Modeling asset returns under time-varying semi-nonparametric distributions

Leon, A. and Ñíguez, T.M.

Modeling asset returns under time-varying semi-nonparametric distributions

Ángel León
Dept. Fundamentos del Análisis Económico, Universidad de Alicante,
Campus San Vicente del Raspeig, Alicante 03080, Spain.
Email: aleon@ua.es

Trino-Manuel Ñíguez*
Westminster Business School, University of Westminster,
35 Marylebone Road, London NW1 5LS, UK.
Email: T.M.Niguez@wmin.ac.uk

Abstract

We extend the semi-nonparametric (SNP) density of León, Mencía and Sentana (2009) to time-varying higher-order moments for daily asset return innovations of stock indexes and foreign-exchange rates. We estimate robust tail-indexes for testing the existence of the unconditional higher-order moments. We obtain closed-form expressions of partial moments and expected shortfall under the time-varying SNP density with the GJR-GARCH for modeling returns. A comparative study between SNP and Hansen's skewed-t, based on skewness-kurtosis frontiers, in-sample and backtesting analyses, is also implemented. Finally, we conduct an out-of-sample portfolio selection exercise for the stocks of the S&P 100 index through an equity screening method based on our parametric one-sided reward/risk performance measures and compare with the Sharpe ratio portfolio.

Keywords: Backtesting; equity screening; expected shortfall; conditional higher-order moments; tail-index.

JEL classification codes: C22, G11, G17.

*Corresponding author

1 Introduction

Optimal asset allocation relies critically on the modeling of asymmetry and tail-fatness of portfolio return distributions.[1] Recent econometric results have shown the importance of clustering and asymmetric response of time-varying (TV) skewness and/or kurtosis to positive and negative shocks. These TV or conditional higher-moments are implicitly modeled through the TV shape parameters of the distribution assumed for the innovations of asset returns. This framework has been employed, among others, by Jondeau and Rockinger (2003) (JR hereafter) with the skewed-t (ST) distribution of Hansen (1994); León and Ñíguez (2020) and León, Rubio and Serna (2005) with a transformation of the Gram-Charlier (GC) distribution in JR (2001); Bali, Mo and Tang (2008) for the skewed generalized t of Theodossiou (1998); Feunou, Jahan-Parvar and Tédongap (2016) who introduce the skewed generalized error distribution (GED); and Lalancette and Simonato (2017) for the Johnson S_u distribution. Following this literature, we propose an analytically tractable probability density function (pdf) for modeling asset return innovations, which extends the semi-nonparametric (SNP) density of León, Mencía and Sentana (2009) by incorporating TV parameters and, as a result, conditional higher-order moments. The SNP density is obtained as an expansion of the standard normal in terms of Hermite polynomials. Besides, it is more flexible than, for example, the sum of independent restricted GC densities in Zoia, Biffi and Nicolussi (2018), as the latter accounts only for excess kurtosis. It also allows to model financial time series that can only take positive values, such as the volatility index VIX, by making use of the SNP expansion of the Gamma density; see Mencía and Sentana (2018) for details.

Our paper proceeds as follows. First, we propose various possible specifications for the TV behavior of the SNP parameters for the conditional distribution of returns with the specific GARCH-family model by Glosten, Jagannathan and Runkle (1993) (GJR hereafter). In short, our model, henceforth referred to as TV-SNP-GJR, aims to capture stylized features in asset returns such as non-constant conditional skewness and kurtosis. Second, we provide evidence on the model suitability through an in-sample analysis for a set of daily stock-index and foreign-exchange (FX) returns. We also perform a comparative analysis between the SNP and ST distributions. For that purpose, we examine the regions of skewness and kurtosis that the distributions can generate. Besides, we use Vuong's closeness test (1989) for the in-sample fit, an analysis for the fit of the distribution tails, and also the backtesting approach of Du and Escanciano (2017) to study differences in forecasting Value-at-Risk (VaR) and expected shortfall (ES) for long and short trading positions. Third, recent evidence in the literature has shown that the tails of the unconditional distributions of daily or weekly returns can be well approximated by power laws,[2] which are characterized by the behavior of the tail index. Its value determines the maximal order of finite moments of return distributions. For instance, if $\zeta \in (2,3)$ then the first two moments (mean and variance) exist; if $\zeta \in (2,4)$ then the skewness also exist; and if $\zeta \in (2,5)$ then the mean, variance, skewness and kurtosis of returns exist. Suppose that $\zeta > 4$, then the lower the tail index, the higher the kurtosis, and vice versa.[3] Hence, the finiteness of first moments is relevant for risk managers, financial regulators and also, for investors with preferences going beyond the typical Markowitz framework, relying on the first two finite moments, to the finite higher-order

[1] See Bernardi and Catania (2018) and references therein.

[2] See Gabaix, Gopikrishnan, Plerou and Stanley (2006), Gabaix (2009) and Ibragimov, Ibragimov and Walden (2015).

[3] The relation between the GARCH unconditional kurtosis and the tail-index can be seen in Mikosch and Starica (2000).

moments of skewness and kurtosis.[4] Thus, the estimation of tail-indexes is of key interest. Here, we estimate robust tail-indexes, based on the methodology of Gabaix and Ibragimov (2011), for testing finiteness in the first unconditional moments.

Furthermore, we derive closed-form expressions of conditional one-sided reward/risk measures under the TV-SNP-GJR specification, and obtain their corresponding performance measures (PMs), as in León and Moreno (2017) for the unconditional return distribution under the GC specification. In an empirical application, we design active portfolio strategies through equity screening rules based on our PMs for ranking stocks and then building portfolios. These PMs extend those in León, Navarro and Nieto (2019) obtained under historical simulation (HS). The PMs we study are listed as follows: (a) The Sharpe ratio (SR) (Sharpe, 1966, 1994) as the benchmark. (b) The skewness and kurtosis ratio (SKR), see Watanabe (2006). (c) PMs based on partial moments, such as (i) the Farinelli-Tibiletti (FT) ratio, which nests the popular Omega and Upside potential ratios, see Farinelli and Tibiletti (2008), and (ii) the Sortino ratio, see Sortino and Van der Meer (1991). (d) Quantile-based PMs, such as the Rachev or expected tail ratio (ETR), and the Value-at-Risk ratio (VaRR). See Biglova, Ortobelli, Rachev and Stoyanov (2004) and Caporin and Lisi (2011) for these two last measures, respectively.

Finally, we implement an out-of-sample (OOS) analysis for portfolios composed from selecting among the stocks that constitute the S&P 100 index using these alternative PM strategies. Cumulative portfolio returns are obtained over the OOS period for each PM strategy and compared with the SR portfolio returns. Our empirical findings show evidence of considerable gains in both SKR and ETR portfolio cumulative returns.[5]

The remainder of the article is organized as follows. In Section 2, we introduce the TV-SNP-GJR models, discuss some statistical properties and obtain closed-form expressions for the ES and partial moments used to build conditional parametric PMs. Section 3 discusses the previous model estimation through an empirical application to stock index and FX returns, as well as it provides a comparative analysis with respect to the ST. Section 4 shows the performance of OOS portfolios by means of equity screening based on PMs for ranking stocks that compose the S&P 100 index. In Section 5, we summarize our conclusions. All proofs are provided in Appendix 1. Appendix 2 contains the conditional PMs used in our analysis. Appendix 3 derives the confidence interval for testing the existence of the unconditional fourth moment, and hence kurtosis, based on the Delta method. Appendix 4 provides both SNP and ST theoretical quantiles for a comparative analysis.

2 Modeling asset returns

Let the asset return r_t be a process characterized by the sequence of conditional densities $f(r_t|I_{t-1};\boldsymbol{\psi})$, where I_{t-1} denotes the information set available prior to the realization of r_t, $\boldsymbol{\psi} = (\boldsymbol{\theta}, \boldsymbol{\nu})$ is the vector of unknown parameters such that $\boldsymbol{\theta}$ is the subset characterizing both the conditional mean and variance of r_t, i.e. $\mu_t(\boldsymbol{\theta}) = \mu(I_{t-1};\boldsymbol{\theta})$ and $\sigma_t^2(\boldsymbol{\theta}) = \sigma(I_{t-1};\boldsymbol{\theta})$, and finally, $\boldsymbol{\nu}$ is the subset characterizing the shape of the

[4] See Ñíguez, Paya, Peel and Perote (2019), Boudt, Lu and Peeters (2015), Xiong and Idzorek (2011) and Jondeau and Rockinger (2006).

[5] Our results are also in line with those in León et al. (2019) where the best portfolio performance is obtained under the Generalized Rachev ratio, which nests the ETR used in this study.

distribution of the innovations, z_t. Thus, we assume that

$$r_t = \mu_t(\boldsymbol{\theta}) + \varepsilon_t, \qquad \varepsilon_t = \sigma_t(\boldsymbol{\theta}) z_t, \qquad z_t \sim i.i.d.\ g(z_t; \boldsymbol{\nu}). \tag{1}$$

So, equation (1) decomposes the return at time t into a conditional mean, μ_t, and the term ε_t defined as the product between the conditional standard deviation, σ_t, and the innovation (or standardized return), z_t, with zero mean and unit variance. It is assumed that $\{z_t\}$ is a sequence of independent identically distributed ($i.i.d.$) random variables with $g(\cdot)$ as pdf. A TV distribution with $g(z_t; \boldsymbol{\nu}_t)$ as pdf is obtained with a dynamic specification of the parameter vector $\boldsymbol{\nu}_t$, then $\{z_t\}$ are neither independent nor identically distributed, with $z_t | I_{t-1} \sim g(z_t; \boldsymbol{\nu}_t)$ as conditional distribution.

2.1 SNP density of z_t

Let us define z_t as a linear transformation of x_t with pdf given by the SNP distribution of León et al. (2019),

$$z_t = a(\boldsymbol{\nu}) + b(\boldsymbol{\nu}) x_t, \qquad b = 1/\sigma_x, \quad a = -b\mu_x, \tag{2}$$

where $\mu_x = E(x_t)$ and $\sigma_x = \sqrt{V(x_t)}$ are, respectively, the mean and the standard deviation of x_t with density function transformed according to the Gallant and Nychka (1987) method:

$$q_n(x_t) = \frac{\phi(x_t)}{\boldsymbol{\nu}'\boldsymbol{\nu}} \left(\sum_{k=0}^{n} \nu_k H_k(x_t) \right)^2, \tag{3}$$

where $\boldsymbol{\nu} = (\nu_0, \nu_1, \ldots, \nu_n)' \in \mathbb{R}^{n+1}$, $\phi(\cdot)$ denotes the pdf of a standard normal random variable and $H_k(\cdot)$ are the normalized Hermite polynomials. These polynomials can be defined recursively for $k \geq 2$ as

$$H_k(x) = \frac{x H_{k-1}(x) - \sqrt{k-1} H_{k-2}(x)}{\sqrt{k}}, \tag{4}$$

with initial conditions $H_0(x) = 1$ and $H_1(x) = x$. The set $\{H_k(x)\}_{k \in N}$ constitutes an orthonormal basis with respect to the weighting function $\phi(x)$. Thus, $E_\phi[H_k(x) H_l(x)] = \mathbf{1}(k = l)$, where $\mathbf{1}(\cdot)$ is the usual indicator function and the operator $E_\phi[\cdot]$ takes the expectation of its argument with respect to $\phi(\cdot)$ as pdf.

Since $q_n(\cdot)$ in (3) is homogeneous of degree zero in $\boldsymbol{\nu}$, we impose $\nu_0 = 1$ to solve the scale indeterminacy. If we consider $n = 2$ and expand the square term expression in (3), we obtain an alternative expression of $q_2(\cdot)$ and, henceforth, denoted as $q(\cdot)$:

$$q(x_t) = \phi(x_t) \sum_{k=0}^{4} \gamma_k(\boldsymbol{\nu}) H_k(x_t), \tag{5}$$

such that $\gamma_0(\boldsymbol{\nu}) = 1$ and

$$\begin{aligned} \gamma_1(\boldsymbol{\nu}) &= \tfrac{2\nu_1(1+\sqrt{2}\nu_2)}{\boldsymbol{\nu}'\boldsymbol{\nu}}, & \gamma_2(\boldsymbol{\nu}) &= \tfrac{\sqrt{2}(\nu_1^2+2\nu_2^2+\sqrt{2}\nu_2)}{\boldsymbol{\nu}'\boldsymbol{\nu}}, \\ \gamma_3(\boldsymbol{\nu}) &= \tfrac{2\sqrt{3}\nu_1\nu_2}{\boldsymbol{\nu}'\boldsymbol{\nu}}, & \gamma_4(\boldsymbol{\nu}) &= \tfrac{\sqrt{6}\nu_2^2}{\boldsymbol{\nu}'\boldsymbol{\nu}}. \end{aligned} \tag{6}$$

2.1.1 Moments

The first four noncentral moments of x_t with pdf in (5) are:

$$\begin{aligned} \mu_x'(1) &= \gamma_1(\boldsymbol{\nu}), & \mu_x'(2) &= \sqrt{2}\gamma_2(\boldsymbol{\nu}) + 1, \\ \mu_x'(3) &= \tfrac{6\nu_1(1+2\sqrt{2}\nu_2)}{\boldsymbol{\nu}'\boldsymbol{\nu}}, & \mu_x'(4) &= \tfrac{12(\nu_1^2+3\nu_2^2+\sqrt{2}\nu_2)}{\boldsymbol{\nu}'\boldsymbol{\nu}} + 3. \end{aligned} \tag{7}$$

Hence, $\mu_x = \mu'_x(1)$ and $\sigma_x^2 = \mu'_x(2) - \mu_x^2$. Therefore, the skewness and kurtosis of z_t are given by

$$s_z \equiv E\left(z_t^3\right) = a^3 + 3a^2 b\mu'_x(1) + 3ab^2\mu'_x(2) + b^3\mu'_x(3), \tag{8}$$

$$k_z \equiv E\left(z_t^4\right) = a^4 + 4a^3 b\mu'_x(1) + 6a^2b^2\mu'_x(2) + 4ab^3\mu'_x(3) + b^4\mu'_x(4). \tag{9}$$

2.1.2 Cumulative distribution function (cdf)

Let $Q(\cdot)$ denote the cdf of x_t with $q(\cdot)$ as the pdf in (5). The pdf of z_t is given by $g(z_t) = \frac{1}{b(\boldsymbol{\nu})} q\left(\frac{z_t - a(\boldsymbol{\nu})}{b(\boldsymbol{\nu})}\right)$. The next result shows the expression of the cdf related to z_t.

Proposition 1. *The cdf of z_t in (2), denoted as $G(\cdot)$, is obtained as*

$$\begin{aligned} G(z_t) &= Q(z_t^*) = \int_{-\infty}^{z_t^*} q(x_t)\, dx_t \\ &= \Phi(z_t^*) - \phi(z_t^*) \sum_{k=1}^{4} \frac{\gamma_k}{\sqrt{k}} H_{k-1}(z_t^*), \end{aligned} \tag{10}$$

where $z_t^* = (z_t - a)/b$, $H_k(\cdot)$ is given in (4) and $\Phi(\cdot)$ denotes the cdf of the standard normal.

Proof. *It is verified that $\int_{-\infty}^{u} H_k(x)\phi(x)dx = -\frac{1}{\sqrt{k}} H_{k-1}(u)\phi(u)$, then (10) is directly obtained.*

2.2 GJR-SNP model

Let $\sigma_t^2 = E\left[\varepsilon_t^2 | I_{t-1}\right]$ be the GJR (1,1) conditional variance model. Then,

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \beta\sigma_{t-1}^2 + \alpha_1^+ \left(\varepsilon_{t-1}^+\right)^2 + \alpha_1^- \left(\varepsilon_{t-1}^-\right)^2 \\ &= \alpha_0 + \beta\sigma_{t-1}^2 + \alpha_1^+ \sigma_{t-1}^2 \left(z_{t-1}^+\right)^2 + \alpha_1^- \sigma_{t-1}^2 \left(z_{t-1}^-\right)^2, \end{aligned} \tag{11}$$

such that $\alpha_0 > 0$, $\beta \geq 0$, $\alpha_1^+ \geq 0$ and $\alpha_1^- \geq 0$. Consider $y_t^+ = \max(y_t, 0)$, $y_t^- = \min(y_t, 0)$ where y_t can be either ε_t or z_t defined in (1). Another representation of (11) is given by $\sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + (\alpha_1 + \gamma D_{t-1})\varepsilon_{t-1}^2$ where $D_{t-1} = 1$ if $\varepsilon_{t-1} < 0$ and $D_{t-1} = 0$ if $\varepsilon_{t-1} \geq 0$. Hence, both expressions are related through $\alpha_1^+ = \alpha_1$ and $\alpha_1^- = \alpha_1 + \gamma$. Henceforth, we denote (11) as simply the GJR model, which nests the GARCH model when $\alpha_1^+ = \alpha_1^-$.

