. Furthermore, the SRM estimates have considerably wider confidence intervals than ES. Thus, the SRM estimation is less precise than that of ES. This discrepancy can be explained by the sample size. If we have n observations in the tail, the ES make use of these observations only. The SRM estimator, on the other hand, makes use of all observations and places considerable weight on a small subset of these tail observations and thus effectively uses a smaller sample size (Cotter & Dowd, 2005). Unfortunately, authors cannot investigate similar precision issues for Lévy models and/or between Lévy and GP models. This is because we do not have analytic formulae for either of the Lévy coherent risk measures, and subjective numerical implementations are not comparable with analytic implementations.

6. DISCUSSION

Full density-based Lévy models are considered for estimating the tail-based coherent risk measure ES (with coverage level as parameter) and entire distribution-based coherent risk measure SRM (with degree of risk aversion R as parameter), alongside the estimation of ES and SRM for a tail targeting EV model. Our findings reveal that between the approaches (Lévy and EV), neither is comprehensively superior to the other.

In Table 5, we report the frequency of significant EV and Levy estimates presented in Table 4. We have 15 SRM estimates of risk under two types of risk models (EV and Lévy), which are the estimates across all five indexes and three coverage levels under consideration. Similarly there

Table 5. Frequency distribution of significant estimates of ES and Lévy spectral risk measures

	Extreme value models		Total	Lévy models		Total
	ES	SRM		ES	SRM	
GP	8 (15)	0 (15)	8(30)	–	–	–
VG	–	–	–	1 (15)	1 (15)	2 (30)
NIG	–	–	–	3 (15)	6 (15)	9 (30)
HYP	–	–	–	1 (15)	2 (15)	3 (30)
GH	–	–	–	2 (15)	6 (15)	8 (30)
Total	8	0	–	7	15	22

Note: This table reports the number of significant estimates found in Tables 4 to 8. Numbers in parentheses are total number of estimates.

are 15 ES estimates. Among 15 ES estimates, the EV model provides 8 most favorable estimates. We have found 7 Lévy-ES and 15 Lévy-SRM estimates that are significant. Among 15 Lévy-SRM estimates, 11 estimates are closer to respective empirical values. Among the Lévy models, however, the NIG (4 out of 11) and GH (6 out of 11) models provide much more closer estimates in the sense of minimum absolute deviation from empirical values compared to the VG and HYP (1 out of 11) Lévy-SRM models. This is consistent with the evidence for derivative pricing (Schouten, 2003). Eberlein and Prouse (2002) also find that the VaR computed parametrically for the GH distribution is closer to the empirically observed Value-at-Risk. This is due to the fact that a fully flexible GH model estimates the quantiles more consistently than its restricted versions. Among the restricted versions, however, the restriction effect is minimal for NIG characterization. Among the Lévy-ES models, 3 out of 7 favorable estimates belong to the NIG model and 2 out of 7 to the GH model. For the remaining 2 favorable estimates, one belongs to the VG and one to the HYP model. Thus, the assumption that a tail-targeting model is more likely to provide superior consistent estimates for tail-based ES is empirically disputed. However, while we cannot say that a particular full density Lévy model is superior for all ranges of the data, we also cannot say that the tail density EV model is adequate for all ranges of data. Given such similarity of estimation performances under both approaches, the choice is likely to be driven by a compromise between simplicity and adequacy.

The Anderson-Darling test emphasizes the fit on the tail. The outcome is that a quantile mismatch

outside the tail is hardly detected by a left-truncated version of the AD test applied to the tail-based EV model. The EV model might have some preference over Lévy models based on AD test results in Table 3. However, we note that the AD test statistics are based on a tail quantile match of EV and bear no information regarding the quantile match outside the tail. For this reason, an AD test value of a solely tail-based model can be deceptive when compared with the AD test value of an entire distribution-based Lévy model. Furthermore, such deception cannot be adequately detected by a tail emphasized goodness of fit test. As a result, the seemingly preferable EV model does not always give us the most 'adequate' (in the sense of closest to empirical values) estimates of risk measures of ES.

