Uncovering the skewness news impact curve

Stanislav Anatolyev*[1], Anton Petukhov[2]

[1]CERGE-EI, Czech Republic and New Economic School, Russia

[2]MIT Sloan School of Management, USA

May 2016

Abstract

We investigate, within flexible semiparametric and parametric frameworks, the shape of the news impact curve (NIC) for the conditional skewness of stock returns, i.e. how past returns affect present skewness. We find that returns may impact skewness in ways that sharply differ from those proposed in earlier literature. The skewness NIC may exhibit sign asymmetry, other types of non-linearity, and even non-monotonicity. In particular, the newly discovered 'rotated S'-shape of the skewness NIC for the S&P500 index is intriguing. We explore, among other things, properties of skewness NIC estimates and conditional density forecasts, the term structure of the skewness NIC, and previously documented approaches to its modeling.

*Corresponding author. Address: Stanislav Anatolyev, CERGE-EI, Politických vězňů 7, 11121 Prague 1, Czech Republic; e-mail stanislav.anatolyev@cerge-ei.cz. We are grateful to the Editor Andrew Patton, the Associate Editor and two anonymous referees for numerous useful suggestions that significantly improved the paper. We have benefited from discussions with Massimiliano Caporin, Peter Reinhard Hansen and Grigory Kosenok. Our thanks also go to seminar audiences at the European University Institute, Einaudi Institute for Economics and Finance, New Economic School, Universita Ca' Foscari Venezia, Universita di Padova, University of Sydney, and at the IAAE 2015 Annual Conference in Thessaloniki.

1 Introduction

It is an established empirical fact that stock returns exhibit asymmetry (see, for example Peiro, 1999; Harvey & Siddique, 1999). Even though the mean-variance paradigm in finance has prevailed for decades, there has been a considerable interest to unconditional and conditional skewness of returns in the context of different financial applications. Kraus & Litzenberger (1976), Simaan (1993), Harvey & Siddique (2000) among others investigate the implications of skewness for the theory and empirics of asset prices. Kane (1982), Athayde & Flôres (2004) develop theoretical models of portfolio choice taking into account the skewness of returns. Patton (2004), Guidolin & Timmermann (2008), and Ghysels et al. (2015), among others, provide empirical evidence of the significance of skewness for portfolio choice. DeMiguel et al. (2013), Jha & Kalimipalli (2010), and Neuberger (2012) investigate the economic importance of option-implied measures of skewness. Recent empirical research (Rehman & Vilkov (2012), Amaya et al. (2015)) finds that firm-level skewness is a significant factor determining heterogeneity in the cross-section of stock returns, and Schneider et al. (2016) argue that firm-level skewness is capable of explaining low-risk anomalies, established in the literature, in the cross-section of stocks. Finally, skewness has received considerable attention in the risk management literature (see, for example, Duffie & Pan, 1997; Wilhelmsson, 2009; Bali et al., 2008; Grigoletto & Lisi, 2009; Engle, 2011).

In his seminal work, Hansen (1994) proposes the autoregressive conditional density (ARCD) framework, where the dynamics of parameters beyond the conditional mean and variance may be conveniently modeled. However, the literature applying it to modeling conditional skewness is pretty limited. The relevant papers use ad-hoc parametric specifications for the *news impact curve (NIC)* of the skewness equation, i.e. how past returns affect present conditional skewness.[1] The voluminous GARCH literature provides lots of versions for the volatility NIC thanks to the abundance of stylized facts related to

[1]The term NIC was introduced by Engle & Ng (1993) in the context of GARCH models for conditional variance. For example, in the GARCH(1,1) model $\sigma_t^2 = \beta_0 + \beta_{-1}\sigma_{t-1}^2 + \beta_1\sigma_{t-1}^2 z_{t-1}^2$ the volatility NIC is the last term $\beta_1\sigma_{t-1}^2 z_{t-1}^2$.

volatility and high informational content of return data about it. Because the former are pretty scarce and the latter is pretty low as far as skewness is concerned, there are few suggestions related to the skewness NIC, and those proposed are weakly motivated.

In this paper we argue that the parametric forms proposed in the previous literature describe the skewness dynamics poorly and demonstrate that the skewness NIC may take on fancier shapes and can, more specifically, exhibit sign asymmetry and other types of non-linearity, even including non-monotonicity. We show this by applying a semiparametric technique that was previously used by Engle & Ng (1993) to study the shape of the volatility NIC, to the skewness equation of the ARCD model with the Skewed Generalized Error (SGE) distribution and EGARCH volatility equation.