2.2.1 Unconditional variance

Following He and Terasvirta (1999), we can rewrite the GJR in (11) as the following stochastic difference equation (SDE):

$$\sigma_t^2 = \alpha_0 + c_t\sigma_{t-1}^2, \tag{12}$$

where $c_t = \beta + \alpha_1^+ \left(z_{t-1}^+\right)^2 + \alpha_1^- \left(z_{t-1}^-\right)^2$. Note that for the GARCH, we have $c_t = \beta + \alpha_1 z_{t-1}^2$. If we assume (11) to be covariance stationary, then the unconditional variance of ε_t is obtained as

$$\sigma_\varepsilon^2 = E\left(\sigma_t^2\right) = \frac{\alpha_0}{1 - E(c_t)}, \tag{13}$$

such that $E(c_t) < 1$ and

$$E(c_t) = \beta + \alpha_1^+ + \left(\alpha_1^- - \alpha_1^+\right) E\left[\left(z_t^-\right)^2\right], \tag{14}$$

where $E\left[\left(z_t^-\right)^2\right]$ is obtained in Proposition 2 below for $k=2$. Note that if we assume $\mu_t=\mu$ in (1), then σ_ε^2 in (13) is the unconditional variance of r_t , and both ε_t and r_t have the same unconditional skewness and kurtosis.

Proposition 2. *Let $z_t=a+bx_t$ be the standardized variable defined in* (2) *and x_t an i.i.d. sequence with pdf given in* (5)*, then*

$$\begin{aligned} E\left[\left(z_t^-\right)^k\right] &= \int_{-\infty}^{0} z_t^k g\left(z_t\right) dz_t = \int_{-\infty}^{-a/b} \left(a+bx_t\right)^k q\left(x_t\right) dx_t \\ &= \sum_{j=0}^{k} \binom{k}{j} a^{k-j} b^j \xi_j\left(-a/b\right), \end{aligned} \tag{15}$$

where $k\in\mathbb{N}$ and $\xi_j\left(u\right)=\int_{-\infty}^{u} x^j q\left(x\right) dx$ is the equation in (45).

Proof. *See section ii) of Appendix 1.*

2.2.2 Unconditional fourth moment and kurtosis

If we square σ_t^2 in (12), then $\sigma_t^4=\alpha_0^2+c_t^2\sigma_{t-1}^4+2\alpha_0 c_t\sigma_{t-1}^2$. By taking expectations and assuming that $E\left(\sigma_t^4\right)=E\left(\sigma_{t-1}^4\right)$, then $\left[1-E\left(c_t^2\right)\right]E\left(\sigma_t^4\right)=\alpha_0^2+2\alpha_0 E\left(c_t\right)E\left(\sigma_t^2\right)$ and finally, we obtain the unconditional kurtosis of ε_t :

$$k_\varepsilon=\frac{E\left(\varepsilon_t^4\right)}{\sigma_\varepsilon^4}=k_z\frac{E\left(\sigma_t^4\right)}{\sigma_\varepsilon^4}=k_z\left(\frac{1-E\left(c_t\right)^2}{1-E\left(c_t^2\right)}\right), \tag{16}$$

such that k_z is defined in (9) and

$$E\left(c_t^2\right)=2\beta E\left(c_t\right)-\beta^2+\left(\alpha_1^+\right)^2 k_z+\left[\left(\alpha_1^-\right)^2-\left(\alpha_1^+\right)^2\right]E\left[\left(z_t^-\right)^4\right], \tag{17}$$

where $E\left[\left(z_t^-\right)^4\right]$ is obtained for $k=4$ in (15). The condition for the existence of the unconditional fourth moment is verified when $E\left(c_t^2\right)<1$.

Note that $z_t\sim i.i.d.\ N\left(0,1\right)$ when $\nu_1=\nu_2=0$ under the SNP distribution for x_t, then $s_z=0$ and $s_\varepsilon=0$. It is verified that $E\left[\left(z_t^-\right)^2\right]=1/2$, $E\left[\left(z_t^-\right)^4\right]=3/2$, $k_z=3$ and so, k_ε in (16) becomes the following expression:

$$k_\varepsilon=3\left(\frac{1-\beta^2-\beta\left(\alpha_1^++\alpha_1^-\right)-\frac{1}{4}\left(\alpha_1^++\alpha_1^-\right)^2}{1-\beta^2-\beta\left(\alpha_1^++\alpha_1^-\right)-\frac{3}{2}\left[\left(\alpha_1^+\right)^2+\left(\alpha_1^-\right)^2\right]}\right). \tag{18}$$

Finally, Ling and McAleer (2002) show expressions of $E\left(c_t^2\right)$ for alternative GARCH-family models (including the GJR) when $z_t\sim i.i.d.\ t\left(v\right)$ with $v\geq 5$ such that $z_t\sim i.i.d.\ N\left(0,1\right)$ when $v\to\infty$.

2.2.3 Power-law tail property

A consequence of using GARCH models is that they exhibit heavy-tails and hence, excess kurtosis for the unconditional distribution of returns regardless of the distribution of z_t, see Bai, Russell and Tiao (2003). Knowledge of the tail behavior of financial returns is in itself of great interest. In this paper, we study the tail shape of the empirical distribution of some returns series. The tail index, or Pareto exponent, is a measure of the fatness or heaviness (the rate of decay) of the tails under power law distributions. The

greater the probability mass in the tails, the smaller the tail index, and vice versa. The value of the tail index characterizes the maximal order of the finite moments of r_t. In our empirical analysis we compare the tail index estimates with the conditions related to the existence of unconditional moments under the estimated parametric distribution driven by the SNP-GJR model. We define heavy-tails by this power-law tail property in more detail below.

Consider the SDE representation for σ_t^2 in (12) which nests alternative GARCH-family equations under different specifications of c_t as a function on z_{t-1}, then $c_t = c(z_{t-1})$ such that $c_t > 0$. Suppose there exists a positive real number $\varrho > 0$ such that $E(c_t^{\varrho}) = 1$. According to the theory in Kesten (1973), the stationary solution of σ_t^2 follows a heavy-tailed distribution:

$$P\left\{\sigma_t^2 > x\right\} \sim Ax^{-\varrho}, \quad \text{as} \quad x \to \infty,$$

where ϱ is the tail index of σ_t^2 and $A > 0$ is the tail scale. Then, $P\{\sigma_t > x\} = P\left\{\sigma_t^2 > x^2\right\} \sim Ax^{-\zeta}$ where $\zeta = 2\varrho$ is the tail index of σ_t. Here, $f(x) \sim g(x)$ means $f(x) = g(x)(1 + o(1))$ as $x \to \infty$. Suppose that $E(|z_t|^{\zeta}) < \infty$, then Mikosch and Starica (2000) derive the following result:

$$P\{|\varepsilon_t| > x\} = P\{|\sigma_t z_t| > x\} \sim E\left(|z_t|^{\zeta}\right) P\{\sigma_t > x\}, \quad \text{as} \quad x \to \infty.$$

In short, $|\varepsilon_t|$ has a similar tail behavior as σ_t, i.e. the tail index of $|\varepsilon_t|$ equals ζ. For the existence of the p-th moment of ε_t, it must be verified that $E(|\varepsilon_t|^p) < \infty$. Since the value of ζ characterizes the maximal order of finite moments of ε_t, then

$$E(|\varepsilon_t|^p) < \infty \quad \text{if} \quad p < \zeta \quad \text{and} \quad E(|\varepsilon_t|^p) = \infty \quad \text{if} \quad p \geq \zeta. \tag{19}$$

According to (19), it is verified that $E(|\varepsilon_t|) < \infty$ if and only if $\zeta > 1$. The second moment $E\left(\varepsilon_t^2\right) < \infty$, and thus $\sigma_\varepsilon^2 < \infty$, if and only if $\zeta > 2$. The fourth moment $E\left(\varepsilon_t^4\right) < \infty$, and thus $k_\varepsilon < \infty$, if and only if $\zeta > 4$. In short, the condition $E\left(c_t^k\right) < 1$ with $k = 1, 2$ in (14) and (17) hold, respectively, if $\zeta > 2$ and $\zeta > 4$. Finally, suppose that $\zeta \in (2, 4)$, then $E(c_t) < 1$ and $E\left(c_t^2\right) \geq 1$ which implies a finite variance but an infinite kurtosis for the unconditional distribution of r_t with $\mu_t = \mu$ in (1).

2.3 Time-varying SNP parameters

Consider $\varepsilon_t = r_t - \mu_t$ in equation (1) and let $\sigma_t^2 = E\left[\varepsilon_t^2 | I_{t-1}\right]$ follow the GJR model in (11). Then, the conditional skewness and kurtosis of r_t are defined, respectively, as

$$s_{r,t} = \frac{E\left(\varepsilon_t^3 | I_{t-1}\right)}{\sigma_t^3}, \qquad k_{r,t} = \frac{E\left(\varepsilon_t^4 | I_{t-1}\right)}{\sigma_t^4}. \tag{20}$$

If we let the SNP distribution exhibit TV parameters, the pdf of x_t in (5) is now defined as $q(x_t | I_{t-1})$ where ν_i is replaced with $\nu_{i,t}$ being measurable with respect to the information set I_{t-1}. Hence, $s_{r,t} = s_{z,t}$ and $k_{r,t} = k_{z,t}$ are now TV such that both $s_{z,t}$ and $k_{z,t}$ are obtained by plugging $\nu_{i,t}$ into equations (8) and (9), respectively. We model $\nu_{i,t}$ according to the following autoregressive specification:

$$\nu_{i,t} = \varphi_{0i} + \varphi_{1i}\nu_{i,t-1} + \Upsilon_i^{\star}(z_{t-1}), \tag{21}$$

where $\Upsilon_i^\star(\cdot)$ is a real-valued function that aims to capture the news impact curve specification of both conditional skewness and kurtosis.[6] We consider a flexible model for $\Upsilon_i^\star(\cdot)$ and specifically, the equation labeled as 'transition model' in Anatolyev and Petukhov (2016):

$$\Upsilon_i^\star(z) = \varphi_{2i}^+(1+\varphi_{3i}|z|)z^+ + \varphi_{2i}^-(1+\varphi_{3i}|z|)z^-, \tag{22}$$

where $z^+ = \max(z,0)$ and $z^- = \min(z,0)$. The equation (22) does account for nonlinear dynamics in $\nu_{i,t}$ through the parameter φ_{3i}. Note that it nests the asymmetric linear model when $\varphi_{3i} = 0$ and $\varphi_{2i}^+ \neq \varphi_{2i}^-$, i.e. $\Upsilon_i^\star(z) = \varphi_{2i}^+ z^+ + \varphi_{2i}^- z^-$. The symmetric linear one corresponds to the case of $\varphi_{3i} = 0$ and $\varphi_{2i}^+ = \varphi_{2i}^-$, i.e. $\Upsilon_i^\star(z) = \varphi_{2i} z$. In short, hereafter any TV-SNP specification adopted here is nested in the general TV (GTV) model driven by equations (21) and (22), and denoted as GTV-SNP.

2.4 Log-likelihood function

Note that we have previously studied the main components that define the asset return equation given in (1). If we now express the conditional density of r_t in terms of the conditional density of x_t, then

$$f(r_t|I_{t-1};\boldsymbol{\psi}) = \frac{q(x_t|I_{t-1})}{b(\boldsymbol{\nu}_t)\sigma_t}, \tag{23}$$

where $\boldsymbol{\psi}$ is the whole parameter vector, $q(\cdot|I_{t-1})$ is the conditional pdf given in (5) with $\nu_{i,t}$ as TV parameters with general expression in (21), $x_t = \frac{z_t(\boldsymbol{\theta})-a(\boldsymbol{\nu}_t)}{b(\boldsymbol{\nu}_t)}$ and $z_t(\boldsymbol{\theta}) = (r_t - \mu_t(\boldsymbol{\theta}))/\sigma_t(\boldsymbol{\theta})$ such that $\boldsymbol{\theta} \subset \boldsymbol{\psi}$ contains the parameters driven by both conditional mean and variance of (1). The log-likelihood function corresponding to a particular observation r_t, denoted as l_t, takes the following form:

$$\begin{aligned} l_t &= -\frac{1}{2}\ln\left(\sigma_t^2(\boldsymbol{\theta})\right) - \ln(b(\boldsymbol{\nu}_t)) - \ln(\boldsymbol{\nu}_t'\boldsymbol{\nu}_t) - \frac{1}{2}\ln(2\pi) \\ &\quad -\frac{1}{2}\left(\frac{z_t(\boldsymbol{\theta})-a(\boldsymbol{\nu}_t)}{b(\boldsymbol{\nu}_t)}\right)^2 + \ln\left[\sum_{k=0}^{2}\nu_{k,t}H_k\left(\frac{z_t(\boldsymbol{\theta})-a(\boldsymbol{\nu}_t)}{b(\boldsymbol{\nu}_t)}\right)\right]^2, \end{aligned} \tag{24}$$

such that $\nu_{0,t} = 1$, $\nu_{i,t} = \nu_{i,t}(\boldsymbol{\vartheta}_i)$ where $\boldsymbol{\vartheta}_i \subset \boldsymbol{\psi}$ is the parameter vector underlying the equation of $\nu_{i,t}$ in (21) and hence, $\boldsymbol{\psi} = (\boldsymbol{\theta}, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)$. If we adopt, for instance, the GTV-SNP model, then $\boldsymbol{\vartheta}_i = \left(\varphi_{0i}, \varphi_{1i}, \varphi_{2i}^+, \varphi_{2i}^-, \varphi_{3i}\right)$. Finally, the log-likelihood function for the model is obtained as $LL = \sum_{t=1}^{T} l_t$ where T is the total sample size.

2.5 Conditional quantile and expected shortfall

Let $F(r_t|I_{t-1})$ denote the conditional cdf corresponding to the TV-SNP model of r_t with pdf in (23),

$$F(r_t|I_{t-1}) = \int_{-\infty}^{r_t} f(r_t|I_{t-1};\boldsymbol{\psi})\,dr_t = \int_{-\infty}^{r_t^*} q(x_t|I_{t-1})\,dx_t = Q(r_t^*|I_{t-1}), \tag{25}$$

where $Q(\cdot|I_{t-1})$ is the conditional cdf, which is just the cdf $Q(\cdot)$ in (10) but with TV-SNP parameters, and $r_t^* = (r_t - \mu_t - a_t\sigma_t)/b_t\sigma_t$ where $a_t = a(v_t)$ and $b_t = b(v_t)$. The α-quantile, or VaR at the α-confidence level, of the distribution of the asset return r_t is $r_{\alpha,t} = F^{-1}(\alpha|I_{t-1})$. So,

$$r_{\alpha,t} = \kappa_{0t} + \kappa_{1t}Q_t^{-1}(\alpha), \tag{26}$$

[6] According to the literature, two choices are suggested for the equations driven by the TV-SNP parameters, $\nu_{i,t}$. The first one is as a function of lags of the standardized returns z_t, and the second one as a function of lags of ε_t. We stick to the former since we are indeed modeling the higher-order moments for the distribution of z_t.

where $\kappa_{0t} = \mu_t + a_t\sigma_t$, $\kappa_{1t} = b_t\sigma_t$ and $Q_t^{-1}(\alpha) = \inf\{x \,|Q(x\,|I_{t-1}) \geq \alpha\}$ is the conditional α-quantile with $q(\cdot\,|I_{t-1})$ as pdf. Since $q(\cdot\,|I_{t-1})$ nests the $N(0,1)$ distribution for $\nu_{1,t} = \nu_{2,t} = 0$, then $Q_t^{-1}(\alpha) = \Phi^{-1}(\alpha)$. Once we have obtained $r_{\alpha,t}$ in (26), the ES is easily computed.

Proposition 3. *Let r_t be the asset return with pdf in* (23) *and let $r_{\alpha,t}$ be the conditional α-quantile in* (26)*, then*

$$\begin{aligned} ES_t(\alpha) &= E_{t-1}(r_t\,|r_t \leq r_{\alpha,t}) \\ &= \kappa_{0t} + \kappa_{1,t}E_{t-1}\left(x_t\,\middle|x_t \leq r^*_{\alpha,t}\right) \\ &= \kappa_{0t} + \frac{\kappa_{1t}}{\alpha}\xi_{1t}\left(r^*_{\alpha,t}\right) \end{aligned} \tag{27}$$

where $r^*_{\alpha,t} = (r_{\alpha,t} - \kappa_{0t})/\kappa_{1t}$ and $\xi_{1t}(u) = \int_{-\infty}^{u} xq(x\,|I_{t-1})\,dx$ is the conditional version of $\xi_1(u)$ given in section ii) in Appendix 1.

Proof. *See section iii) in Appendix 1.*

2.6 Conditional partial moments

The lower partial moments (LPMs), see Fishburn (1977), measure risk by negative deviations of the asset return in relation to a return threshold, θ. The conditional LPM of order m where the asset return r_t follows a TV-SNP process, i.e. with pdf given by (23), is defined as

$$LPM_t(\theta, m) = \int_{-\infty}^{\theta} (\theta - r_t)^m f(r_t\,|I_{t-1})dr_t. \tag{28}$$

The conditional upper partial moment (UPM) of order m and return threshold θ is defined as

$$UPM_t(\theta, m) = \int_{\theta}^{\infty} (r_t - \theta)^m f(r_t\,|I_{t-1})dr_t. \tag{29}$$

In this paper we are only interested in the LPMs of orders 1 and 2, that is, $LPM_t(\theta, 1)$ and $LPM_t(\theta, 2)$ in (28). Respecting the UPMs, we use $UPM_t(\theta, 1)$ in (29). Next, we obtain the closed-form expressions of these two LPMs and the UPM.