We find that the extreme value model of ES outperforms Lévy models in terms of precision. This is in spite of the fact that the Lévy densities are calibrated on the entire data set and the EV models are calibrated using only the extreme observations. Thus, the tail-based risk measures of ES perform better if the model is calibrated on the tail alone. Additionally, SRM apparently incorporates all quantiles with corresponding probability mass spread on the entire spectrum. When compared to empirical values, the SRMs of Lévy models often outperform those of the EV model. To explain this feature, we note that SRM is not a tail based risk measure, whereas EV parameters are calibrated using only extreme quantiles. We observe that such a calibration could yield misleading quantiles, especially when outcomes fall in the most extreme end of the opposite tail. On the other hand, Lévy models that consider the entire data set in

calibration are expected to generate consistent values of quantiles far from the extreme tail⁹.

The practical implication of our findings is that portfolio managers or hedge fund managers should consider Lévy SRM models, as they provide the opportunity to choose appropriate risk aversion parameters to determine the expected loss in the event of market turmoil. In addition, the SRMs of Lévy models often outperform those of the EV model when compared to their empiri-

cal values. This becomes increasingly apparent as investors become more risk averse. We note that even though VaR is more precise than the ES risk measure, professional investors these days are more likely to use ES on account of that such measure has the benefit of coherence. Similarly, even though the SRM measure is less precise than the extreme value measure, it has the benefit of ascertaining investor's risk aversion while most closely matching losses empirically, especially in the case of the Lévy-SRM.

CONCLUSION

Authors apply various methodologies to estimate the coherent risk measures ES and SRM for Lévy models. We focus on a subclass of Lévy processes: variance gamma (VG), normal inverse Gaussian (NIG), hyperbolic distribution (HYP) and generalized hyperbolic (GH). On the basis of empirical evidence from all major indexes, our study suggests that the extreme value spectral risk measures are inadequate. The tail-based risk measures of ES are often found to be more consistent than those of Lévy models. Observations discarded by EV but incorporated by Lévy models are found not to make a significant improvement in the performance of tail-based risk measures. However, the EV calibration is restricted to the tails alone, and provides estimates of quantiles outside the fixed tails, which in turn yield poor estimates of the spectral risk measure itself. Lévy spectral risk measures, in contrast, often perform better than extreme spectral risk measures when compared with empirical values. This becomes increasingly apparent as investors become more risk averse.

REFERENCES

- Acharya, V., Engle, R., Richardson, M. (2012). Capital shortfall: a new approach to ranking and regulating systemic risks. *American Economic Review*, 102(3), 59-64. Retrieved from <http://www.jstor.org/stable/23245505>
- Acharya, V., Pedersen, L., Philippon T., Richardson, M. (2010). *Measuring Systemic Risk* (Technical report, Department of Finance, NYU).
- Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, 26, 1505-1518. [https://doi.org/10.1016/S0378-4266\(02\)00281-9](https://doi.org/10.1016/S0378-4266(02)00281-9)
- Acerbi, C. (2004). Coherent representation of subjective risk aversion. In G. Szego (Ed.), *Risk Measures for the 21st Century* (pp. 147-207). Chichester: John Wiley and Sons Ltd.
- Adam, A., Houkari, M., Laurent, J. (2008). Spectral risk measures and portfolio optimization. *Journal of Banking and Finance*, 32(9), 1870-1882. <https://doi.org/10.1016/j.jbankfin.2007.12.032>
- Anna, C., Rachev, S., Fabozzi, F. (2005). *Composite Goodness-of-Fit Tests for Left-Truncated Loss Samples* (Working Paper). Department of Statistics and Applied Probability. University of California, Santa Barbara.
- Artzner, P., Delbaen, F., Eber, J. M., Heath, D. (1999). Coherent Measure or Risk. *Mathematical Finance*, 9, 203-228.
- Bertoin, Jean. (1996). *Lévy Processes*. Cambridge University Press: Cambridge, UK.
- Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society. London A, Issue 353*, 401-419.
- Barndorff-Nielsen, O. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics*, 5, 151-157.
- Barndorff-Nielsen, O. (1995). *Normal Inverse Gaussian distributions and the modeling of stock returns* (pp. 401-419). Research report no 300. Department of

⁹ Note that we consider only 100 slices in evaluating the integrals of SRM, as is uniformly considered for all models. Cotter and Dowd (2006) reports the SRM for EV considering one million slices. This is almost impossible to apply for Lévy models. Even so, the Lévy SRM with 100 slicing for numerical integration is in general fairly comparable with their EV SRM and stands superior for the NIG and GH models in particular.