Furthermore, within this ARCD framework, we fit a series of parametric specifications for the skewness NIC that allow for its non-monotonicity and various degrees of non-linearity. It turns out that stock return indexes exhibit diverse patterns of the skewness NIC, some symmetric, some asymmetric, some monotone, some non-monotone; very few are accordant with specifications used earlier in the literature. In particular, the S&P500 index reveals an interesting 'rotated S'-shape of the skewness NIC, which is robust to various perturbations, such as the removal of extreme events, mean, volatility or density specification changes, and explicit accounting for jumps.

We also run a series of Monte-Carlo experiments confirming that the in-sample model selection tools we use tend to locate the genuine skewness NIC. The estimates of the NIC parameters are tightly concentrated around the true values, in sharp contrast to unconditional skewness estimates (see Kim & White, 2004). But, as expected, in out-of-sample multiperiod density forecasting experiments (similar to Maheu & McCurdy, 2011) simpler NIC shapes yield a slightly better performance than those that are optimal for describing the skewness dynamics in-sample.

We also estimate the skewness NIC using the 'direct approach' of Harvey & Siddique (1999) and León et al. (2005), who model skewness as a time varying parameter of a peculiar conditional distribution. Several complications arise from this approach, in particular, the presence of complex nonlinear mapping that has to be repeatedly inverted and theoretical bounds for the skewness-kurtosis pair, as well as dubious highly curved skewness NIC. In addition, the empirical results do not strongly support these specifications.

Some of our empirical findings on the shape of the skewness NIC are consistent with and can be explained using the time-varying expected return hypothesis of Pindyck (1984), French et al. (1987) and Campbell & Hentschel (1992). However, some of the newly discovered phenomena, such as the non-monotonicity of the skewness NIC, do not seem to be easily explainable using off-the-shelf financial theories.

Although our main analysis is focused on stock indexes, we also use data on individual stocks. The shape of the skewness NIC turns out to be heterogeneous across stocks as well. Moreover, we explore the term structure of the skewness NIC and skewness itself. We find that skewness of the S&P500 is negative on average and increases in absolute value with the return horizon.

The article is organized as follows: in Section 2 we review the previously developed approaches to modeling skewness and present the results of the semiparametric analysis, which motivates the parametric models introduced in Section 3, which also contains density forecasting and Monte-Carlo experiments. In Section 4, we discuss possible explanations of empirical findings, and Section 5 concludes. The online Appendix available at `is.gd/skewnic` contains auxiliary analyses (robustness of skewness NIC, skewness NIC for individual stocks, term structure of skewness NIC) and technical details on direct models for skewness.

2 A semiparametric look at skewness NIC

In this Section, we describe setups that analyze the dynamics of skewness, and present the framework within which we extract the form of the skewness NIC in a semiparametric fashion. Our interest is in seeing what shape the skewness NIC may take and reconcile this shape with parametric forms encountered in the previous literature. This will also be useful in our further parametric analysis.

Let $\{r_t\}$ be a series of returns. The Hansen (1994) ARCD framework starts from the following representation:

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

where $\mu_t = \mathbb{E}[r_t|I_{t-1}]$, $\sigma_t^2 = \mathbb{E}[(r_t - \mu_t)^2|I_{t-1}]$, and I_t is information available at time t. The conditional mean μ_t is typically a constant, a simple linear autoregression or seasonal dummies. The conditional variance σ_t^2, a proxy for volatility, is assumed to follow some GARCH dynamics. The standardized return z_t has zero mean, unit variance and distributed with a density that allows nonzero skewness. Sometimes researchers assume distributions for ε_t that have a nonzero mean; then the conditional expectation is not μ_t (e.g., Harvey & Siddique, 1999).

2.1 Existing setups for modeling skewness dynamics

There are two approaches to modeling the dynamics of skewness in this framework. In the *direct approach*, the evolution of skewness is defined by an explicit equation for skewness, say s_t. This approach was proposed by Harvey & Siddique (1999). The authors consider the noncentral t distribution for standardized returns and, using an analogy with the GARCH model for volatility, develop the *autoregressive conditional skewness* model. In this model the conditional skewness s_t follows the process

$$s_t = \kappa_0 + \kappa_{-1} s_{t-1} + \kappa_1 \varepsilon_{t-1}^3. \tag{2}$$

This approach was followed by León et al. (2005). The authors utilize a modification of the Gram–Charlier density for standardized returns and extend the model of Harvey

& Siddique (1999) by adding autoregressive dynamics for the conditional kurtosis k_t:

$$s_t = \kappa_0 + \kappa_{-1} s_{t-1} + \kappa_1 z_{t-1}^3, \tag{3}$$

$$k_t = \delta_0 + \delta_{-1} k_{t-1} + \delta_1 z_{t-1}^4. \tag{4}$$

White et al. (2010) also consider this model and compare it with their multi-quantile CAViaR model of time-varying skewness and kurtosis.