Proposition 4. *Let r_t be the asset return driven by the TV-SNP process with pdf in* (23)*, then*

$$LPM_t(\theta, 1) = (\theta - \kappa_{0t})\,\xi_{0t}(\theta^*_t) - \kappa_{1t}\xi_{1t}(\theta^*_t), \tag{30}$$

$$LPM_t(\theta, 2) = (\theta - \kappa_{0t})^2\,\xi_{0t}(\theta^*_t) + \left(\kappa_{1t}^2 - 2\theta\kappa_{1t}\right)\xi_{1t}(\theta^*_t) + \kappa_{1t}^2\xi_{2t}(\theta^*_t), \tag{31}$$

$$UPM_t(\theta, 1) = \mu_t - \theta + LPM_t(\theta, 1), \tag{32}$$

where $\kappa_{0t} = \mu_t + a_t\sigma_t$, $\kappa_{1t} = b_t\sigma_t$, $\theta^*_t = (\theta - \kappa_{0t})/\kappa_{1t}$, $\xi_{0t}(u) = Q(u\,|I_{t-1})$ and $\xi_{jt}(u) = \int_{-\infty}^{u} x^j q(x\,|I_{t-1})\,dx$ is the conditional version of $\xi_j(u)$ given in section ii) of Appendix 1.

Proof. *It is obtained straightforwardly.*

3 Empirical application

3.1 Dataset and summary statistics

We start analyzing the time-series behavior of six stock indexes and four FX rates. The data employed were daily percentage log returns, which were computed as $r_t = 100 \log (P_t/P_{t-1})$ from series $\{P_t\}_{t=1}^{T}$ of daily closing prices for Nasdaq, TAIEX, Bovespa, CAC, DAX and Eurostoxx stock indexes; and pound sterling to euro (UK-EU), Japanese yen to U.S. dollar (JAP-US), Canadian dollar to U.S. dollar (CAN-US) and pound sterling to U.S. dollar (UK-US) FX rates. All of the price series were sampled from September 28, 1997 to September 27, 2017 to obtain a total of $T = 5,219$ observations. The data were downloaded from Datastream.

Table 1 exhibits summary statistics of both stock-index and FX returns series. We can see that all series of stock-index returns exhibit much higher standard deviations than FX ones. The same goes for the maximum and minimum values, all series of stock-index returns exhibit both the lowest minimum and highest maximum values. Clearly, all series show high leptokurtosis with the UK-US returns presenting the largest kurtosis (14.7), and the TAIEX the smallest (6.79). The degree of unconditional skewness is heterogeneous among the series, with the largest positive and negative (in absolute value) skewness corresponding to the UK-US (0.57) and JAP-US (-0.47) returns, respectively, and the smallest (in absolute value) to the Nasdaq (-0.06). The UK-EU and UK-US returns are positively skewed whilst the rest of the series present negative skewness. In all cases, the Jarque-Bera (J-B) test rejects the null of normality, motivating the use of our SNP distribution.

Table 1: Summary statistics for daily percent stock-index and foreign-exchange log returns

	Nasdaq	TAIEX	Bovespa	CAC	DAX	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
Mean	0.02	0.00	0.01	0.01	0.02	0.01	0.00	0.00	0.00	0.00
Std. dev.	1.59	1.49	2.47	1.57	1.62	1.59	0.51	0.68	0.56	0.58
Min	-10.16	-11.34	-17.96	-11.74	-9.60	-11.10	-2.67	-6.58	-5.05	-4.47
Max	13.25	8.26	18.01	12.14	12.37	11.96	6.22	3.71	4.34	8.31
Skewness	-0.06	-0.22	-0.24	-0.06	-0.11	-0.08	0.50	-0.47	-0.10	0.57
Kurtosis	8.59	6.79	9.35	9.02	7.54	8.40	9.07	8.10	8.44	14.75

This table presents the summary statistics for stock-index and FX daily percent log returns from September 29, 1997 to September 27, 2017 ($T = 5,218$ obs.).

3.2 Estimation results

The parameters of the SNP models we considered in this analysis were estimated using maximum likelihood (ML) according to equation (24). To account for the small structure in the return conditional means, we filtered the r_t series with autoregressive processes of different orders for the conditional mean, μ_t. Since the estimations, under either filtered $(r_t - \widehat{\mu}_t)$ or non-filtered returns, yielded rather similar results, we decided to report only the results for non-filtered data. Therefore, we assume a constant mean equation for r_t, i.e.

$\mu_t = \mu$. The stylized features of returns volatility were described through the GJR process in (11). For the SNP distribution driven by the innovations, z_t, in equation (1), we take different specifications for $\nu_{i,t}$ nested in the GTV-SNP driven by equations (21) and (22).

3.2.1 C-SNP-GJR model

Table 2 presents the estimation results under the constant SNP (C-SNP) model, which is the restricted GTV-SNP with $\varphi_{1i} = \varphi_{2i}^{+} = \varphi_{2i}^{-} = \varphi_{3i} = 0$. The unconditional mean parameter, μ, is not significant for any of the return series, except for the DAX returns for which it is significant at the five per cent level. The parameter estimates of the conditional variance equation (11) show that, for all series, the model correctly captures the asset returns stylized features of (i) clustering and high persistence in volatility, and (ii) asymmetric response of volatility to positive and negative shocks. Indeed, both persistence, β, and asymmetry, $\alpha_1^{-} \neq \alpha_1^{+}$, in (11) are not altered either through the Normal or the different SNP specifications besides C-SNP (available upon request). For all series, the C-SNP parameters, denoted as φ_{01} and φ_{02}, are significant at least at the one per cent level. The last row of Table 2 presents the likelihood ratio (LR) test for the Normal versus C-SNP models. The LR test null is rejected for all series at any reasonable significance level, which shows that the SNP distribution significantly improves the Normal in fitting the skewness and leptokurtosis levels exhibited in the empirical returns distributions in Table 1.

The point estimation, not reported, of $E(c_t)$ in (14) to check the condition for the existence of the unconditional second moment of ε_t is always lower than one and hence, $\sigma_\varepsilon^2 < \infty$ in (13). In all series, the unconditional standard deviations implied by model C-SNP, i.e. $\sigma_{C\text{-}SNP} = \sigma_\varepsilon$, are very close to the sample ones. For instance, the estimated $\sigma_{C\text{-}SNP}$ is equal to 2.46 and the sample standard deviation is 2.47 for Bovespa. The condition for the existence of the unconditional fourth moment of ε_t, given by $E\left(c_t^2\right) < 1$ in (17), seems to be satisfied according to the point estimates of $E\left(c_t^2\right)$ for all FX return series, whilst among the stock index series only Bovespa satisfies this condition. Furthermore, to be more precise, we obtained a confidence interval at 95% for $E\left(c_t^2\right)$ by using a numerical derivative based on the Delta method. The null hypothesis $H_o : E\left(c_t^2\right) = 1$ is only rejected for JAP-US in favor of the alternative $H_a : E\left(c_t^2\right) < 1$. For the remaining assets, $H_o : E\left(c_t^2\right) = 1$ is not rejected in favor of either the one-sided alternative $H_a : E\left(c_t^2\right) < 1$ or $H_a : E\left(c_t^2\right) > 1$. In short, the infinite unconditional kurtosis will not be rejected in most cases. Stronger evidence of possibly infinite fourth unconditional moments will be discussed later in Table 3 through the analysis of confidence intervals from tail-index estimates. Finally, the confidence interval at 95% for $E\left(c_t^2\right)$ is derived in Appendix 3 by using now a closed-form expression for the Delta method under the Normal-GJR model since their partial derivatives are obtained immediately unlike the C-SNP-GJR model. The results from the confidence intervals are similar in both models.

3.2.2 Further results on the existence of unconditional moments

Next, we study the finiteness of the first moments of asset returns so as to understand some previous results in Table 2 related to the existence of the unconditional variance and kurtosis through the estimates of $\sigma_{C\text{-}SNP}$ and $E\left(c_t^2\right)$, respectively. In particular, we analyze the heavy-tailedness property by estimating the

tail indexes. A number of studies have concluded that inference on the tail index ζ using the popular Hill's estimator suffers from several problems. Table 3 exhibits the robust estimates of ζ from the OLS approach using the log-log rank-size (RS) regressions from Gabaix and Ibragimov (2011) and denoted as $\widehat{\zeta}_{RS}$. The standard error of $\widehat{\zeta}_{RS}$ is equal to $\widehat{\zeta}_{RS}\sqrt{2/n}$ where $n < T$ with T as the total number of observations. We set for n the most commonly used values for extreme observations, i.e. $n = k_0 \times T$ with $k_0 = 5\%, 10\%$. The corresponding 95% confidence interval for ζ is

$$\left(\widehat{\zeta}_{RS} - 1.96 \times \widehat{\zeta}_{RS}\sqrt{2/n}\ ,\ \widehat{\zeta}_{RS} + 1.96 \times \widehat{\zeta}_{RS}\sqrt{2/n}\right). \tag{33}$$

The point estimates $\widehat{\zeta}_{RS}$ for the returns of our stock indexes and FX lie between 3.20 and 4.31 for the 5% truncation level, and between 3.06 and 3.86 for the 10%. The conclusions drawn from this table are the following. First, the null hypothesis $H_o : \zeta = j$ where $j = 1, 2$ is rejected in favor of the one-sided alternative $H_a : \zeta > j$ for the two truncation levels and all assets, so the the first moment and the variance are finite for the ten assets. Second, it is verified that $H_o : \zeta = 3$ is not rejected but $H_o : \zeta = 3$ is rejected in favor of $H_a : \zeta > 3$ for (i) TAIEX, UK-EU, CAN-US and UK-US at 10% truncation level, and (ii) TAIEX, DAX, UK-EU, JAP-US and CAN-US at 5% truncation level. Hence, there is a strong evidence of finite skewness for most of the FX series. Third, the null hypothesis $H_o : \zeta = 4$ is rejected in most cases for the 10% truncation level except for TAIEX and UK-EU, while it is only rejected at 5% truncation level for Bovespa and UK-US. Fourth, for those cases where $H_o : \zeta = 4$ is not rejected, we cannot reject $H_o : \zeta = 4$ in favor of $H_a : \zeta > 4$. Fifth, those assets rejecting $H_o : \zeta = 4$ at 10% truncation level also do not reject $E\left(c_t^2\right) = 1$ under C-SNP-GJR except for JAP-US; see the confidence intervals in Table 2. In short, there seems to be more evidence in favor of infinite kurtosis ($\zeta < 4$) for most asset returns.[7]

3.2.3 TV-SNP-GJR models

Finally, we implement the following TV-SNP models nested in the GTV-SNP, defined according to (21) and (22): (i) the asymmetric linear SNP in (22) with and without AR(1), denoted as AL1-SNP (restricted GTV with $\varphi_{3i} = 0$) and AL0-SNP (restricted GTV with $\varphi_{1i} = \varphi_{3i} = 0$), respectively; (ii) the transition specification in (22) with and without AR(1), denoted as T1-SNP (non restricted GTV) and T0-SNP (restricted GTV with $\varphi_{1i} = 0$), respectively. Table 4 reports the Akaike information criterion (AIC) corresponding to the four candidate TV-SNP models for all series and also, the parameter estimates for the TV-SNP-GJR with the best fit according to AIC. First, we find that the parameter estimates of the GJR model remain similar in magnitude to those of the C-SNP and their statistical significance does not seem to be affected, so they are not reported in the table. Second, the AL1-SNP provides a better fit for TAIEX, Bovespa, and UK-US series; AL0-SNP for Eurostoxx, T0-SNP for CAN-US and T1-SNP for the rest.[8]

[7] We do not obtain the implicit tail index ζ from equation $E(c_t^{\zeta}) = 1$ since it is beyond the scope of this paper. Our empirical analysis follows the approach, among others, in Gu and Ibragimov (2018), Ankudinov, Ibragimov and Lebedev (2017) and Ibragimov, Ibragimov and Kattuman (2003).

[8] We also employed the Schwarz information criterion (SIC), which penalizes model complexity more heavily compared to AIC, and found that T1-SNP provides better fit for Nasdaq, whilst AL0-SNP is selected for the other series. These results are not included but are available upon request.

Table 2: C-SNP-GJR model estimation results

	Nasdaq	TAIEX	Bovespa	CAC	DAX	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
μ	0.034	0.024	0.031	0.019	0.034**	0.016	0.001	0.004	0.004	-0.003
	(0.015)	(0.015)	(0.027)	(0.016)	(0.017)	(0.016)	(0.006)	(0.008)	(0.006)	(0.007)
α_0	0.013***	0.007**	0.126***	0.030***	0.029***	0.031***	0.001**	0.005***	0.001***	0.003**
	(0.004)	(0.003)	(0.035)	(0.007)	(0.007)	(0.008)	(0.001)	(0.002)	(0.001)	(0.001)
β	0.926***	0.953***	0.899***	0.908***	0.919***	0.912***	0.957***	0.951***	0.953***	0.945***
	(0.011)	(0.009)	(0.015)	(0.013)	(0.009)	(0.012)	(0.008)	(0.009)	(0.005)	(0.014)
α_1^+	0.010***	0.021***	0.029***	0.016**	0.012	0.012*	0.037***	0.032***	0.052***	0.053***
	(0.008)	(0.006)	(0.008)	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)	(0.007)	(0.012)
α_1^-	0.125***	0.065***	0.125***	0.140***	0.122***	0.136***	0.043***	0.043***	0.038***	0.041**
	(0.018)	(0.012)	(0.019)	(0.019)	(0.014)	(0.019)	(0.013)	(0.010)	(0.005)	(0.018)
φ_{01}	0.487***	0.579***	0.598***	0.553***	0.547***	0.585***	0.647***	0.671***	0.596***	0.649***
	(0.041)	(0.027)	(0.043)	(0.056)	(0.054)	(0.059)	(0.041)	(0.029)	(0.045)	(0.038)
φ_{02}	0.245***	0.287***	0.280***	0.223***	0.222***	0.234***	0.235***	0.335***	0.217***	0.265***
	(0.018)	(0.019)	(0.028)	(0.031)	(0.029)	(0.035)	(0.026)	(0.024)	(0.029)	(0.028)
$\sigma_{C\text{-}SNP}$	1.767	1.979	2.456	1.673	1.626	1.636	0.583	0.683	0.613	0.586
$E\left(c_t^2\right)$	1.021	1.007	0.995	1.015	1.007	1.012	0.998	0.984	0.999	0.991
$95\%CI$	(0.99,1.04)	(0.99,1.02)	(0.97,1.02)	(0.99,1.04)	(0.98,1.03)	(0.98,1.04)	(0.98,1.01)	(0.97,0.99)	(0.99,1.01)	(0.97,1.01)
$LL_{C\text{-}SNP}$	-8469.4	-8689.0	-11307.8	-8823.9	-9004.3	-8881.6	-3480.8	-4992.6	-3676.3	-4025.5
LL_N	-8550.6	-8769.6	-11386.1	-8862.8	-9043.7	-8920.5	-3493.7	-5103.1	-3691.7	-4053.4
LR	162.41***	161.27***	156.68***	77.74***	78.85***	77.74***	25.92***	220.94***	30.71***	55.83***

Model: $r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t(\boldsymbol{\theta})\, z_t, \quad \sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \alpha_1^+(\varepsilon_{t-1}^+)^2 + \alpha_1^-(\varepsilon_{t-1}^-)^2, \;\; z_t \sim i.i.d.\; g(z_t;\nu), \;\; \nu = (\varphi_{01}, \varphi_{02})\,.$

This table presents ML estimates of the parameters of the C-SNP-GJR model (conditional skewness and kurtosis are constant) for stock index and FX rate percent log returns ($T = 5,218$ obs.). Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (***) indicates significance at 1% level; (**) indicates significance at 5% level and (*) indicates significance at 10% level. $\sigma_{C\text{-}SNP}$ denotes the implied unconditional standard deviation obtained from the C-SNP-GJR model estimates. Let $E\left(c_t^2\right) < 1$ be the condition for the existence of the unconditional fourth moment. The confidence interval at 95% for $E\left(c_t^2\right)$ is obtained by using a numerical Delta method $LR = 2(LL_{C\text{-}SNP} - LL_N)$ denotes likelihood ratio statistic where $LL_{C\text{-}SNP}$ and LL_N give log-likelihood values (constant terms included) of C-SNP-GJR and Normal-GJR respectively. In this case, LR is χ^2 asymptotically distributed with two degrees of freedom. The critical values of the χ_2^2 for significance levels of 1%, 5% and 10% are 9.21, 5.99 and 4.61, respectively.

Table 3: Tail index estimates for daily percent stock-index and foreign-exchange log returns

	Nasdaq	TAIEX	Bovespa	CAC	DAX	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
					Panel 1: Truncation 10%					
$\widehat{\zeta}_{RS}$	3.24	3.60	3.06	3.16	3.33	3.21	3.86	3.33	3.48	3.50
$s.e.\,(\widehat{\zeta}_{RS})$	0.20	0.22	0.19	0.20	0.21	0.20	0.24	0.21	0.22	0.22
$95\%\,CI_{RS}$	$(2.84, 3.63)$	$(3.17, 4.04)$	$(2.69, 3.43)$	$(2.77, 3.54)$	$(2.92, 3.73)$	$(2.82, 3.60)$	$(3.39, 4.32)$	$(2.93, 3.74)$	$(3.06, 3.91)$	$(3.07, 3.92)$
					Panel 2: Truncation 5%					
$\widehat{\zeta}_{RS}$	3.58	4.31	3.20	3.52	3.75	3.64	4.08	3.65	3.74	3.39
$s.e.\,(\widehat{\zeta}_{RS})$	0.31	0.38	0.28	0.31	0.33	0.32	0.36	0.32	0.33	0.30
$95\%\,CI_{RS}$	$(2.96, 4.19)$	$(3.57, 5.06)$	$(2.65, 3.75)$	$(2.91, 4.12)$	$(3.11, 4.40)$	$(3.02, 4.27)$	$(3.38, 4.78)$	$(3.02, 4.28)$	$(3.10, 4.39)$	$(2.81, 3.97)$

$\widehat{\zeta}_{RS}$ denotes the tail-index estimate using the log-log rank-size (RS) regression in Gabaix-Ibragimov (2011). The corresponding standard error $s.e.\,(\widehat{\zeta}_{RS})$ equals $\sqrt{2/n}\;\widehat{\zeta}_{RS}$ where $n = k_0 \times T$ such that k_0 is the percent truncation size ($k_0 = 5\%, 10\%$) and $T = 5,218$ is the total sample. Finally, $95\%\,CI_{RS}$ denotes the confidence level at 95% for the tail index ζ in (33).