- Theoretical Statistics, Aarhus University, A 353.
12. Bingham, N. H., Kiesel, R. (2001). Modeling asset returns with hyperbolic distributions. In *Return Distributions on finance* (pp. 1-20). Butterworth-Heinemann.
 13. Brownlees, C., Engle, R. (2012). *Volatility, Correlation and Tails for Systemic Risk Measurement* (Working Paper Series). Department of Finance, NYU. Retrieved from <http://ssrn.com/abstract=1611229>
 14. Chen, Z., Wang, Y. (2008). Two sided coherent risk measures and their application in realistic portfolio optimization. *Journal of Banking and Finance*, 32(12), 2667-2673. <https://doi.org/10.1016/j.jbankfin.2008.07.004>
 15. Christoffersen, P. (2003). *Element of Financial Risk Management*. Academic Press.
 16. Cotter, J., Dowd, K. (2006). Extreme spectral risk measures: An application to futures clearinghouse margin requirements. *Journal of Banking and Finance*, 30, 3469-3485. <https://doi.org/10.1016/j.jbankfin.2006.01.008>
 17. Deng, X., Zheng, Y., Shao, P. (2009). Portfolio optimization based on spectral risk measures. *International Journal of Mathematical Analysis*, 34, 1657-1668. Retrieved from <http://www.m-hikari.com/ijma/ijma-password-2009/ijma-password33-36-2009/zhaopeibiao1-JMA33-36-2009.pdf>
 18. Dowd, K. (2005). *Measuring Market Risk*. John Wiley & Sons Ltd.
 19. Dowd, K., Cotter, J., Sorwar, C. (2008). Spectral Risk Measures: Properties and Limitations. *Journal of Financial Services and Research*, 34, 61-75. Retrieved from https://www.researchgate.net/publication/5150942_Spectral_Risk_Measures_Properties_and_Limitations
 20. Dowd, K., Cairns, A. J. C., Blake, D., Coughlan, C.D., Epstein, D., Khalaf Allah, M. (2010). Backtesting Stochastic Mortality Models: An Ex Post Evaluation of Multiperiod-Ahead Density Forecasts. *North American Actuarial Journal*, 14(3), 281-298. Retrieved from [http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/e6158fd2566f5dccc125780600434cbc/\\$FILE/LM_backtesting_models_dec2010.pdf](http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/e6158fd2566f5dccc125780600434cbc/$FILE/LM_backtesting_models_dec2010.pdf)
 21. Eberlein, E., Hammerstein, E. A. (2002). *The Generalized Hyperbolic and Inverse Gaussian Distributions: limiting cases and approximation of processes* (FDM preprint 80). University of Freiburg.
 22. Eberlein, E., Prause, K. (2002). The Generalized Hyperbolic model: financial derivatives and risk measures. *Proceedings of the Mathematical Finance – Bachelier Congress 2000, Springer*, 245-267. Retrieved from https://link.springer.com/chapter/10.1007/978-3-662-12429-1_12
 23. Fajardo, J. (2015). Barrier style contracts under Lévy processes: An alternative approach. *Journal of Banking & Finance*, 53, 179-187. <https://doi.org/10.1016/j.jbankfin.2015.01.002>
 24. Fajardo, J., Mordecki, E. (2006). Symmetry and duality in Lévy markets. *Quantitative Finance*, 6(3), 219-227.
 25. Fajardo, J., Mordecki, E. (2014). Skewness premium with Lévy processes. *Quantitative Finance*, 14(9), 1619-1626. <https://doi.org/10.1080/14697688.2011.618809>
 26. Fuse, G., Meucci, A. (2008). Pricing discretely monitored Asian options under Lévy processes. *Journal of Banking & Finance*, 32(10), 2076-2088. <https://doi.org/10.1016/j.jbankfin.2007.12.027>
 27. German, H. (2002). Pure jump Lévy processes for asset price modeling. *Journal of Banking & Finance*, 26(7), 1297-1316.
 28. Giannopoulos, K., Tunaru, R. (2005). Coherent risk measures under filtered historical simulation. *Journal of Banking and Finance*, 29(4), 979-996.
 29. Grootveld, H., Hallerbach, W. G. (2004). Upgrading value-at-risk from diagnostic metric to decision variable: A wise thing to do? In Szego (Ed.), *Risk Measures for the 21st Century* (pp. 33-50). Wiley, New York.
 30. Inui, K., Kijima, M. (2005). On the significance of expected shortfall as a coherent risk measure. *Journal of Banking and Finance*, 29(4), 853-864.
 31. Kim, Y. S., Rachev, S. T., Bianchi, M. L., Fabozzi, F. J. (2008). Financial market models with Lévy processes and time-varying volatility. *Journal of Banking and Finance*, 32(7), 1363-1378. <https://doi.org/10.1016/j.jbankfin.2007.11.004>
 32. Kyprianou, A. (2006). *Introductory Lectures on Fluctuations of Levy Processes with Applications*. Springer.
 33. Prause, K. (1999). *The Generalized Hyperbolic Model: Estimation, financial derivatives and risk measures* (PhD Thesis, University of Freiburg).
 34. Sorwar, G., Dowd, K. (2010). Estimating financial risk measures for options. *Journal of Banking and Finance*, 34(8), 1982-1992. <https://doi.org/10.1016/j.jbankfin.2010.01.005>
 35. Sato, Ken-Iti. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press: Cambridge, UK.
 36. Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley & Sons Ltd.
 37. Wächter, H. P., Mazzoni, T. (2013). Consistent modeling of risk aversion behavior with spectral risk measures. *European Journal of Operational Research*, 229, 487-495. <https://doi.org/10.1016/j.ejor.2013.03.001>
 38. Wong, H. Y., Guan, P. (2011). An FFT-network for Lévy option pricing. *Journal of Banking & Finance*, 35(4), 968-999. <https://doi.org/10.1016/j.jbankfin.2010.09.014>