The direct approach would be very attractive if not for a couple of major complications that make it rarely used in the literature.[2] First, there are few distributions that have skewness and kurtosis as parameters; usually these depend on deep parameters through a complex nonlinear mapping. This complicates the maximum likelihood procedure as one needs to invert this mapping for each observation at each iteration. Second, there exists a theoretical bound, within which all possible values of the skewness-kurtosis combination must lie (Jondeau & Rockinger, 2003), while the dynamics specified in (2), (3) and (4) do not restrict values of skewness and kurtosis. Harvey & Siddique (1999) perform repeated inversions of the mapping but ignore the second complication, which may lead to the non-invertibility of the mapping for some observations. In addition, in the Harvey & Siddique (1999) model the conditional skewness is closely tied to the conditional mean specification (see Appendix B). León et al. (2005) utilize the Gram–Charlier distribution, which has skewness s_t and kurtosis k_t as parameters. However, to overcome the boundedness problem, they modify the density so that it becomes defined for any pair of s_t and k_t,[3] which compromises the idea of the direct approach because these parameters are no longer skewness and kurtosis with respect to the modified density. Regarding the dynamic equations (2)–(3), the embedded skewness NICs are weakly motivated (solely by analogy with GARCH dynamics) and exhibit dubious highly curved shapes. Below,

[2] To our knowledge, Harvey & Siddique (1999) and León et al. (2005) are the only papers that propose models for the dynamics of conditional skewness utilizing the direct approach; aside from these, Brooks et al. (2005) use the direct approach to model conditional kurtosis.

[3] The formula defining the Gram–Charlier density can yield negative values when values of s_t and k_t that are outside of the bounds are used. A detailed description of the Gram–Charlier density and its modification can be found in Appendices A and B.

we also implement the models of Harvey & Siddique (1999) and León et al. (2005) and compare the results to those arising in our setup.

The more popular *indirect approach* to modeling conditional skewness is to utilize a flexible distribution with a parameter reflecting asymmetry. Typically, such a parameter has to lie between certain levels; if so, it is replaced, via some transformation, with an unrestricted one whose dynamics are modeled instead.

Jondeau & Rockinger (2003) is one of the first papers to use this approach to investigate the dynamics of conditional skewness and kurtosis. The authors use the Hansen (1994) Skewed-t distribution, which has two parameters: asymmetry, say λ, and degrees of freedom, say η. They consider different parametric specifications for the dynamics of λ_t and η_t, some of which are frequently used in the subsequent literature. Hashmi & Tay (2007) extend one of these specifications to explore spillover effects on several Asian stock markets. Bali & Theodossiou (2008) utilize one of the specifications proposed by Jondeau & Rockinger (2003) to model conditional value at risk.

Feunou et al. (2014) take this approach to modeling the dynamics of parameters of the SGE, Skewed-t and Skewed Binormal distributions, and find that SGE shows the best performance. Brännäs & Nordman (2003) model the dynamics of conditional skewness through parameters of the Log Generalized Gamma (LGG) and Pearson type IV (PIV) distributions. The LGG distribution has one parameter associated with skewness. The PIV distribution has two parameters, one of which is closely tied to skewness. However, the authors find that time-varying skewness in the PIV model does not yield significant enhancement compared to the model with a constant asymmetry parameter. Yan (2005) and Grigoletto & Lisi (2009) also exploit the PIV distribution and find strong evidence of time-varying conditional skewness in stock index returns.

We would like to mark out two models considered in the 'indirect approach' literature that are special cases in our analysis. In both, the time varying asymmetry parameter λ_t

is replaced by $\tilde{\lambda}_t$ via the following logistic transformation:

$$\lambda_t = -1 + \frac{2}{1 + \exp(-\tilde{\lambda}_t)}. \tag{5}$$

The first model is considered in Feunou et al. (2014), where standardized returns follow the SGE distribution with two parameters. The one that reflects tail thickness is kept constant; the other one reflects time varying asymmetry:

$$\tilde{\lambda}_t = \kappa_0 + \kappa_{-1}\tilde{\lambda}_{t-1} + \kappa_{0,+}z_{t-1}^{+} + \kappa_{0,-}z_{t-1}^{-}, \tag{6}$$

where $x^{-} = \min(0, x)$, $x^{+} = \max(0, x)$. Among the models considered in the present paper, this model is also included as a special case.