Table 4: GTV-SNP-GJR estimation results

	Nasdaq	TAIEX	Bovespa	CAC	DAX	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
φ_{01}	0.132	0.160***	0.081***	0.594***	0.613***	0.462***	0.498***	0.742***	0.583***	0.068
	(0.096)	(0.063)	(0.041)	(0.094)	(0.131)	(0.080)	(0.076)	(0.148)	(0.052)	(0.051)
φ_{11}	0.780***	0.620***	0.793***	-0.273***	-0.289***		0.295***	-0.216***		0.845***
	(0.195)	(0.115)	(0.081)	(0.108)	(0.122)		(0.110)	(0.166)		(0.079)
φ_{21}^{+}	0.003	0.152***	0.109***	0.191***	0.042	0.216***	-0.003	0.006	-0.156	0.016
	(0.063)	(0.046)	(0.035)	(0.106)	(0.066)	(0.064)	(0.004)	(0.013)	(0.103)	(0.025)
φ_{21}^{-}	0.070	-0.024	-0.023	0.058*	0.003	-0.061	0.047	-0.018	0.013	-0.064*
	(0.168)	(0.032)	(0.027)	(0.023)	(0.004)	(0.058)	(0.056)	(0.016)	(0.020)	(0.034)
φ_{31}	0.186			0.566	3.123***		2.157	2.520	-0.638***	
	(0.788)			(0.955)	(5.412)		(2.996)	(2.381)	(0.179)	
φ_{02}	0.036***	0.071***	0.052***	0.227***	0.168***	0.129***	0.045*	0.134***	0.226***	0.047
	(0.015)	(0.033)	(0.023)	(0.048)	(0.042)	(0.046)	(0.025)	(0.059)	(0.041)	(0.032)
φ_{12}	0.865***	0.505***	0.592***	-0.332***	-0.078		0.859***	0.336*		0.566***
	(0.132)	(0.128)	(0.063)	(0.113)	(0.074)		(0.087)	(0.188)		(0.104)
φ_{22}^{+}	0.093	0.186***	0.197***	0.056	0.015***	0.241***	-0.001	0.055	-0.125	0.047
	(0.145)	(0.038)	(0.033)	(0.072)	(0.006)	(0.048)	(0.001)	(0.046)	(0.085)	(0.031)
φ_{22}^{-}	0.110***	-0.050*	-0.031	-0.005	-0.001	-0.063	0.002***	-0.165***	0.130***	-0.126***
	(0.053)	(0.029)	(0.027)	(0.009)	(0.001)	(0.046)	(0.001)	(0.043)	(0.061)	(0.040)
φ_{32}	-0.281***			1.394	11.389***		9.900*	-0.110	-0.568***	
	(0.129)			(2.255)	(4.154)		(5.923)	(0.110)	(0.127)	
AIC_{T1}	3.2381‡	3.3294	4.3299	3.3816‡	3.4462‡	3.4039	1.3358‡	1.9138‡	1.4127	1.5439
AIC_{T0}	3.2449	3.3312	4.3324	3.3817	3.4479	3.4036	1.3373	1.9145	1.4122‡	1.5454
AIC_{AL1}	3.2419	3.3292‡	4.3297‡	3.3821	3.4481	3.4033	1.3360	1.9154	1.4136	1.5437‡
AIC_{AL0}	3.2448	3.3305	4.3319	3.3819	3.4480	3.4029‡	1.3367	1.9153	1.4132	1.5450

Model: $r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t(\boldsymbol{\theta}) z_t, \quad \sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \alpha_1^+ (z_{t-1}^+)^2 + \alpha_1^- (z_{t-1}^-)^2, \quad z_t \sim g(z_t; \boldsymbol{\nu}_t), \quad \boldsymbol{\nu}_t = (\nu_{1,t}, \nu_{2,t})$.
GTV-SNP: $\nu_{i,t} = \varphi_{0i} + \varphi_{1i}\nu_{i,t-1} + \varphi_{2i}^+(1 + \varphi_{3i}|z_{t-1}|)z_{t-1}^+ + \varphi_{2i}^-(1 + \varphi_{3i}|z_{t-1}|)z_{t-1}^-$.

This table presents ML estimates of the parameters of different TV-SNP models nested in GTV-SNP for stock index and FX rate percent log returns ($T = 5,218$ obs.). Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (***) indicates significance at 1% level; (**) indicates significance at 5% level and (*) indicates significance at 10% level. The parameter estimates exhibited correspond to the selected model under the Akaike information criterion (AIC). The lower value of AIC_j (indicated by the symbol ‡), the closer is the fit of model j to the empirical data. The models, denoted as j, are the following: $AL0 : (\varphi_{1i} = \varphi_{3i} = 0)$, $AL1 : (\varphi_{3i} = 0)$, $T0 : (\varphi_{1i} = 0)$ and $T1$ is the GTV model (no parameter restrictions).

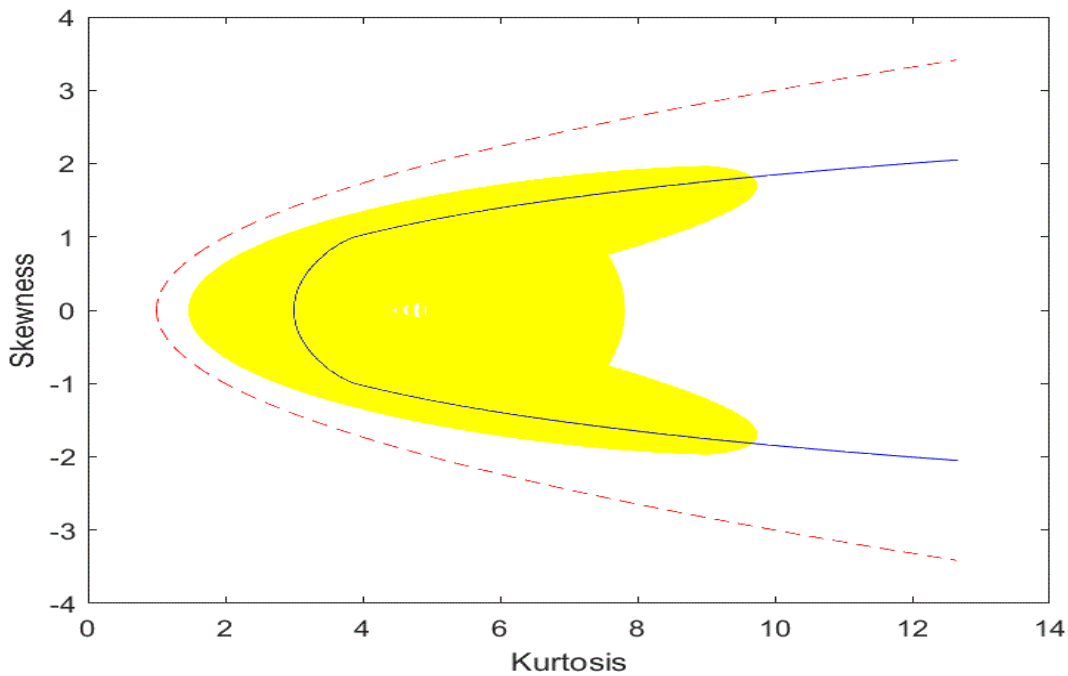
3.3 Comparative analysis

This section provides, for modeling the innovation z_t in (1), a comparison of the C-SNP distribution and the popular ST distribution of Hansen (1994) with parameters $\delta \in (-1, 1)$ and $\upsilon > 2$ that control for skewness and kurtosis, respectively.[9]

3.3.1 Skewness-kurtosis frontiers

We examine the differences of these two densities through the regions of skewness and kurtosis they can generate. Figure 1 provides the envelope of all the combinations of skewness and kurtosis for ST and SNP densities, which are symmetric respecting the x-axis (kurtosis). For the sake of neatness, the ST envelope is plotted up to a suitable level of kurtosis so that its scale is compatible with the SNP envelope (frontier of the yellow area). The figure shows the differences in the levels of skewness and kurtosis these distributions can achieve. First, if we consider levels of kurtosis from three to about ten (positive excess kurtosis), the SNP clearly admits a higher range of skewness than the ST. Note that the sample kurtosis for most returns series in Table 1 are within that range. Second, although the ST can generate larger kurtosis values than our two-parameter SNP distribution, it is worth remarking that the SNP can produce tailored levels of kurtosis through adjusting the truncation order of its polynomial series expansion, as shown in Figure 1 in León, Mencía and Sentana (2009). Third, the ST cannot generate kurtosis below 3 (negative excess kurtosis). This might be restrictive for modeling asset returns over longer time horizons since kurtosis tend to decrease as the return horizon increases. Fourth, if we consider the kurtosis value of 3, the ST only allows for a zero-valued skewness (standard Normal), whilst the SNP allows for a range of possible skewness values.

Figure 1: Regions of skewness and kurtosis



This figure exhibits the regions of skewness and kurtosis for SNP (yellow area) and ST (blue-dashed line) densities. The region enveloped by the red-dashed line contains all possible combinations of skewness and kurtosis.

[9] We thank the referee for pointing out this analysis of comparing distributions.

3.3.2 In-sample analysis

We use Vuong's likelihood ratio test (LRV) (1989) for the nonnested ST versus C-SNP models. The LRV statistic is based on the null hypothesis of being the two candidate models equally close to the true specification, and it is defined as

$$T^{-1/2}\left(LL_{ST} - LL_{C\text{-}SNP}\right)/\widehat{\omega}_T \xrightarrow{d} N\left(0,1\right), \tag{34}$$

such that LL_j denotes the log-likelihood value for model j, and $\widehat{\omega}_T$ is the sample standard deviation obtained in the usual manner, i.e.

$$\widehat{\omega}_T^2 = T^{-1}\sum_{t=1}^{T}\left(\ln\left(l_t^{ST}/l_t^{C\text{-}SNP}\right)\right)^2 - \left(T^{-1}\sum_{t=1}^{T}\ln(l_t^{ST}/l_t^{C\text{-}SNP})\right)^2, \tag{35}$$

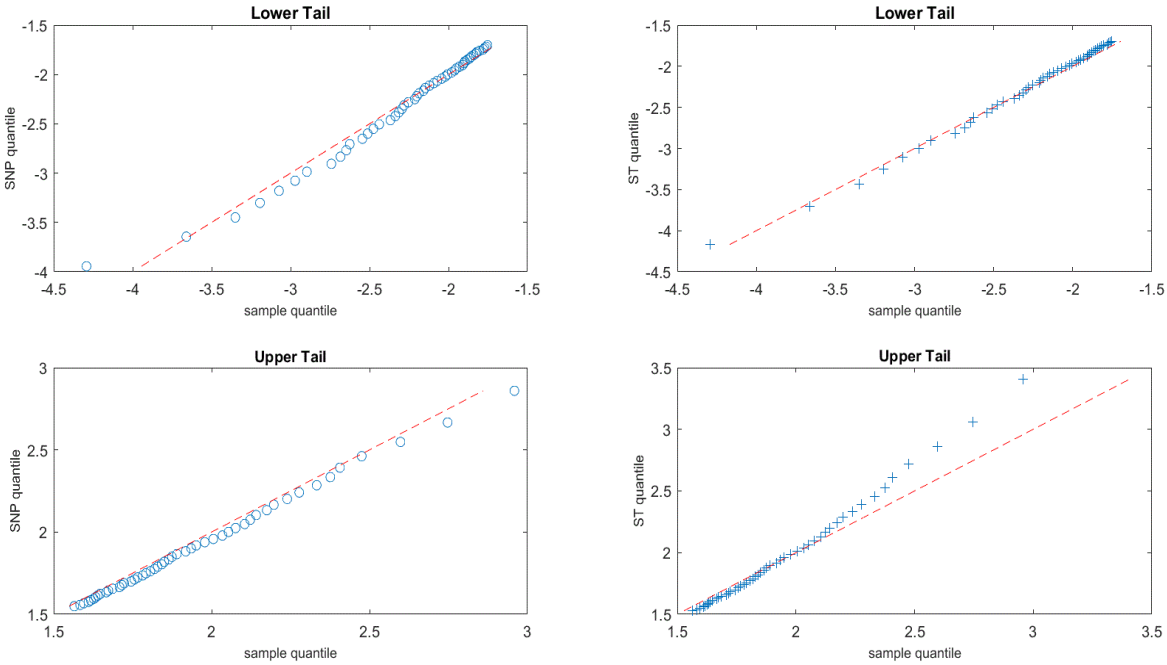
where l_t^j denotes the log-likelihood corresponding to a particular observation t for model j. We consider here the whole sample of $T = 5,218$ observations. Table 5 reports the parameter estimates of the ST model. Similar to C-SNP in Table 2, the ST parameters are significant except δ, which is not significant for CAN-US and UK-US. Thus, we can reject a symmetric distribution for their innovations z_t. Our LRV results are mixed as follows: (i) for Nasdaq and Bovespa the C-SNP provides better fit, (ii) the null of no difference is not rejected for TAIEX, CAC and DAX; (iii) for Eurostoxx, UK-EU, JAP-US, CAN-US there is a greater evidence (and less for UK-US) that ST provides a better fit than C-SNP.

In order to study more exhaustively the differences between the SNP and the ST, we have performed an analysis for their fit of the return distribution tails. To isolate the effect of the distribution, we have applied a two-stage ML estimation procedure. In the first stage we obtain the standardized residuals from a Normal-GJR model, which are employed in the second stage to estimate the parameters of each density.[10] Figure 2 presents a comparison of each density theoretical quantiles with the sample standardized residuals ones for both distribution tails: left or lower tail (quantiles from 0.001 to 0.05) and right or upper tail (quantiles from 0.95 to 0.999). The results of this analysis do not throw a clear-cut better model for all series. For instance, for Nasdaq we find that the SNP seems to provide a better fit for the right tail whilst for the left tail the fit is rather similar. So, as a result, more evidence is found in favor of the SNP as best candidate. For CAN-US returns, both SNP and ST fits are rather similar for the left tail whilst for the right tail the ST performs better, which suggests the ST provides a better fit. These findings are in line with our Vuong test results. Finally, Table A4 (in Appendix 4) exhibits both SNP and ST theoretical quantiles for the specific levels of 1, 5, 95 and 99 per cent, extracted from Figure 1. The results show small differences in the quantiles from both distributions. In particular, the ST systematically overestimates the SNP quantiles for 99% level.

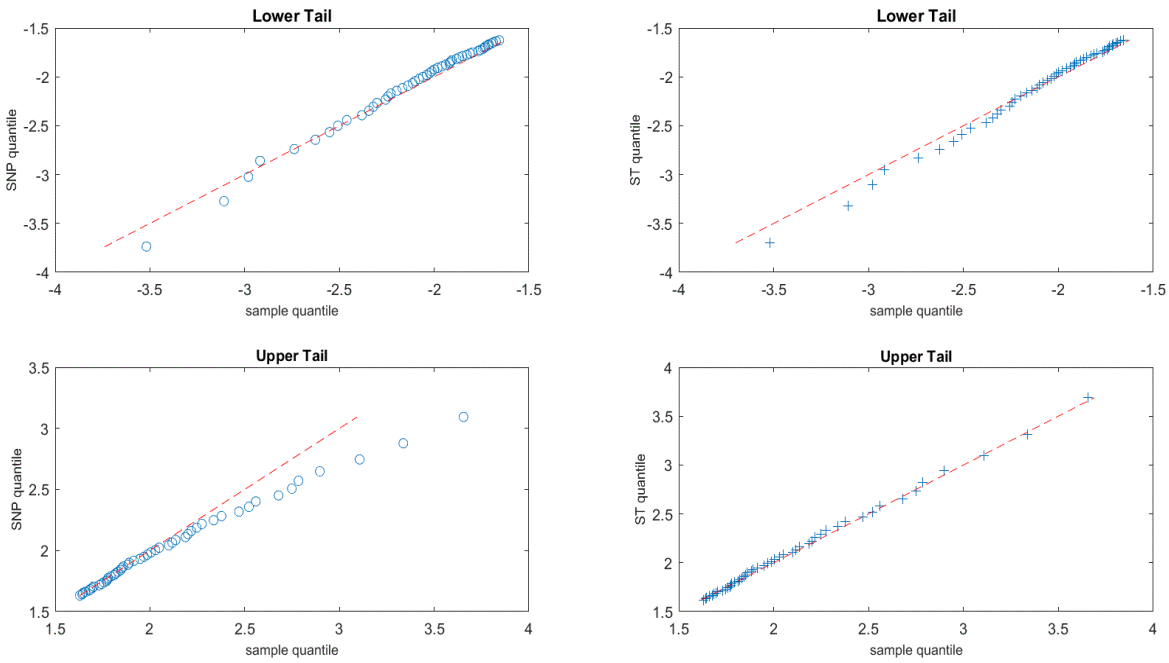
[10] The parameter estimates from the second stage are rather similar to those from the one-stage ML in Tables 2 and 5.