APPENDIX

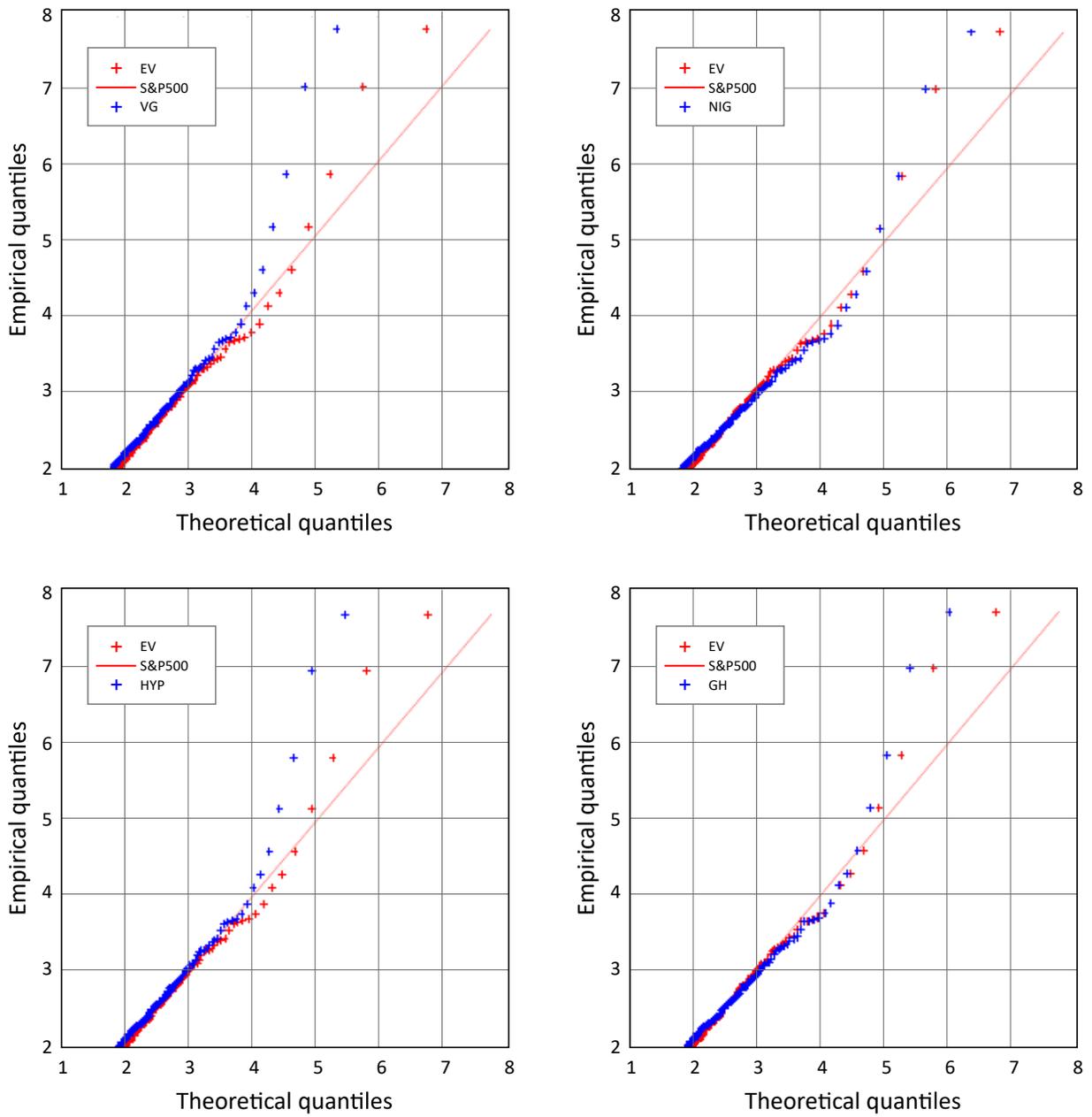


Figure 1. Conditional and unconditional quantiles in excess of threshold (2): long position in S&P 500