The second model is considered in Jondeau & Rockinger (2003) where the standardized returns are Skewed-t distributed. This distribution also has two parameters: the one reflecting tail thickness is kept constant, and the asymmetry parameter is time varying:

$$\tilde{\lambda}_t = \kappa_0 + \kappa_{-1}\tilde{\lambda}_{t-1} + \kappa_1\varepsilon_{t-1}. \tag{7}$$

The literature suggests two choices for the driving process in the equations for parameters like (6) or (7). The first is standardized return z_t and its lags (e.g., Feunou et al., 2014). Another choice is gross return innovation $\varepsilon_t = \sigma_t z_t$ and its lags (e.g., Jondeau & Rockinger, 2003). We stick to the former choice, because it is a distribution of standardized returns whose skewness is modeled; in addition, all these measures are unitless unlike the gross returns. Now we turn to the description of our model.

2.2 Specification of parametric part

Within the ARCD framework (1) we specify parametric forms for the conditional mean, conditional variance and (shape of) conditional density, while leaving the conditional skewness part incompletely specified. Below, in Section 3, while doing full parametric analysis, we check for the robustness of the skewness equation to pertubations of functional forms for conditional mean, variance, and density.

The conditional mean is set constant; the conditional variance is assumed to follow exponential GARCH (EGARCH) introduced in Nelson (1991):

$$\log \sigma_t^2 = \beta_0 + \beta_{-1} \log \sigma_{t-1}^2 + \beta_1 z_{t-1} + \beta_{|1|} |z_{t-1}|.$$

This is one of the most preferable conditional variance models among asymmetric GARCH in terms of flexibility of dynamics (Rodríguez & Ruiz, 2012) and is empirically one of the most frequently selected for stock market data (Cappiello et al., 2006).

We utilize the Skewed Generalized Error distribution for the standardized residuals. Depending on parameter values, it may exhibit either fat or thin tails and nonzero skewness. It is widely used in the empirical finance literature: Anatolyev & Shakin (2007) use it to model intertrade durations in stock exchanges; Bali & Theodossiou (2008) apply this distribution to model value at risk; Feunou et al. (2014) compare this model with those based on the Skewed-t distribution of Hansen (1994) and Binormal distribution of Feunou et al. (2013) and find that the SGE-based model performs best.

Following the notation of Feunou et al. (2014), one can write the SGE density as

$$f(z, \lambda, \eta) = C \exp\left(-\frac{|z+m|^\eta}{(1+\operatorname{sgn}(z+m)\lambda)^\eta \theta^\eta}\right), \tag{8}$$

$$C = \frac{\eta}{2\theta}\Gamma\left(\frac{1}{\eta}\right)^{-1}, \quad \theta = \Gamma\left(\frac{1}{\eta}\right)^{1/2}\Gamma\left(\frac{3}{\eta}\right)^{-1/2} S(\lambda)^{-1}, \quad m = 2\lambda A S(\lambda)^{-1},$$

$$S(\lambda) = \sqrt{1+3\lambda^2-4A^2\lambda^2}, \quad A = \Gamma\left(\frac{2}{\eta}\right)\Gamma\left(\frac{1}{\eta}\right)^{-1/2}\Gamma\left(\frac{3}{\eta}\right)^{-1/2},$$

where $\Gamma(\cdot)$ is the Gamma function; η and λ are parameters of the distribution, which are subject to restrictions $\eta > 0$, $-1 < \lambda < 1$. Parameters η and λ control tail thickness and asymmetry respectively. Figure 1 graphs the SGE density for different values of parameters. When $\lambda = 0$, the distribution is symmetric, when $\lambda > 0$, it is skewed to the right, and when $\lambda < 0$, it is skewed to the left. When $\lambda = 0$ and $\eta = 2$, it coincides with the standard normal distribution. The skewness and kurtosis for this distribution can be expressed in terms of parameters λ and η as

$$Sk = A_3 - 3m - m^3, \quad Ku = A_4 - 4A_3m + 6m^2 + 3m^4, \tag{9}$$

where $A_3 = 4\lambda(1+\lambda^2)\Gamma(4/\eta)\Gamma(1/\eta)^{-1}\theta^3$ and $A_4 = (1+10\lambda^2+5\lambda^4)\Gamma(5/\eta)\Gamma(1/\eta)^{-1}\theta^4$.

The left panel of Figure 2 depicts the dependence of skewness on parameter λ for different values of η. For reasonable values of η, skewness is an increasing function of λ, hence it is convenient to model skewness dynamics through time-varying parameter λ.