Figure 2: Distribution tail fit analysis

Nasdaq standardized returns



CAN-US standardized returns



This figure provides SNP and ST theoretical quantiles versus sample standardized return quantiles for both distribution tails: lower tail (quantiles from 0.001 to 0.05) and upper tail (quantiles from 0.95 to 0.999). The theoretical quantiles correspond to the parameter estimates obtained through a two-stage ML procedure. Series: Nasdaq and CAN-US standardized returns ($T = 5,218$ obs.).

3.3.3 Backtesting

For the out-of-sample analysis, we implement the backtesting approach of Escanciano and Olmo (2010) for VaR and Du and Escanciano (2017) for ES. We also study the performance for long and short trading positions which are related, respectively, to the left and right tails of the return distribution; see, e.g., Giot and Laurent (2003) for a similar analysis. We are interested in both the unconditional and conditional backtests for VaR and ES.

The backtest implementation involves the first T - N observations for the first in-sample window and the OOS period of length $N = 1,000$ using a constant-sized rolling window. We use two-step estimation procedure as shown, among others, in Zhu and Galbraith (2011) and Komunjer (2007). In the first stage, the mean and GJR parameters are estimated by quasi-maximum likelihood (QML). Then, the SNP and ST density parameters are obtained by ML using the estimations of the standardized residuals, z_t, from the first stage. We have done this for all asset return series presented above under several coverage levels (denoted as α): 1%, 2.5%, 5% and 10%. The one-day-ahead VaR for the α-quantile is given by

$$VaR_t(\alpha) = \kappa_{0,t} + \kappa_{1,t} Q^{-1}(\alpha), \tag{36}$$

where $\kappa_{0,t} = \mu + a\sigma_t$ and $\kappa_{1,t} = b\sigma_t$. Let

$$h_t(\alpha) = \mathbf{1}(r_t < VaR_t(\alpha)) \tag{37}$$

denote the violation or hit variable. We obtain the quadratic loss function,[11] which incorporates the exception magnitude and provides useful information to discriminate among similar models in terms of the unconditional coverage criterion. Thus,

$$QL_t(\alpha) = (r_t - VaR_t(\alpha))^2 \times h_t(\alpha). \tag{38}$$

We estimate the sample averages VIOL and MSE corresponding, respectively, to the daily violations in (37) and the daily quadratic losses in (38) for the OOS period of N days, i.e.

$$VIOL(\alpha) = \frac{1}{N}\sum_{t=1}^{N} h_t(\alpha), \qquad MSE(\alpha) = \frac{1}{N}\sum_{t=1}^{N} QL_t(\alpha). \tag{39}$$

Backtesting VaR The probability $P(r_t < VaR_t(\alpha)|I_{t-1}) = \alpha$ suggests that violations are Bernoulli variables with mean α and hence, the centered violations $\{h_t(\alpha) - \alpha\}_{t=1}^{\infty}$ follow a martingale difference sequence (MDS) that implies the zero mean property and its uncorrelation. Testing MDS leads to the unconditional and conditional backtests initially proposed by Kupiec (1995) and Christoffersen (1998),

[11] For a comparison of VaR models under different loss functions, see Abad, Benito and López (2015).

respectively. The null hypothesis for the unconditional backtest, $H_{0,U} : E\left[h_t(\alpha)\right] = \alpha$, corresponds to the following sample test statistic:

$$U_{VaR}\left(\alpha\right) = \frac{\sqrt{N}\left(\overline{h}(\alpha) - \alpha\right)}{\sqrt{\alpha\left(1-\alpha\right)}} \overset{a}{\sim} N\left(0,1\right), \tag{40}$$

where $\overline{h}(\alpha)$ is the sample average of $\left\{\widehat{h}_t\left(\alpha\right)\right\}_{t=1}^{N}$ such that $\widehat{h}_t\left(\alpha\right) = \mathbf{1}\left(\widehat{u}_t \leq \alpha\right)$ with $\widehat{u}_t$ as the estimation of $u_t = F\left(r_t\left|I_{t-1}\right.\right)$ in (25). To test the null hypothesis for the conditional backtest, $H_{0,C} : E\left[h_t(\alpha) - \alpha\left|I_{t-1}\right.\right] = 0$, we implement the approach by Escanciano and Olmo (2010) based on the Box-Pierce (BP) test statistic defined as $C_{VaR}(m) = N\sum_{i=1}^{m}\widehat{\rho}_j^2 \overset{a}{\sim} \chi_m^2$, which is asymptotically a chi-square distribution with m degrees of freedom such that $\hat{\rho}_j$ is the j-th lag of the sample autocorrelation given by $\hat{\rho}_j = \hat{\gamma}_j/\hat{\gamma}_0$ with $\hat{\gamma}_j = \frac{1}{N-j}\sum_{t=1+j}^{N}\left(\widehat{h}_t\left(\alpha\right) - \alpha\right)\left(\widehat{h}_{t-j}\left(\alpha\right) - \alpha\right)$.

Backtesting ES The ES backtest by Du and Escanciano (2017) is based on the notion of cumulative violations (CV) defined as $\mathcal{H}_t\left(\alpha\right) = \int_0^{\alpha} h_t\left(u\right)du$, which accumulates the violations across the tail distribution. Note that $h_t\left(u\right) = \mathbf{1}\left(r_t < VaR_t\left(u\right)\right) = \mathbf{1}\left(u_t \leq u\right)$, then $\mathcal{H}_t\left(\alpha\right)$ can be rewritten as $\mathcal{H}_t\left(\alpha\right) = \left(1 - u_t/\alpha\right)\mathbf{1}\left(u_t \leq \alpha\right)$. This equation provides a better insight of the notion of CV since it measures the distance of the returns from the corresponding α-quantile in (36) for the violations. It is also verified that $\{\mathcal{H}_t(\alpha) - \alpha/2\}_{t=1}^{\infty}$ follows a MDS. The null hypothesis for the unconditional backtest is given by $H_{0,U} : E\left[\mathcal{H}_t(\alpha)\right] = \alpha/2$ and the related sample test statistic is obtained as

$$U_{ES}\left(\alpha\right) = \frac{\sqrt{N}\left(\overline{\mathcal{H}}(\alpha) - \alpha/2\right)}{\sqrt{\alpha\left(1/3-\alpha/4\right)}} \overset{a}{\sim} N\left(0,1\right), \tag{41}$$

where $\overline{\mathcal{H}}(\alpha)$ is the mean of $\left\{\widehat{\mathcal{H}}_t\left(\alpha\right)\right\}_{t=1}^{N}$ such that $\widehat{\mathcal{H}}_t\left(\alpha\right) = \left(1 - \widehat{u}_t/\alpha\right)\mathbf{1}\left(\widehat{u}_t \leq \alpha\right)$. The null hypothesis for the conditional backtest is $H_{0,C} : E\left[\mathcal{H}_t(\alpha) - \alpha/2\left|I_{t-1}\right.\right] = 0$ with BP as the test statistic where $\hat{\gamma}_j$ is now obtained as $\hat{\gamma}_j = \frac{1}{N-j}\sum_{t=1+j}^{N}\left(\widehat{\mathcal{H}}_t\left(\alpha\right) - \alpha/2\right)\left(\widehat{\mathcal{H}}_{t-j}\left(\alpha\right) - \alpha/2\right)$.

Backtesting results Table 6 exhibits a descriptive analysis of VaR violations obtained from C-SNP-GJR and ST-GJR models. The columns of VIOL and MSE correspond to the equations in (39) with $N = 1,000$ and different coverage levels for long (Panel 1) and short (Panel 2) positions. We can conclude that both models provide rather similar performance for long and short trading positions in terms of VIOL. Respecting the MSE, we can observe that (i) for long position, C-SNP is better (lower MSE) than ST regardless the level of α for Nasdaq and TAIEX; (ii) for short position, ST is always better than C-SNP at levels of $\alpha = 1\%$, 2.5%; and (iii) for both positions, C-SNP is always better than ST at level $\alpha = 10\%$.

Table 7 reports the p-values for the unconditional backtesting of VaR and ES associated with the sample test statistics of (40) and (41), respectively. The following conclusions are based on a significance level of

5%. For the long position (Panel 1), there is hardly any difference between the two models. In all cases, both null hypotheses are not rejected except for one case. In particular, for UK-EU the null is always rejected at level $\alpha = 1\%$ for both VaR and ES, nevertheless there is difference between C-SNP and ST at level $\alpha = 2.5\%$ where the null is rejected for the ES with C-SNP density. For the short position (Panel 2), we find more differences respecting the long position. First, for Nasdaq and TAIEX the two null hypotheses are rejected in most situations regardless the model. Second, for CAC, both null hypotheses are never rejected at the $\alpha = 1\%$ level, nevertheless both are always rejected at levels $\alpha = 5\%, 10\%$, irrespective of the model. Third, we always reject both null hypotheses at level $\alpha = 10\%$ for DAX and Eurostoxx. Fourth, for the FX series, both null hypotheses are not rejected in any situation except for CAN-US at the $\alpha = 10\%$ level under ST for VaR. As a result, we can conclude that both models do really perform very similarly in backtesting VaR and ES. Finally, the conditional backtests for VaR and ES yield stronger evidence of similar performance.[12]

In summary, the results of our analysis in this section show that the SNP can be a good alternative density to the ST for modeling asymmetric and heavy-tailed distributions.

4 Equity screening and portfolio selection

Once we have proposed the SNP distribution for modeling asset return innovations and discussed its properties and estimation, we apply the TV-SNP-GJR specification to derive closed-form expressions of alternative conditional performance measures based on the one-sided risk and reward measures obtained in previous sections. These PMs will be appropriate to create portfolios through an equity screening approach as in León et al. (2019).

4.1 Dataset description

We study the performance of portfolios formed from choosing stocks that were constituents of the S&P 100 index in October 2017. The data series used are sampled over the period November 4, 2004 to October 18, 2017, a total of $T = 3,262$ daily percent log return observations. After filtering, we restrict to 90 stocks that continuously belonged to S&P 100 during our sample period. We split the series into two subsamples, one for the in-sample and another for the OOS period. The in-sample period goes from November 4, 2004 to December 7, 2009. We always use a constant-sized rolling window of $1,282$ observations for the in-sample period and, also, when estimating across the OOS period.

[12] The few cases in which the null hypothesis for the conditional backtesting is rejected (at 5% level) are marked in Table 7 with the symbol.† The p-values of the full conditional backtesting analysis are not reported but they are available upon request.

Table 5: Hansen's ST estimation results

	Nasdaq	TAIEX	Bovespa	CAC	DAX	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
υ	8.3521***	5.6698***	7.8554***	9.2800***	9.0422***	9.0975***	8.8033***	5.5519***	10.1232***	8.8229***
	(1.0006)	(0.4684)	(0.8720)	(1.1858)	(1.1376)	(1.1924)	(1.1224)	(0.4266)	(1.3697)	(1.2179)
δ	-0.1343***	-0.0324**	-0.0739***	-0.0858***	-0.0848***	-0.0730***	0.0637***	-0.0373**	-0.0021	0.0148
	(0.0170)	(0.0148)	(0.0179)	(0.0190)	(0.0186)	(0.0192)	(0.0179)	(0.0179)	(0.0192)	(0.0167)
LL_{ST}	-8547.57	-8707.44	-11358.44	-8815.18	-8994.48	-8848.21	-3417.90	-4925.46	-3651.67	-3973.39
LRV	-5.95***	-1.24	-2.42***	0.65	0.65	1.99**	2.22**	3.26***	2.60***	1.61*

This table presents the ML estimates of Hansen's ST parameters where $\delta \in (-1, 1)$ and $\upsilon > 2$ that control for skewness and kurtosis, respectively. Mean and GJR equation estimates for ST are very similar to the reported under C-SNP (Table 2) and Normal models, they are not presented here to avoid duplicity. The data used are stock index and FX rate percent log returns. The sample consists of $T = 5,218$ observations. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (***) indicates significance at 1% level; (**) indicates significance at 5% level and (*) indicates significance at 10% level. LRV denote the nonnested LR test statistic of Vuong (1989) for ST and C-SNP models, i.e. $T^{-1/2}(LL_{ST} - LL_{C\text{-}SNP})/\widehat{\omega}_T \xrightarrow{d} N(0,1)$ under the null (both models are equally close to the true specification) where LL_j denote the log-likelihood value for model j and $\widehat{\omega}_T$ is the standard deviation in (35).

Table 6: Descriptive analysis of violations and MSE

	VIOL		MSE			VIOL		MSE	
α	ST	C-SNP	ST	C-SNP	α	ST	C-SNP	ST	C-SNP
Panel 1: Long									
Nasdaq					TAIEX				
0.01	0.012	0.012	0.0077	0.0072	0.01	0.013	0.012	0.0043	0.0027
0.025	0.026	0.025	0.0163	0.0154	0.025	0.022	0.021	0.0178	0.0160
0.05	0.047	0.046	0.0321	0.0312	0.05	0.041	0.039	0.0375	0.0354
0.1	0.083	0.081	0.0655	0.0647	0.1	0.093	0.083	0.0722	0.0682
Bovespa					CAC				
0.01	0.007	0.006	0.1503	0.1467	0.01	0.009	0.009	0.0462	0.0465
0.025	0.022	0.023	0.1975	0.1976	0.025	0.021	0.022	0.0640	0.0653
0.05	0.047	0.047	0.2670	0.2687	0.05	0.041	0.046	0.0878	0.0893
0.1	0.098	0.095	0.4135	0.4122	0.1	0.092	0.092	0.1339	0.1338
DAX					Eurostoxx				
0.01	0.010	0.011	0.0372	0.0373	0.01	0.008	0.008	0.0574	0.0583
0.025	0.022	0.022	0.0553	0.0564	0.025	0.022	0.022	0.0780	0.0799
0.05	0.043	0.043	0.0790	0.0801	0.05	0.040	0.045	0.1036	0.1053
0.1	0.087	0.086	0.1241	0.1237	0.1	0.092	0.090	0.1514	0.1509
UK-EU					JAP-US				
0.01	0.017	0.018	0.0012	0.0014	0.01	0.011	0.011	0.0055	0.0058
0.025	0.031	0.031	0.0038	0.0037	0.025	0.027	0.028	0.0118	0.0129
0.05	0.053	0.050	0.0080	0.0074	0.05	0.044	0.044	0.0209	0.0212
0.1	0.102	0.097	0.0166	0.0149	0.1	0.085	0.079	0.0364	0.0346
CAN-US					UK-US				
0.01	0.008	0.008	0.0030	0.0032	0.01	0.012	0.014	0.0038	0.0041
0.025	0.020	0.021	0.0051	0.0053	0.025	0.027	0.027	0.0074	0.0076
0.05	0.049	0.049	0.0086	0.0085	0.05	0.044	0.044	0.0124	0.0122
0.1	0.090	0.084	0.0167	0.0159	0.1	0.083	0.075	0.0217	0.0207
Panel 2: Short									
Nasdaq					TAIEX				
0.01	0.001	0.002	0.0001	0.0014	0.01	0.003	0.005	0.0005	0.0014
0.025	0.013	0.014	0.0007	0.0009	0.025	0.012	0.013	0.0030	0.0039
0.05	0.029	0.029	0.0042	0.0040	0.05	0.036	0.035	0.0091	0.0090
0.1	0.084	0.076	0.0165	0.0147	0.1	0.093	0.079	0.0271	0.0233
Bovespa					CAC				
0.01	0.014	0.016	0.0380	0.0501	0.01	0.007	0.008	0.0139	0.0152
0.025	0.024	0.027	0.0887	0.0943	0.025	0.013	0.014	0.0224	0.0225
0.05	0.050	0.048	0.1594	0.1561	0.05	0.035	0.032	0.0346	0.0332
0.1	0.102	0.096	0.2981	0.2798	0.1	0.076	0.070	0.0625	0.0580
DAX					Eurostoxx				
0.01	0.008	0.008	0.0112	0.0124	0.01	0.007	0.008	0.0154	0.0167
0.025	0.021	0.021	0.0190	0.0194	0.025	0.012	0.012	0.0234	0.0237
0.05	0.039	0.036	0.0330	0.0317	0.05	0.039	0.038	0.0350	0.0337
0.1	0.074	0.066	0.0634	0.0591	0.1	0.077	0.076	0.0636	0.0590
UK-EU					JAP-US				
0.01	0.011	0.015	0.0216	0.0235	0.01	0.010	0.021	0.0051	0.0066
0.025	0.028	0.033	0.0266	0.0277	0.025	0.027	0.030	0.0097	0.0104
0.05	0.048	0.050	0.0330	0.0332	0.05	0.052	0.048	0.0166	0.0157
0.1	0.104	0.092	0.0450	0.0436	0.1	0.097	0.087	0.0297	0.0262
CAN-US					UK-US				
0.01	0.009	0.013	0.0019	0.0023	0.01	0.010	0.014	0.0424	0.0445
0.025	0.026	0.027	0.0038	0.0040	0.025	0.025	0.028	0.0489	0.0499
0.05	0.056	0.056	0.0070	0.0069	0.05	0.058	0.058	0.0560	0.0561
0.1	0.121	0.116	0.0151	0.0140	0.1	0.101	0.098	0.0688	0.0675

This Table presents a descriptive analysis of one-day-ahead VaR violations for both C-SNP-GJR and ST-GJR models. Both VIOL and MSE denote, respectively, average violations and mean square error in (39). The coverage level is $\alpha = \{0.01, 0.025, 0.05, 0.1\}$ for both long and short positions in Panels 1 and 2, respectively. The data consists of daily return series from five stock indexes: Nasdaq, TAIEX, Bovespa, CAC, DAX, Eurostoxx, and four FX: UK-EU, JAP-US, CAN-US and UK-US. Total sample: $5,218$ observations from September 29, 1997 to September 27, 2017. Predictions: $1,000$.