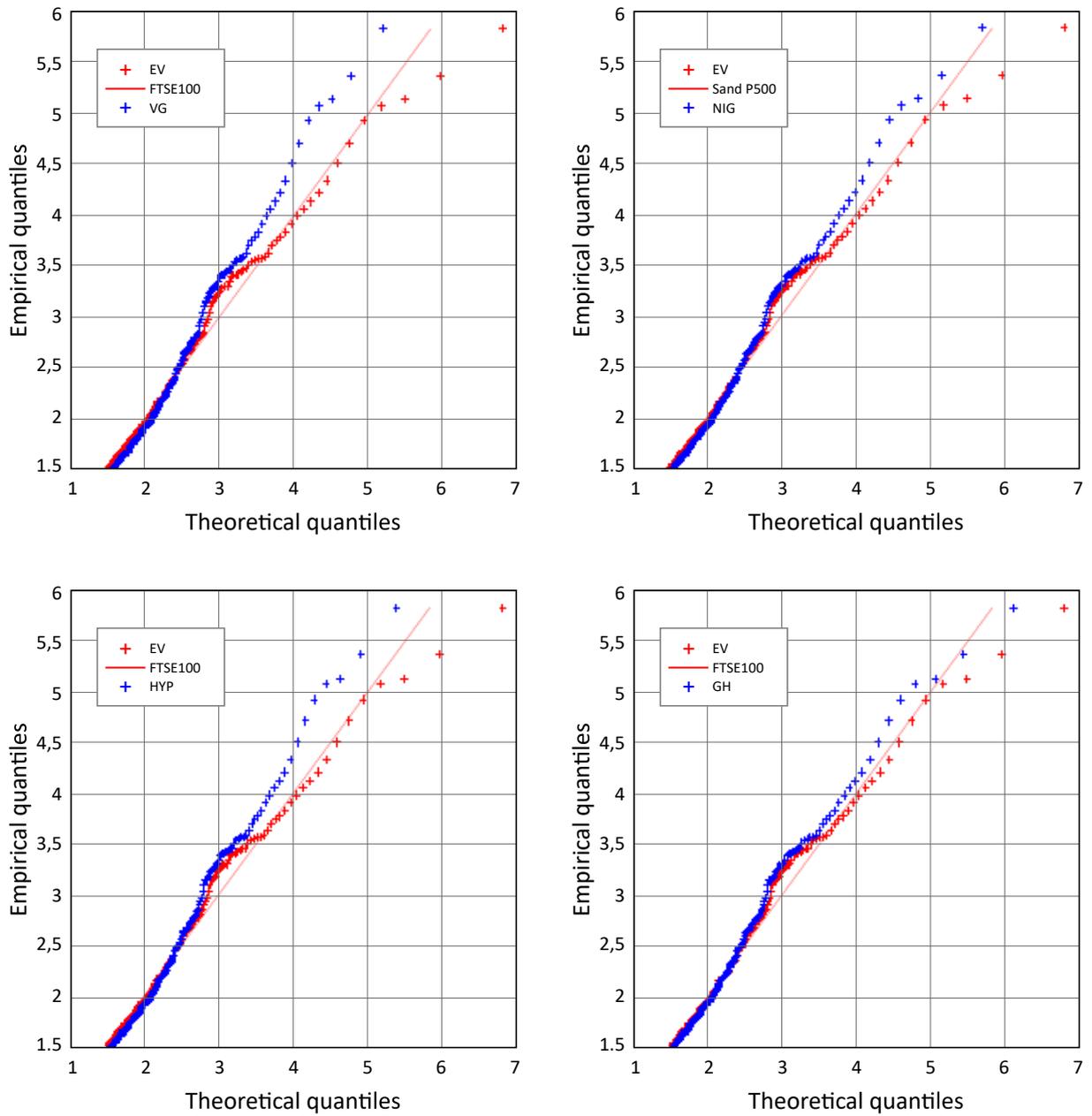


Figure 2. Conditional and unconditional quantiles in excess of threshold (1.5): long position in FTSE 100