Table 7: P-values for unconditional backtesting ES and VaR

	VaR		ES			VaR		ES	
α	ST	C-SNP	ST	C-SNP	α	ST	C-SNP	ST	C-SNP
Panel 1: Long									
Nasdaq					TAIEX				
0.01	0.5250	0.5250	0.3049	0.3613	0.01	0.3404	0.5250	0.4981	0.9013
0.025	0.8395	1.0000	0.5906	0.7869	0.025	0.5434	0.4178	0.7639	0.8450
0.05	0.6634	0.5617	0.6639	0.8521	0.05	0.1916	0.1105	0.5821	0.3399
0.1	0.0731	0.0452	0.3804	0.2953	0.1	0.4606	0.0731	0.2100	0.0707
Bovespa					CAC				
0.01	0.3404	0.2036	0.6822	0.5912	0.01	0.7506	0.7596	0.9867	0.9725
0.025	0.5434	0.6854	0.7320	0.6359	0.025	0.4178	0.5434	0.3757	0.4929
0.05	0.6634	0.6634	0.6063	0.6127	0.05	0.1916	0.5617	0.2960	0.4520
0.1	0.8330	0.5982	0.4852	0.4992	0.1	0.3991	0.3991	0.3192	0.4369
DAX					Eurostoxx				
0.01	1.0000	0.7506	0.9466	0.8279	0.01	0.5250	0.5250	0.8139	0.7801
0.025	0.5434	0.5434	0.5792	0.7153	0.025	0.5434	0.5434	0.4316	0.6417
0.05	0.3098	0.3098	0.3974	0.5223	0.05	0.1468	0.4682	0.2391	0.3987
0.1	0.1706	0.1400	0.1458	0.1795	0.1	0.3991	0.2918	0.2782	0.3804
UK-EU					JAP-US				
0.01	0.0261	0.0110	0.0405	0.0056	0.01	0.7506	0.7506	0.7835	0.6933
0.025	0.2243	0.2243	0.0536	0.0273	0.025	0.6854	0.5434	0.7139	0.4552
0.05	0.6634	1.0000	0.2337	0.2970	0.05	0.3840	0.3840	0.7533	0.9973
0.1	0.8330	0.7518	0.5014	0.9911	0.1	0.1138	0.0269	0.2744	0.2612
CAN-US					UK-US				
0.01	0.5250	0.5250	0.6791	0.9069	0.01	0.5250	0.2036	0.6439	0.4498
0.025	0.3112	0.4178	0.2289	0.3698	0.025	0.6854	0.6854	0.3571	0.2175
0.05	0.8846	0.8846	0.4781	0.5771	0.05	0.3840	0.3840	0.9611	0.9400
0.1	0.2918	0.0917	0.4200	0.3458	0.1	0.0731	$0.0084^{\dagger}$	0.2138	0.1612
Panel 2: Short									
Nasdaq					TAIEX				
0.01	$0.0042^{\dagger}$	0.0110	$0.0123^{\dagger}$	0.0210	0.01	0.0261	0.1120	0.0377	0.2859
0.025	0.0151	0.0259	0.0016	0.0068	0.025	0.0085	0.0151	0.0068	0.0845
0.05	0.0023	0.0023	0.0015	0.0027	0.05	0.0422	0.0295	0.0014	0.0083
0.1	0.0917	0.0114	0.0023	0.0007	0.1	0.4606	0.0269	0.0160	0.0054
Bovespa					CAC				
0.01	0.2036	0.0565	0.0707	0.0003	0.01	0.3404	$0.5250^{\dagger}$	0.5995	0.9249
0.025	0.8395	0.6854	0.6566	0.1933	0.025	0.0151	0.0259	0.0655	0.1101
0.05	1.0000	0.7717	0.9283	0.8175	0.05	0.0295	0.0090	0.0205	0.0172
0.1	0.8330	0.6733	0.8464	0.9332	0.1	0.0114	0.0016	0.0077	0.0019
DAX					Eurostoxx				
0.01	$0.5250^{\dagger}$	$0.5250^{\dagger}$	0.5922	0.7986	0.01	0.3404	0.5250	0.5490	0.8820
0.025	0.4178	0.4178	0.3493	0.6370	0.025	0.0085	0.0085	0.0525	0.0906
0.05	0.1105	0.0422	0.1811	0.2122	0.05	0.1105	0.0817	0.0199	0.0149
0.1	0.0061	0.0003	0.0201	0.0093	0.1	0.0153	0.0114	0.0228	0.0063
UK-EU					JAP-US				
0.01	0.7506	0.1120	0.7722	0.0365	0.01	$1.0000^{\dagger}$	0.0005	0.7174	0.0006
0.025	0.5434	0.1052	0.8718	0.0461	0.025	0.6854	0.3112	0.5265	0.0193
0.05	0.7717	1.0000	0.8511	0.2698	0.05	0.7717	0.7717	0.7282	0.3822
0.1	0.6733	0.3991	0.8902	0.7994	0.1	0.7518	0.1706	0.9344	0.6414
CAN-US					UK-US				
0.01	0.7506	0.3404	0.8799	0.2476	0.01	1.0000	0.2036	0.9944	0.1202
0.025	0.8395	0.6845	0.8825	0.2682	0.025	1.0000	0.5434	0.9799	0.2548
0.05	0.3840	0.3840	0.7402	0.4472	0.05	0.2457	0.2457	0.5641	0.1940
0.1	0.0269	0.0917	0.1016	0.1571	0.1	0.9161	$0.8330^{\dagger}$	0.6027	0.5475

This table reports the p-values of the tests for the VaR and ES unconditional backtesting for both C-SNP-GJR and ST-GJR models. The coverage level is $\alpha = \{0.01, 0.025, 0.05, 0.1\}$ for both long and short positions in Panels 1 and 2, respectively. The symbol † denotes rejection of the null at 5% level for the *conditional* backtest (these p-values, with Box-Pierce test statistic asymptotically distributed χ^2_5, are available upon request). The data consists of daily return series from five stock indexes: Nasdaq, TAIEX, Bovespa, CAC, DAX, Eurostoxx, and four FX: UK-EU, JAP-US, CAN-US and UK-US. $T = 5,218$ observations from September 29, 1997 to September 27, 2017. Predictions: $1,000$.

Table 8 presents some summary statistics of the data analyzed in this section. The top panel presents sample moments only for the in-sample daily percent log returns. The kurtosis coefficients reveal the stock return distributions are highly leptokurtic (median kurtosis is 10.883). In contrast with the stock indexes in Table 1, the skewness of the single stocks is predominantly positive (median skewness is 0.069).

4.2 Model and estimation of individual stock returns

The estimation of the parameters for each stock return series, $r_{j,t}$, is in two stages as in our backtesting procedure in subsection 3.3. For the first stage, we estimate by QML the conditional mean and variance under the specification from Oh and Patton (2017), i.e. the AR(1) - GJR(1,1) model augmented with lagged market (S&P 100) return information for the stock return series, then

$$r_{j,t} = \gamma_{0j} + \gamma_{1j} r_{j,t-1} + \gamma_{mj} r_{m,t-1} + \varepsilon_{j,t}, \qquad \varepsilon_{j,t} = \sigma_{j,t} z_{j,t}, \qquad j = 1, ..., 90 \tag{42}$$

$$\sigma^2_{j,t} = \alpha_{0j} + \beta_j \sigma^2_{j,t-1} + \alpha^+_{1j} \left(\varepsilon^+_{j,t-1}\right)^2 + \alpha^-_{1j} \left(\varepsilon^-_{j,t-1}\right)^2 + \delta^+_{mj} \left(\varepsilon^+_{m,t-1}\right)^2 + \delta^-_{mj} \left(\varepsilon^-_{m,t-1}\right)^2, \tag{43}$$

such that $\varepsilon_{m,t}$ is the demeaned market return. Onwards, we refer to model in (42) - (43) as AR1-GJRA where AR1 and GJRA denote the conditional mean and variance in (42) and (43), respectively. The bottom panel of Table 8 exhibits information on the parameter estimates of the AR1-GJRA for the in-sample period. Our estimates of the mean equation show a small positive AR(1) coefficient, γ_{1j}, that is significant only for 7% of the stocks, and an estimate for the lagged market return parameter, γ_{mj}, that is predominantly negative, larger in magnitude than γ_{1j} and significant for 31% of the stocks. The GJRA parameter estimates show that most stock returns exhibit typical volatility clustering and high persistence in volatility, as well as asymmetric response of volatility to positive and negative news. Furthermore, the GJRA asymmetric response is in average greater to market shocks than to individual ones. We find evidence of leverage effect since the average estimate of α^-_{1j} is higher than that of α^+_{1j}, namely $0.068 > 0.023$. The second stage consists of using the QML standardized returns, i.e. $z_{j,t} = \varepsilon_{j,t}/\sigma_{j,t}$, to estimate by ML the parameters of alternative specifications of the SNP distribution. It can be seen in the bottom panel that the estimated standardized returns are non-Gaussian since the C-SNP parameters φ_{01j} and φ_{02j} are significant for 100% and 98% of the stocks, respectively. In the next subsections, we will only consider for building portfolios the AL0-SNP out of all TV-SNP models we used in section 3. Note that AL0-SNP is the most parsimonious model and a very good candidate according to the SIC for model selection (see footnote 8). The parameter estimates show evidence of larger response of $\nu_{j,2t}$ to positive rather than to negative shocks, whilst the response of $\nu_{j,1t}$ to shocks is more symmetric (see median values).

Table 8: Summary statistics of S&P 100 stocks and estimation results

	Cross-sectional distribution						
	In-sample period: 11/4/2004-12/7/2009						
	Mean	5%	25%	Median	75%	95%	M
Daily obs. 1,282							
Mean	0.010	-0.054	-0.014	0.013	0.029	0.098	
Std. dev.	2.318	1.310	1.697	2.057	2.688	4.105	
Skewness	-0.018	-0.767	-0.213	0.069	0.283	0.837	
Kurtosis	13.936	7.190	8.650	10.883	14.893	24.162	
Conditional mean							
γ_0	0.032	-0.026	0.001	0.024	0.051	0.126	0.03
γ_1	0.010	-0.094	-0.024	0.012	0.052	0.096	0.07
γ_m	-0.114	-0.240	-0.152	-0.126	-0.067	0.018	0.31
Conditional variance							
α_0	0.158	0.016	0.044	0.074	0.126	0.577	0.42
β	0.875	0.724	0.856	0.888	0.916	0.953	0.76
α_1^+	0.023	0.000	0.000	0.013	0.037	0.071	0.11
α_1^-	0.068	0.000	0.026	0.062	0.094	0.182	0.37
δ_m^+	0.035	0.000	0.000	0.000	0.024	0.176	0.09
δ_m^-	0.176	0.020	0.070	0.128	0.203	0.533	0.42
C-SNP							
φ_{01}	0.748	0.406	0.647	0.724	0.807	1.218	1
φ_{02}	0.347	0.170	0.270	0.334	0.398	0.618	0.98
AL0-SNP							
φ_{01}	0.732	0.202	0.614	0.711	0.840	1.260	0.96
φ_{21}^+	0.003	-0.524	-0.185	0.029	0.190	0.459	0.66
φ_{21}^-	-0.019	-0.629	-0.195	0.026	0.182	0.461	0.57
φ_{02}	0.355	0.012	0.246	0.345	0.481	0.628	0.88
φ_{22}^+	-0.054	-0.420	-0.173	-0.052	0.043	0.302	0.53
φ_{22}^-	-0.038	-0.422	-0.147	0.006	0.083	0.200	0.37

Model (AR1-AL0-SNP-GJRA):
$r_{j,t} = \gamma_{0j} + \gamma_{1j} r_{j,t-1} + \gamma_{mj} r_{m,t-1} + \varepsilon_{j,t}, \;\; \varepsilon_{j,t} = \sigma_{j,t}(\boldsymbol{\theta}) z_{j,t}, \;\; j = 1, \ldots, 90, \;\; z_{j,t} \sim g(z_{j,t}; \boldsymbol{\nu}_{j,t}), \;\; \boldsymbol{\nu}_{j,t} = (\nu_{j,1t}, \nu_{j,2t}),$
$\sigma_{j,t}^2 = \alpha_{0j} + \beta_j \sigma_{j,t-1}^2 + \alpha_{1j}^+ \left(\varepsilon_{j,t-1}^+\right)^2 + \alpha_{1j}^- \left(\varepsilon_{j,t-1}^-\right)^2 + \delta_{mj}^+ \left(\varepsilon_{m,t-1}^+\right)^2 + \delta_{mj}^- \left(\varepsilon_{m,t-1}^-\right)^2,$
$\nu_{j,it} = \varphi_{0ij} + \varphi_{2ij}^+ \left(z_{j,t-1}^+\right)^2 + \varphi_{2ij}^- \left(z_{j,t-1}^-\right)^2, \quad i = 1, 2.$

The top panel presents some summary statistics of the in-sample daily log returns of the stocks that constitute the S&P 100 index used in this study. The columns present the mean, median and percentiles from the cross-sectional distribution of the measures listed in the rows. The bottom panel presents the associated cross-sectional analysis for (i) the QML parameter estimates from the AR1-GJRA model for the conditional mean and variance and (ii) the ML parameter estimates for the C-SNP and AL0-SNP specifications, which are listed in the rows. Note that j denotes an individual stock from S&P 100, and M denotes the number of stocks out of 90 (in %) with significant parameter estimates at 5% level.

4.3 Time-varying portfolio selection

Through our constant-sized rolling window, we obtain the estimations of a battery of PMs across the OOS period for each individual stock and setting a zero mean return as the threshold, $\theta = 0$. We compute a total of thirteen conditional PMs, namely: Sharpe ratio (SR), skewness-kurtosis ratio (SKR),[13] Sortino, Omega and Upside potential ratios, as well as VaR ratio (VaRR) and the Rachev or expected tail ratio (ETR) for the levels of α: 1%, 5%, 10% and 20%. These PMs can be seen in detail in Appendix 2.

Next, we explain the steps to construct the different portfolios. First, the last day of each window, we compute all PMs based on the one-day-ahead forecast of the conditional mean, variance and $\nu_{j,it}$ for each stock assuming the AR1-AL0-SNP-GJRA specification. Second, the stocks are ranked on the basis of each PM and then, we select the ten best-ranked stocks to build initially an equally-weighted (EW) portfolio, i.e. $w_{k,t} = 1/10$ where $k = 1, ..., 10$. We keep this portfolio for the next 5 days to then, compute the daily portfolio returns for these five days. Third, by rolling the window every five days, we repeat the previous two steps a total of 396 times and change each time the portfolio composition according to the equity screening from the different PMs. Fourth, we obtain thirteen OOS portfolio return series of $1,980$ daily observations. We label each of these return series according to the selected PM.

We also repeat the above procedure but changing now the rebalancing frequency. So, we estimate each stock return model under the OOS period every 22 days (monthly frequency) and 10 days (biweekly frequency). Thus, these two rebalancing horizons account for 90 and 198 estimations, respectively.

Figure 3 represents the spreads between the cumulative returns on each portfolio and the SR during the OOS period for the three different rebalancing periods. It is exhibited that the size of spreads - notice the scale in the vertical axis - becomes much higher under both SKR and ETRs, except for the ETR (95,5). Negative spreads, displayed the majority of days, are obtained under monthly frequencies in many portfolios. We also find that VaRR portfolios show positive spreads in most cases under biweekly frequency except for the VaRR (80,20) where, surprisingly, the monthly frequency cumulative returns are consistently higher. Finally, unlike the Omega portfolio, we obtain positive spreads under both Sortino and Upside potential portfolios for weekly frequency.

Finally, a similar analysis, not presented here to save space, was carried out for the spread with respect the S&P 100 index returns. Again, the portfolios with the best performance correspond to both ETR and SKR strategies. The monthly rebalancing yields lower performance although rather better than in the case displayed in Figure 1.

[13] The unconditional version of the SKR may be inapplicable in portfolio choice mainly due to the problem of possibly infinite unconditional kurtosis (see Table 3). Nevertheless, this problem can be overcome using the conditional version of the SKR.

4.4 Weighting schemes and robustness analysis

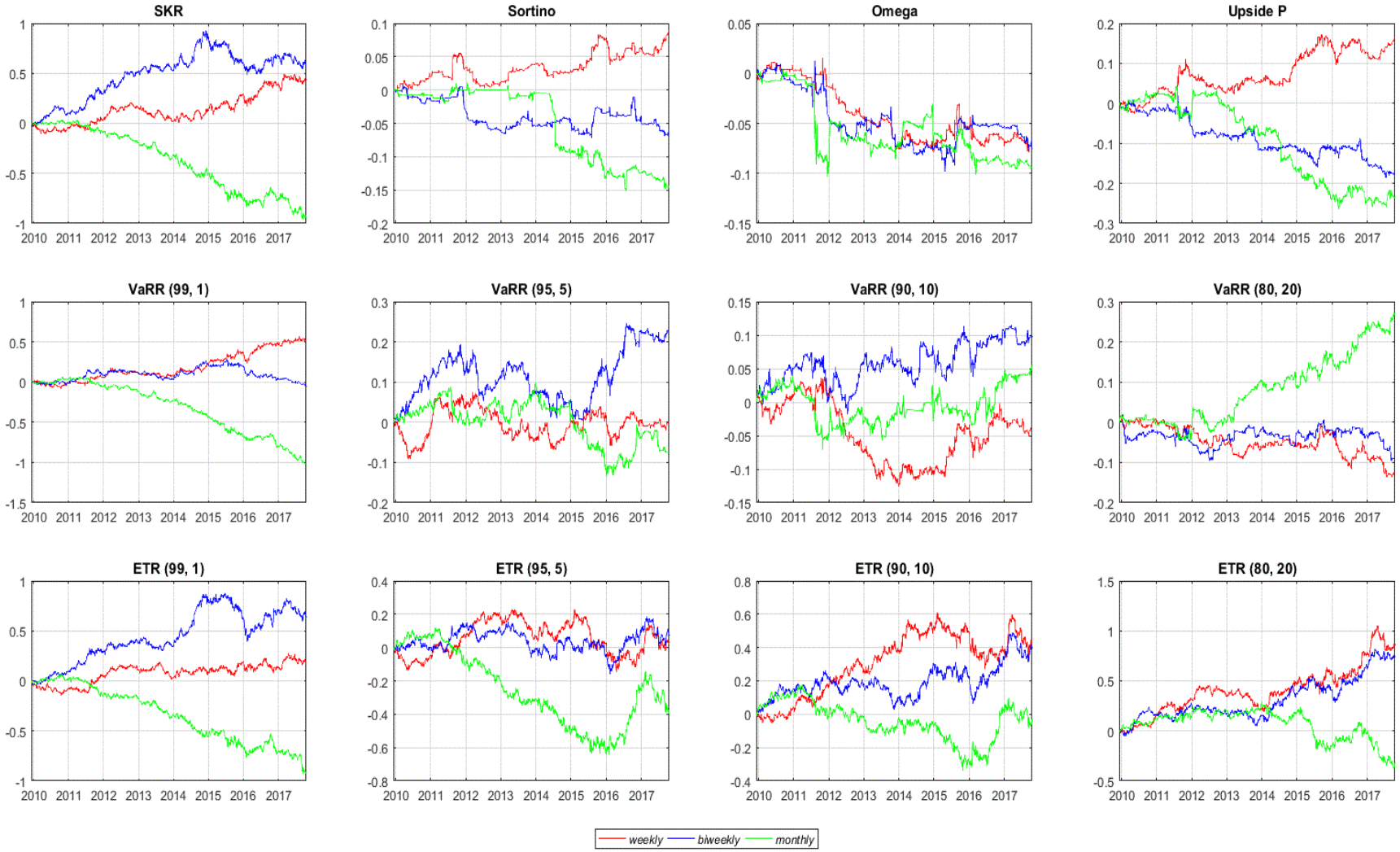
We have previously applied the naive EW portfolio rule. Here, we are interested in the relative portfolio performances, under the PMs used in the previous section, but now adopting different rules to set up portfolio weights. Thus, we consider the following schemes. First, the shortsale-constrained global-minimum-variance (GMV) portfolio, i.e. $\widehat{w}_t = \arg\min w_t' \widehat{\Omega}_t w_t \quad s.t. \; w_t' l = 1$ and $w_t \geq 0$, where $\widehat{\Omega}_t$ is the estimated conditional covariance matrix of order 10 and l is a column vector of ones. Second, the volatility timing (VT) portfolio, i.e. $\widehat{w}_{k,t} = \left(1/\widehat{\sigma}^2_{k,t}\right) / \sum_{k=1}^{10} \left(1/\widehat{\sigma}^2_{k,t}\right)$ where $\widehat{\sigma}^2_{k,t}$ is the estimated conditional variance. Third, the reward-to-risk timing (RRT) portfolio, i.e. $\widehat{w}_{k,t} = \left(\widehat{\mu}^+_{k,t}/\widehat{\sigma}^2_{k,t}\right) / \sum_{k=1}^{10} \left(\widehat{\mu}^+_{k,t}/\widehat{\sigma}^2_{k,t}\right)$ where $\widehat{\mu}^+_{k,t} = \max(\widehat{\mu}_{k,t}, 0)$ with $\widehat{\mu}_{k,t}$ denoting the estimated conditional mean. For more details about these weighting schemes, see Kirby and Ostdiek (2012).

To proceed with the weighting scheme's comparison, we compute the cumulative portfolio daily return spreads for each PM strategy under the GMV, VT and RRT schemes with respect to the EW one. These spread series are exhibited in Figure 4, and only for the weekly rebalancing frequency.[14] Our results show that (i) the RRT scheme overall dominates the rest of the weighting schemes for all PMs consistently across the OOS period except for both ETR (80,20) and, sometimes, VaR (80,20); (ii) the GMV tends to significantly underperform the other schemes for all PMs except for both ETR (80,20) and, sometimes, VaR (80,20); (iii) the VT scheme yields portfolio returns between those obtained under the previous two weighting methods; and (iv) GMV performs less well (negative spread) than the EW portfolio for most PMs. As a result, we show that portfolio performance is significantly sensitive to alternative schemes to the naive diversification. Besides, we find similar results for the SR portfolios which are not displayed in Figure 4.

Finally, as a robustness check although not reported here, we have also provided a comparative analysis of the AR1-AL0-SNP-GJRA model with the HS approach. To do so, we have repeated the exercise presented in the previous section but now using HS to obtain PM portfolio return spreads with respect to SR. For instance, for weekly rebalancing and the EW scheme under either constant-sized rolling or expanding window methods, we find that the weekly portfolio return series in Figure 3 tend to dominate the corresponding HS ones over the OOS period. This finding provides evidence on the superior performance of our parametric model in regard to the HS method.

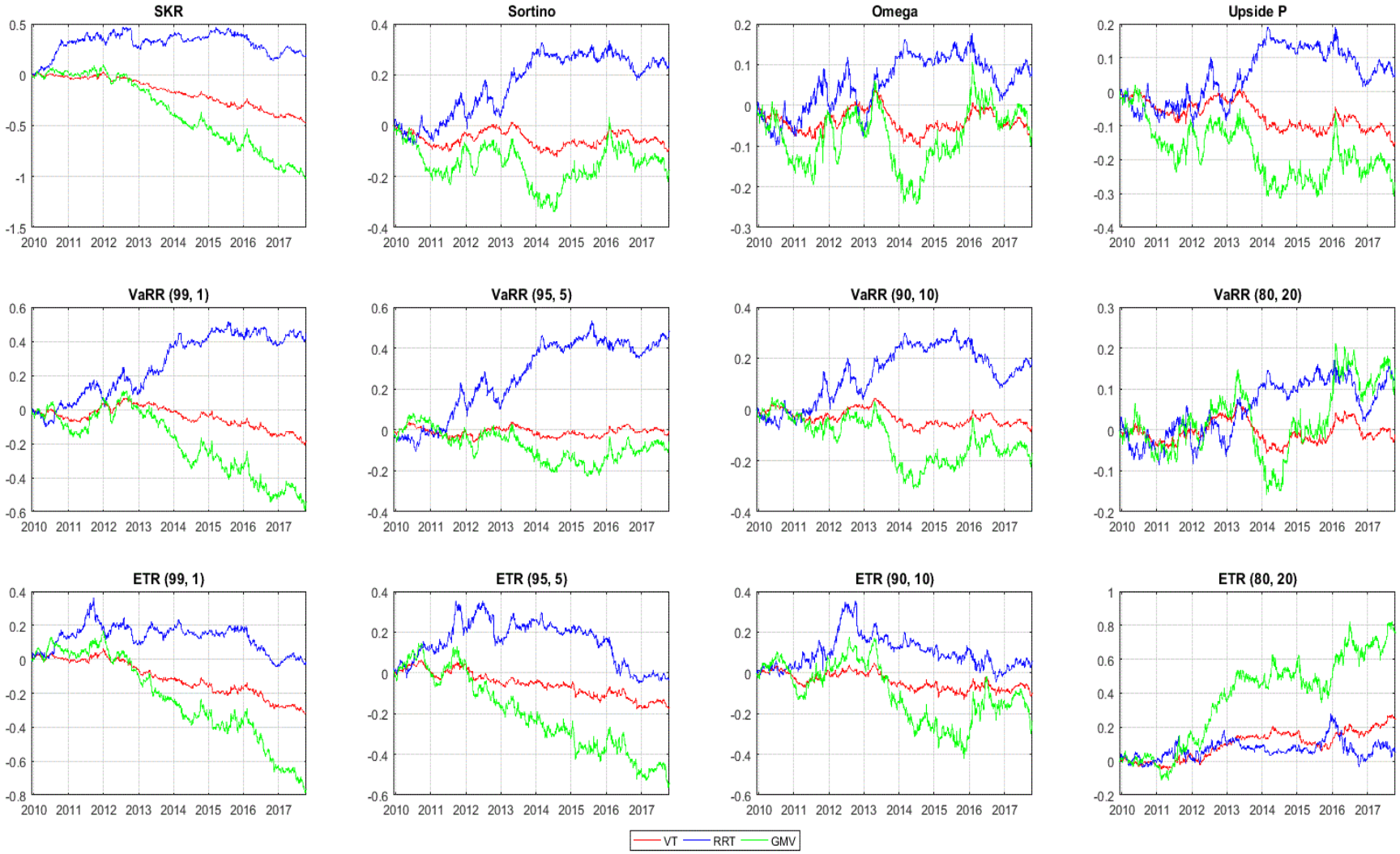
[14] The results for biweekly and monthly rebalancing are not presented but they are available from the authors.

Figure 3: PM portfolio return spread with respect to SR



Plots of PM portfolio return spreads with respect to SR. Rebalancing frequency is weekly, biweekly and monthly over the out-of-sample period to construct equally weighted portfolios of best ten stocks. Model: AR1-AL0-SNP-GJRA.

Figure 4: VT, RRT, GMV portfolio return spreads with respect to the EW portfolio



Plots of portfolio return spreads of GMV, RRT and VT weighting schemes with respect to the naive EW portfolio for each PM. We use AR1-AL0-SNP-GJRA model under weekly rebalancing.

5 Conclusions

This paper develops the SNP density of León, Mencía and Sentana (2009) to incorporate time-varying higher-order moments. First, we derive closed-form expressions of the unconditional variance and kurtosis by assuming a conditional heteroscedastic variance model such as the GJR model, and the SNP density with constant parameters for the innovations of asset returns, i.e. the C-SNP-GJR model. Second, we aim to analyze better the finiteness of the unconditional kurtosis under the heavy-tailed power-law models for financial returns. We obtain expressions for the conditional partial moments, quantiles and expected shortfall (ES) under the C-SNP-GJR. Furthermore, relying on skewness-kurtosis frontiers, in-sample and backtesting analyses, we compare the performance of forecasting VaR and ES between the C-SNP and the popular skewed-t (ST) distribution of Hansen (1994) for modeling the innovations of several return series of stock-indexes and foreign-exchange rates. We estimate robust tail-indexes based on the methodology of Gabaix and Ibragimov (2011) for testing the finiteness of the first unconditional moments. We extend the SNP to time-varying (TV) higher-order moments accounting for nonlinearity and asymmetric effects. Finally, we implement performance measures (PM), based on our closed-form expressions of one-sided reward/risk measures under the TV-SNP-GJR framework, and then carry out an equity screening exercise for ranking stocks from the S&P 100 index depending on the selected PM. We examine the portfolios based on different PM strategies with respect to the benchmark Sharpe ratio portfolio. Our results show that portfolio asset allocation depends critically on the PM considered, as well as on the rebalancing periods and weighting strategies.

Several interesting avenues for further research would be the following. First, obtaining the tail index implied in the heavy-tailed unconditional distribution under the C-SNP-GJR model for testing better the finiteness of the unconditional skewness and kurtosis. See, among others, Zhang, Li and Peng (2019) and Su and Zhou (2014). Second, implementing the generalized autoregressive score for the SNP density as in Thiele (2020). Finally, we could extend the consumption-based asset pricing model with higher-order cumulants in Martin (2013) by assuming the SNP distribution to account for pure higher-order effects on the consumption decision.

Acknowledgements

We are grateful to two anonymous referees and participants at the XXVII Annual Symposium of the Society of Nonlinear Dynamics and Econometrics, University of Alicante Workshop in Quantitative Economics and 36th International Symposium on Forecasting for helpful comments, discussions and suggestions. Of course, the usual caveat applies. Financial support from the Spanish Ministry of Economy and Competitiveness through grants ECO2017-87069-P and ECO2016-75631-P is gratefully acknowledged by the first and second authors, respectively.

Appendix 1: Proofs

i) Obtain the expression of $m_k(\cdot)$:

Let $x \sim N(0,1)$ with $\phi(\cdot)$ and $\Phi(\cdot)$ as pdf and cdf, respectively. We are interested in the moments of the truncated Normal random variable defined as $x \mid x \leq u$ where $u \in \mathbb{R}$. Thus, $m_k(u) = E_\phi\left[x^k \mid x \leq u\right]$ where $k \in \mathbb{N}$. A recursive formula for the truncated normal moments can be obtained as

$$m_k(u) = (k-1)\, m_{k-2}(u) - \frac{u^{k-1}\phi(u)}{\Phi(u)}, \qquad k = 1, 2, 3, ... \tag{44}$$

where $m_{-1}(u) = 0$ and $m_0(u) = 1$. For more details, see Liquet and Nazarathy (2015).

ii) Obtain the expression of $\xi_j(\cdot)$:

Let $\xi_j(u) = \int_{-\infty}^{u} x^j q(x)\, dx$ where $j \in \mathbb{N}$ and $q(\cdot)$ is the pdf in (5), then

$$\begin{aligned}
\xi_j(u) &= \int_{-\infty}^{u} x^j q(x)\, dx \\
&= \sum_{k=0}^{4} \gamma_k \int_{-\infty}^{u} x^j H_k(x)\, \phi(x)\, dx \\
&= \Phi(u) \sum_{i=1}^{5} \eta_i m_{j+i-1}(u), \qquad (45)
\end{aligned}$$

such that $m_k(u)$ is defined in (44) and

$$\begin{array}{lll}
\eta_1 = 1 - \frac{\gamma_2}{\sqrt{2}} + \frac{3\gamma_4}{\sqrt{4!}}, & \eta_2 = \gamma_1 - \frac{3\gamma_3}{\sqrt{3!}}, & \\
\eta_3 = \frac{\gamma_2}{\sqrt{2}} - \frac{6\gamma_4}{\sqrt{4!}}, & \eta_4 = \frac{\gamma_3}{\sqrt{3!}}, & \eta_5 = \frac{\gamma_4}{\sqrt{4!}},
\end{array} \tag{46}$$

where γ_k can be seen in (6). Note that $\xi_0(u) = \Phi(u) \sum_{i=1}^{5} \eta_i m_{i-1}(u)$ is just the SNP cdf given in (10).

iii) Proof of Proposition 3: The expected shortfall, $ES_t(\alpha)$, is obtained as

$$\begin{aligned}
E_{t-1}(r_t \mid r_t \leq r_{\alpha,t}) &= \frac{1}{\alpha} \int_{-\infty}^{r_{\alpha,t}} r_t f(r_t \mid I_{t-1}; \boldsymbol{\psi})\, dr_t \\
&= \frac{1}{\alpha} \int_{-\infty}^{r^*_{\alpha,t}} (\mu_t + a_t\sigma_t + b_t\sigma_t x_t)\, q(x_t \mid I_{t-1})\, dx_t \\
&= \kappa_{0t} + \frac{\kappa_{1t}}{\alpha} \int_{-\infty}^{r^*_{\alpha,t}} x_t q(x_t \mid I_{t-1})\, dx_t \\
&= \kappa_{0t} + \frac{\kappa_{1t}}{\alpha} \xi_{1t}\left(r^*_{\alpha,t}\right) \qquad (47) \\
&= \kappa_{0t} + \frac{\kappa_{1t}}{\alpha} \Phi\left(r^*_{\alpha,t}\right) \sum_{i=1}^{5} \eta_{it} m_{1+i-1}\left(r^*_{\alpha,t}\right) \qquad (48)
\end{aligned}$$

where $r^*_{\alpha,t} = (r_{\alpha,t} - \kappa_{0t})/\kappa_{1t}$, $\kappa_{0t} = \mu_t + a_t\sigma_t$, $\kappa_{1t} = b_t\sigma_t$ and $\xi_{1t}(u)$ in (47) is computed according to $\xi_1(u)$ in (45) such that η_{it} in (48) is given by the expression of η_i in (46) but replacing ν_i with $\nu_{i,t}$, and finally, the expression $m_k(\cdot)$ in (48) is defined in (44).

Appendix 2: Conditional performance measures

Sharpe ratio

We start with the Sharpe (1966, 1994) ratio, denoted as SR, as our benchmark PM. A slightly different version of SR is defined as $(\mu_t - \theta)/\sigma_t$, where θ is the return threshold (e.g., risk-free rate, zero return,...), $\mu_t = E[r_t | I_{t-1}]$ and $\sigma_t = \sqrt{V[r_t | I_{t-1}]}$ denote the conditional mean and volatility of the asset return. A drawback of using SR for ranking assets occurs when the numerator is negative. Israelsen (2005) suggests a modified version to overcome that problem: $SR_t(\theta) = \frac{\mu_t - \theta}{\sigma_t^{sgn(\mu_t - \theta)}}$, where $sgn(z) = z/|z|$ if $z \neq 0$ and $sgn(z) = 0$ if $z = 0$.