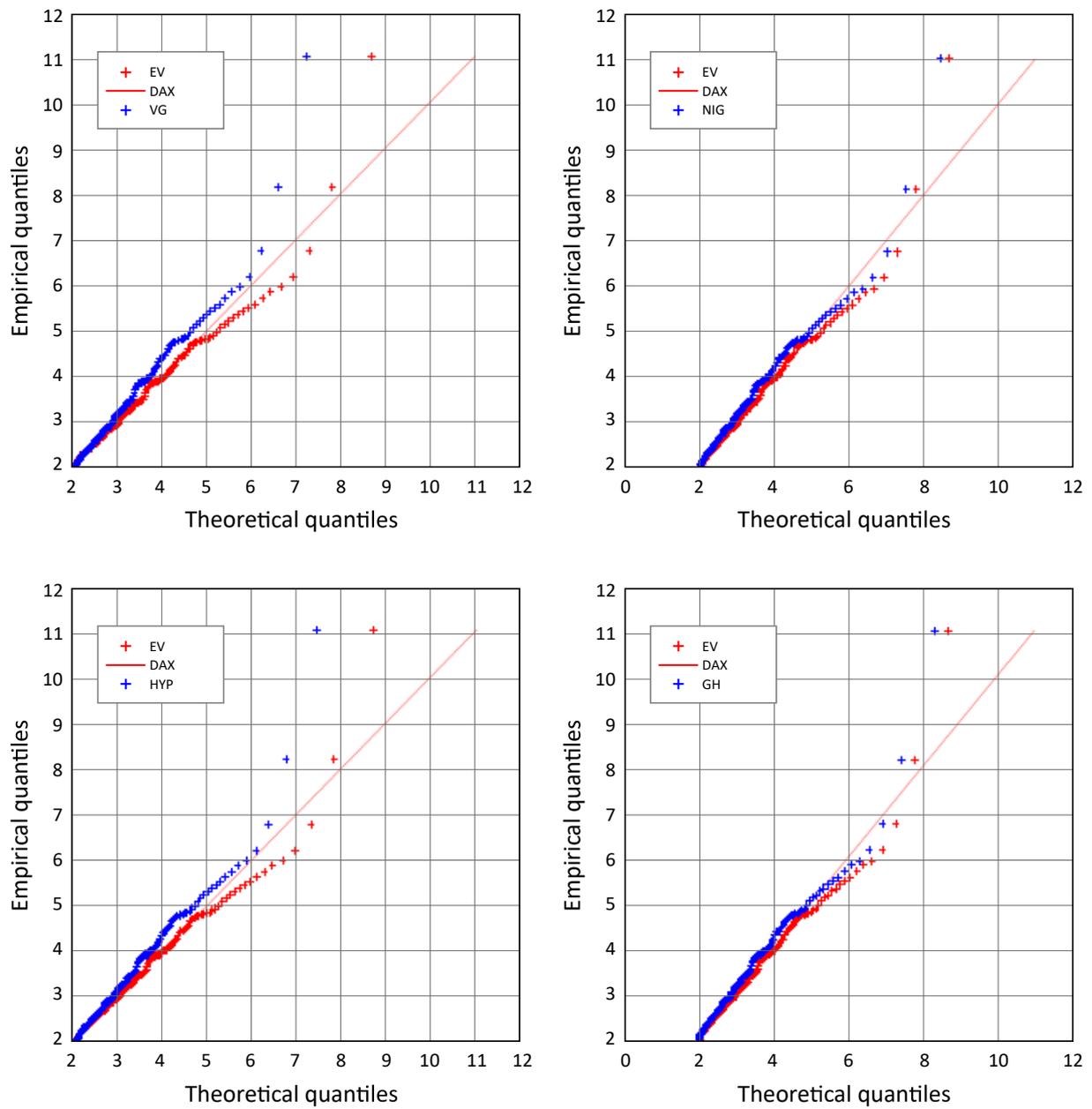


Figure 3. Conditional and unconditional quantiles in excess of threshold (2): long position in DAX