Skewness-kurtosis ratio

Watanabe (2006) suggests the simple skewness-kurtosis ratio, $s_{r,t}/k_{r,t}$, where $s_{r,t}$ and $k_{r,t}$ are defined in (20). Again, higher rather than lower ratios are preferred. Since this PM may lead to ranking problems when the numerator becomes negative, we propose a modified version based on Israelsen's idea. Hereafter, our conditional skewness-kurtosis ratio is $SKR_t = \frac{s_{r,t}}{k_{r,t}^{sgn(s_{r,t})}}$.

PMs based on partial moments

First, the Sortino ratio (Sortino and van der Meer, 1991) is the mean excess return, $\mu_t - \theta$, per unit of risk measured by the square root of LPM of order 2 in (31). Note that this PM presents the same problem as the previous measures since the numerator may be negative. As a solution, we propose the conditional modified Sortino ratio: $Sortino_t(\theta) = \frac{\mu_t - \theta}{\left(\sqrt{LPM_t(\theta,2)}\right)^{sgn(\mu_t - \theta)}}$. Second, we use two conditional PMs which are special cases of the Farinelli and Tibiletti (2008) ratio: $FT_t(\theta, q, m) = \frac{\sqrt[q]{UPM_t(\theta,q)}}{\sqrt[m]{LPM_t(\theta,m)}}$, with $q > 0$ and $m > 0$. The higher the value for q, the greater the investor's preference for expected gain, and the higher the value for m the greater the investor's dislike of expected losses. If $q = m = 1$, we have the Omega ratio (Keating and Shadwick, 2002) and for $q = 1$ and $m = 2$, we have the Upside potential ratio (Sortino, van der Meer and Platinga, 1999). These PMs will be represented as $FT_t(\theta, 1, 1)$ and $FT_t(\theta, 1, 2)$, respectively.

PMs based on quantiles

First, the VaRR (Caporin and Lisi, 2011) is defined as the ratio of the upper and lower quantiles given by $VaRR_t(\alpha) = \left|\frac{VaR_t(1-\alpha)}{VaR_t(\alpha)}\right|$, where $VaR_t(\alpha) = Q_t^{-1}(\alpha)$ and $VaR_t(1-\alpha) = Q_t^{-1}(1-\alpha)$ are, respectively, the conditional lower and upper quantiles of r_t in (26). Second, the ETR or Rachev ratio (Biglova et al., 2004) is defined as $ETR_t(\alpha) = \left|\frac{ES_t(-r_t,\alpha)}{ES_t(r_t,\alpha)}\right|$, where $ES_t(r_t, \alpha)$ is just the conditional ES in (27) with r_t as the random variable, while $ES_t(-r_t, \alpha)$ is the same but replacing r_t with $-r_t$. Thus, the numerator is the reward measure corresponding to the right-hand side (gains) of the return distribution, $E_{t-1}(r_t | r_t \geq VaR_t(1-\alpha))$, while the denominator is the risk measure defined as $E_{t-1}(r_t | r_t \leq VaR_t(\alpha))$. Finally, we can rewrite $VaRR_t(\alpha)$ as a quotient of conditional lower quantiles, i.e. $VaR_t(\alpha) = VaR_t(r_t, \alpha)$ and $VaR_t(1-\alpha) = VaR_t(-r_t, \alpha)$.

Appendix 3: Confidence interval of $E\left(c_t^2\right)$

The expression of $E\left(c_t^2\right)$ in (17) under the Normal-GJR model is obtained as

$$E\left(c_t^2\right) = \beta^2 + \beta\alpha_1^+ + \beta\alpha_1^- + \frac{3}{2}\left(\alpha_1^+\right)^2 + \frac{3}{2}\left(\alpha_1^-\right)^2.$$

To shorten, let $E\left(c_t^2\right)$ be denoted as Γ. Note that $\Gamma = \Gamma\left(\boldsymbol{\Upsilon}\right)$ is a nonlinear function of $\boldsymbol{\Upsilon} = \left(\beta, \alpha_1^+, \alpha_1^-\right)'$. The asymptotic distribution of the ML estimate $\widehat{\boldsymbol{\Upsilon}}$ is given by the following result: $\sqrt{T}\left(\widehat{\boldsymbol{\Upsilon}} - \boldsymbol{\Upsilon}\right) \overset{a}{\sim} N\left(0, \Omega_{\boldsymbol{\Upsilon}}\right)$, where $\Omega_{\boldsymbol{\Upsilon}}$ is the asymptotic covariance matrix. By applying the Delta method, the function $\widehat{\Gamma} = \Gamma\left(\widehat{\boldsymbol{\Upsilon}}\right)$ has the following asymptotic distribution:

$$\sqrt{T}\left(\widehat{\Gamma} - \Gamma\right) \overset{a}{\sim} N\left(0, \sigma_\Gamma^2\right), \qquad \sigma_\Gamma^2 = \frac{\partial\Gamma}{\partial\boldsymbol{\Upsilon}'}\Omega_{\boldsymbol{\Upsilon}}\frac{\partial\Gamma}{\partial\boldsymbol{\Upsilon}},$$

such that $\partial\Gamma/\partial\boldsymbol{\Upsilon}'$ is a row vector containing all the partial derivatives of Γ, then

$$\partial\Gamma/\partial\boldsymbol{\Upsilon} = \begin{pmatrix} \partial\Gamma/\partial\beta \\ \partial\Gamma/\partial\alpha^+ \\ \partial\Gamma/\partial\alpha^- \end{pmatrix} = \begin{pmatrix} 2\beta + \alpha_1^+ + \alpha_1^- \\ \beta + 3\alpha_1^+ \\ \beta + 3\alpha_1^- \end{pmatrix}.$$

The following table shows the 95% confidence interval for the true $\Gamma = E\left(c_t^2\right)$ corresponding to the asset returns in our paper:

Table A3: Normal-GJR point-estimates and confidence intervals of $E\left(c_t^2\right)$

	Nasdaq	TAIEX	Bovespa	CAC	DAX
$E\left(c_t^2\right)$	0.996	0.993	0.956	0.996	0.989
$95\%\,CI$	(0.982, 1.010)	(0.981, 1.005)	(0.927, 0.984)	(0.976, 1.016)	(0.971, 1.007)
	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
$E\left(c_t^2\right)$	0.995	0.997	0.983	0.998	0.990
$95\%\,CI$	(0.974, 1.015)	(0.985, 1.008)	(0.963, 1.002)	(0.990, 1.007)	(0.973, 1.007)

Appendix 4: Theoretical quantiles from SNP and ST densities

Table A4: Theoretical quantiles from SNP and ST densities

		Nasdaq	TAIEX	Bovespa	CAC	DAX	Eurostoxx	UK-EU	JAP-US	CAN-US	UK-US
Panel 1: Left tail											
SNP	1%	-2.7676	-2.7689	-2.6701	-2.5310	-2.5450	-2.4830	-2.3259	-2.6548	-2.3925	-2.4030
	5%	-1.7044	-1.6435	-1.6306	-1.6416	-1.6450	-1.6305	-1.6154	-1.5883	-1.6265	-1.6117
ST	1%	-2.6796	-2.6264	-2.6268	-2.6046	-2.6059	-2.5917	-2.3976	-2.6482	-2.4725	-2.4743
	5%	-1.6949	-1.6082	-1.6552	-1.6696	-1.6681	-1.6610	-1.5743	-1.6028	-1.6231	-1.6067
Panel 2: Right tail											
SNP	95%	1.5486	1.5662	1.5822	1.6036	1.6006	1.6131	1.6472	1.5852	1.6322	1.6276
	99%	2.1659	2.2073	2.2355	2.2669	2.2612	2.2856	2.3468	2.2493	2.3182	2.3148
ST	95%	1.5280	1.5618	1.5590	1.5631	1.5631	1.5702	1.6538	1.5511	1.6205	1.6232
	99%	2.2849	2.5088	2.3949	2.3547	2.3590	2.3780	2.5857	2.5150	2.4664	2.5133

This table exhibits theoretical quantiles for 1, 5, 95 and 99 per cent levels from SNP and ST models for stock-index and FX standardized returns ($T = 5,218$ observations).

References

[1] Abad, P., Muela, S. B., and Martín, C. L., 2015. The role of the loss function in value-at-risk comparisons. *The Journal of Risk Model Validation* 9 (1), 1-19.

[2] Anatolyev, S., Petukhov, A., 2016. Uncovering the skewness news impact curve. *Journal of Financial Econometrics* 14 (4), 746-771.

[3] Ankudinov, A., Ibragimov, R., Lebedev, O., 2017. Heavy tails and asymmetry of returns in the Russian stock market. *Emerging Markets Review* 32, 200-219.

[4] Bai, X., Russell, J.R., Tiao, G.C., 2003. Kurtosis of GARCH and stochastic volatility models with non-normal innovations. *Journal of Econometrics*, 114 (2), 349-360.

[5] Bali, T. G., Mo, H., Tang, Y., 2008. The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking and Finance* 32 (2), 269-282.

[6] Bernardi, M., Catania, L., 2018. Portfolio optimisation under flexible dynamic dependence modelling. *Journal of Empirical Finance* 48, 1-18.

[7] Biglova, A., Ortobelli, S., Rachev, S., Stoyanov, S., 2004. Different approaches to risk estimation in portfolio theory. *Journal of Portfolio Management* 31 (1), 103-112.

[8] Boudt, K., Lu, W., Peeters, B., 2015. Higher order comoments of multifactor models and asset allocation. *Finance Research Letters* 13, 225-233.

[9] Caporin, M., Lisi, F., 2011. Comparing and selecting performance measures using rank correlations. *The Open-Access, Open-Assessment E-Journal* 5, 1-34.

[10] Christoffersen, P.F., 1998. Evaluating interval forecasts. *International Economic Review* 39, 841-862.

[11] Du, Z., and Escanciano, J. C., 2017. Backtesting expected shortfall: accounting for tail risk. *Management Science* 63 (4), 940-958.

[12] Escanciano, J. C., Olmo, J., 2010. Backtesting parametric value-at-risk with estimation risk. *Journal of Business and Economic Statistics* 28 (1), 36-51.

[13] Farinelli, S., Tibiletti, L., 2008. Sharpe thinking in asset ranking with one-sided measures. *European Journal of Operational Research* 185, 1542-1547.

[14] Feunou, B., Jahan-Parvar, M.R., Tédongap, R., 2016. Which parametric model for conditional skewness? *European Journal of Finance* 22 (13), 1237-1271.

[15] Fishburn, P.C., 1977. Mean-risk analysis with risk associated with below-target returns. *American Economic Review* 67 (2), 116-126.

[16] Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H.E., 2006. Institutional investors and stock market volatility. *The Quarterly Journal of Economics* 121 (2), 461-504.

[17] Gabaix, X., 2009. Power laws in economics and finance. *Annual Review of Economics* 1 (1), 255-294.

[18] Gabaix, X., Ibragimov, R., 2011. Rank-1/2: a simple way to improve the OLS estimation of tail exponents. *Journal of Business and Economic Statistics* 29, 24-39.

[19] Gallant, A.R., Nychka, D.W., 1987. Semi-nonparametric maximum likelihood estimation. *Econometrica* 55, 363-390.

[20] Giot, P., Laurent, S., 2003. Value-at-risk for long and short trading positions. *Journal of Applied Econometrics* 18 (6), 641-663.

[21] Glosten, R.T., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48 (5), 1779-1801.

[22] Gu, Z., Ibragimov, R., 2018. The "Cubic Law of the Stock Returns" in emerging markets. *Journal of Empirical Finance* 46, 182-190.

[23] Hansen, B.E., 1994. Autoregressive conditional density estimation. *International Economic Review* 35 (3), 705-730.

[24] He, C., Terasvirta, T., 1999. Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173-192.

[25] Ibragimov, M., Ibragimov, R., Kattuman, P., 2013. Emerging markets and heavy tails. *Journal of Banking and Finance* 37 (7), 2546-2559.

[26] Ibragimov, M., Ibragimov, R., Walden, J., 2015. Heavy-tailed distributions and robustness in economics and finance. *Springer*, New York.

[27] Israelsen, C.L., 2005. A refinement of the Sharpe ratio and information ratio. *Journal of Asset Management* 5 (6), 423-427.

[28] Jondeau, E., Rockinger, M., 2001. Gram-Charlier densities. *Journal of Economic Dynamics and Control* 25 (10), 1457-1483.

[29] Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness and kurtosis: Existence, persistence, and comovements. *Journal of Economic Dynamics and Control* 27, 1699-1737.

[30] Jondeau, E., Rockinger, M., 2006. Optimal portfolio allocation under higher moments. *European Financial Management*, 12 (1), 29-55.

[31] Keating, C., Shadwick, W.F., 2002. A universal performance measure. *Journal of Performance Measurement* 6 (3), 59-84.

[32] Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica* 131, 207-248.

[33] Kirby, C., Ostdiek, B., 2012. It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis* 47, 437-467.

[34] Komunjer, I., 2007. Asymmetric power distribution: Theory and applications to risk measurement. *Journal of Applied econometrics* 22 (5), 891-921.

[35] Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 2, 174-184.

[36] Lalancette, S., Simonato, J.G., 2017. The role of the conditional skewness and kurtosis in VIX index valuation. *European Financial Management* 23 (2), 325-354.

[37] León, A., Rubio, G., Serna, G., 2005. Autoregressive conditional volatility, skewness and kurtosis. *Quarterly Review of Economics and Finance* 45, 599-618.

[38] León, A., Mencía, J., Sentana, E., 2009. Parametric properties of semi-nonparametric distribution, with applications to option valuation. *Journal of Business & Economic Statistics* 27 (2), 176-192.

[39] León, A., Moreno, M., 2017. One-sided performance measures under Gram-Charlier distributions. *Journal of Banking & Finance* 74, 38-50.

[40] León, A., Navarro, L., Nieto, B., 2019. Screening rules and portfolio performance. *North American Journal of Economics and Finance* 48, 642-662.

[41] León, A., Ñíguez, T.M., 2020. Modeling asset returns under transformed Gram Charlier. *Westminster Business School*, Working Paper No 1.

[42] Ling, S., McAleer, M., 2002. Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics* 106 (1), 109-117.

[43] Liquet, B., Nazarathy, Y., 2015. A dynamic view to moment matching of truncated distributions. *Statistics and Probability Letters* 104, 87-93.

[44] Martin, I.W., 2013. Consumption-based asset pricing with higher cumulants. *Review of Economic Studies*, 80 (2), 745-773.

[45] Mencía, J., Sentana, E., 2018. Volatility-related exchange traded assets: An econometric investigation. *Journal of Business and Economic Statistics* 36 (4), 599-614.

[46] Mikosch, T., Starica, C., 2000. Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process. *The Annals of Statistics* 28 (5), 1427-1451.

[47] Ñíguez, T.M., Paya, I., Peel, D., Perote, J., 2019. Flexible distribution functions, higher-order preferences and optimal portfolio allocation. *Quantitative Finance* 19 (4), 699-703.

[48] Oh, D.H., Patton, A.J., 2017. Modeling dependence in high dimensions with factor copulas. *Journal of Business and Economic Statistics* 35 (1), 139-154.

[49] Pézier, J., White, A., 2008. The relative merits of alternative investments in passive portfolios. *Journal of Alternative Investments* 10 (4), 37-49.

[50] Sharpe, W.F., 1966. Mutual fund performance. *Journal of Business* 39, 119-138.

[51] Sharpe, W.F., 1994. The Sharpe ratio. *Journal of Portfolio Management* 21, 49-58.

[52] Sortino, F.A., van der Meer, R., 1991. Downside risk. *Journal of Portfolio Management* 17 (4), 27-31.

[53] Sortino, F.A., van der Meer, R., Platinga, A., 1999. The Dutch triangle. *Journal of Portfolio Management* 26, 50-57.

[54] Sun, P., Zhou, C., 2014. Diagnosing the distribution of GARCH innovations. *Journal of Empirical Finance* 29, 287-303.

[55] Theodossiou, P., 1998. Financial data and the skewed generalized t distribution. *Management Science* 44, 1650-1661.

[56] Thiele, S., 2020. Modelling the conditional distribution of financial returns with asymmetric tails. *Journal of Applied Econometrics* 35 (1), 46-60.

[57] Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica,* 307-333.

[58] Watanabe, Y., 2006. Is Sharpe ratio still effective? *Journal of Portfolio Measurement* 11 (1), 55-66.

[59] Xiong, J.X., Idzorek, T. M., 2011. The impact of skewness and fat tails on the asset allocation decision. *Financial Analysts Journal* 67 (2), 23-35.

[60] Zhang, R., Li, C., Peng, L., 2019. Inference for the tail index of a GARCH (1, 1) model and an AR (1) model with ARCH (1) errors. *Econometric Reviews* 38 (2), 151-169.

[61] Zhu, D., Galbraith, J.W., 2011. Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions. *Journal of Empirical Finance* 18, 765-778

[62] Zoia, M.G., Biffi, P., Nicolussi, F., 2018. Value at Risk and Expected Shortfall based on Gram-Charlier-like expansions. *Journal of Banking and Finance* 93, 92-104.