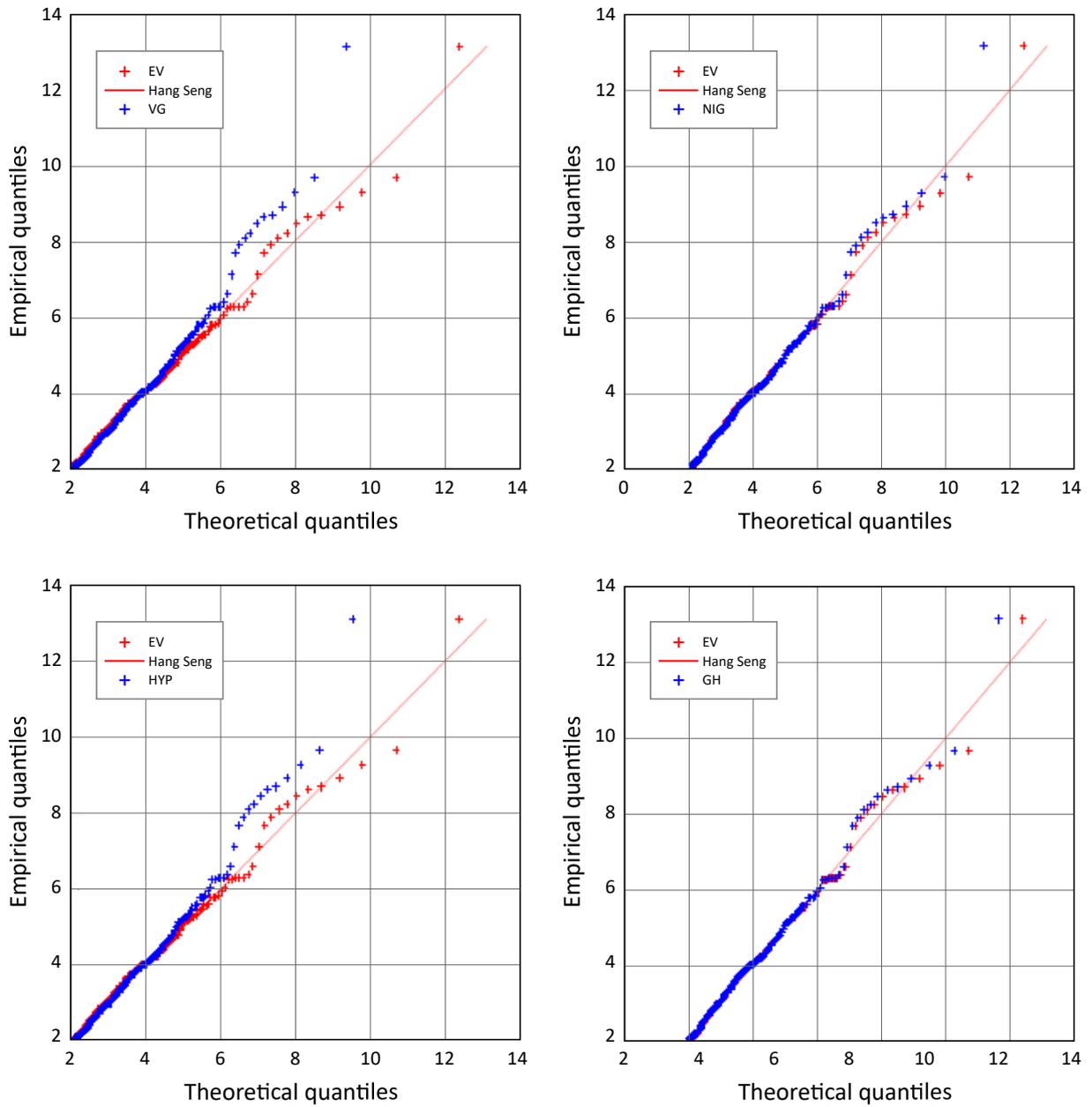


Figure 4. Conditional and unconditional quantiles in excess of threshold (2): long position in Hang Seng

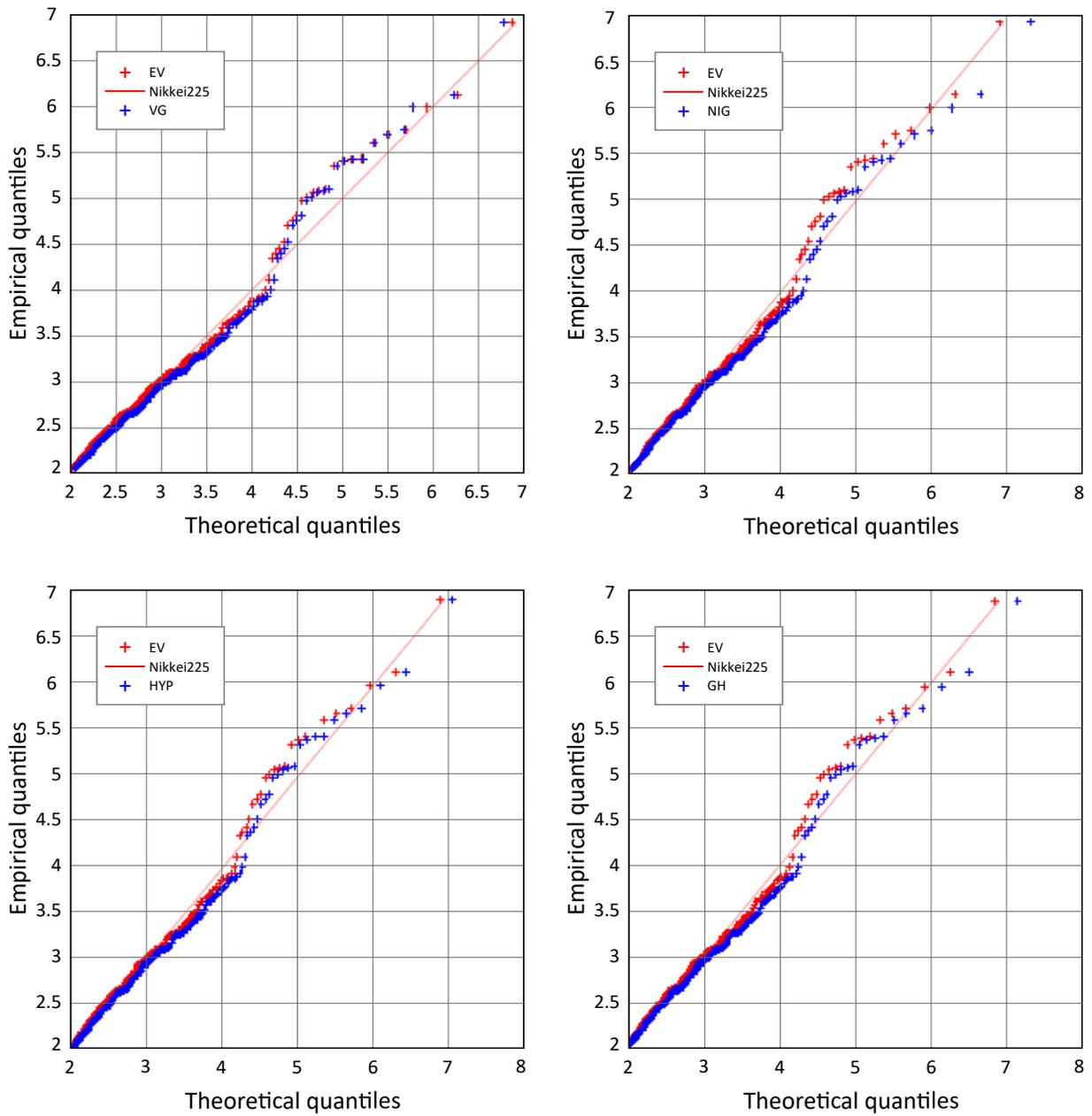


Figure 5. Conditional and unconditional quantiles in excess of threshold (2): long position in Nikkei 225