2.3 Skewness dynamics

Now we describe the semiparametric specification of skewness dynamics. While we do it for the case when the returns follow the SGE distribution, this specification can be easily adapted for another asymmetric distribution with a scalar asymmetry parameter (see Appendix A).

The skewness for the SGE distribution is determined by the parameters η and λ that are responsible for asymmetry and tail heaviness. For moderate values of η the skewness monotonically and strongly depends on the asymmetry parameter λ, while the dependence on η is much weaker (see Figure 2). Therefore, we model the skewness dynamics through the dynamics of λ only.[4]

The parameter λ is specified as a function of past standardized returns through the logistic transformation (5), where $\tilde{\lambda}_t$ is a function of $z_{t-1}, z_{t-2}, \ldots$. The transformation ensures that λ_t lies in the interval $(-1, 1)$. The right panel of Figure 2 depicts the mapping from parameter $\tilde{\lambda}$ to the skewness measure for the SGE distribution for a range of typical values of $\tilde{\lambda}$. One can see that not only is skewness its increasing function, but also this function is quite close to a linear one. Thus, modeling $\tilde{\lambda}$ in an additive fashion

[4]We leave it for future research to allow the parameter η to evolve dynamically in order to model the dynamics of both conditional skewness and conditional kurtosis.

is (almost directly) modeling skewness in an additive fashion.

To model the function $\tilde{\lambda}(z_{t-1}, z_{t-2}, \dots)$ we exploit the partially nonparametric technique used in Engle & Ng (1993) to estimate the impact of news on volatility. Apart from a constant, there are two additive terms in this specification: one is an (autoregressive) persistence term, and the other represents the skewness NIC. The inclusion of the persistence term is justified by the evidence of skewness clustering for stock indexes presented in Jondeau & Rockinger (2003). We expect persistence of skewness to be positive although moderate or even weak compared to persistence of volatility, which is typically very high. Indeed, Jondeau & Rockinger (2003) and Feunou et al. (2014) find that the persistence parameter lies in the range 0.4–0.8.

The skewness NIC, i.e. dependence on z_{t-1}, is modeled by a piecewise linear function in a way similar to Engle & Ng (1993). Let m_+, m_- be some nonnegative integers; $\{\tau_i\}_{i=-m_-}^{m_+}$ be a set of real numbers satisfying $\tau_{-m_-} < \tau_{(-m_-+1)} < \cdots < \tau_{(m_+-1)} < \tau_{m_+}$. Define the dynamics of $\tilde{\lambda}_t$ by

$$\tilde{\lambda}_t = \kappa_0 + \kappa_{-1}\tilde{\lambda}_{t-1} + \psi(z_{t-1}), \tag{10}$$

where the skewness NIC function is

$$\psi(z) = \sum_{i=0}^{m_+} \kappa_{i,+}(z - \tau_i)^+ + \sum_{i=0}^{m_-} \kappa_{i,-}(z - \tau_{-i})^-, \tag{11}$$

and $\kappa_0,\ \kappa_{-1},\ \kappa_{i,+}\quad (i = 0, 1, \dots, m_+),\ \kappa_{i,-}\quad (i = 0, 1, \dots, m_-)$ are parameters to be estimated; $x^+ = \max(0, x),\ x^- = \min(0, x)$. When $m_+ = m_- = 0,\ \tau_0 = 0$ this specification coincides with (6). For fixed $\tilde{\lambda}_{t-1}$ the functional form (11) defines $\tilde{\lambda}_t$ as a continuous piecewise linear function of z_{t-1}. For $z_{t-1} \in (\tau_0,\ \tau_1]$ this function has slope $\kappa_{0,+}$; for $z_{t-1} \in (\tau_1,\ \tau_2]$ this function has slope $(\kappa_{0,+} + \kappa_{1,+})$; for $z_{t-1} \in (\tau_i,\ \tau_{i+1}]\quad (i \geq 0)$ this function has slope $\sum_{j=0}^{i} \kappa_{j,+}$; for $z_{t-1} \in (\tau_{-(i+1)},\ \tau_{-i}]\quad (i \geq 0)$ this function has slope $\sum_{j=0}^{i} \kappa_{j,-}$.

The parameters m_+, m_-, $\{\tau_i\}_{i=-m_-}^{m_+}$ are usually chosen by a researcher, although some automated algorithms can be used. Higher values of m_+ and m_- provide higher flexibility for the model, but at the same time this may lead to less precise parameter estimation. In the case of volatility modeling, Engle & Ng (1993) propose two simple methods for choosing $\{\tau_i\}_{i=-m_-}^{m_+}$. The first method is to assign the values of τ_i based on the order statistics of the explanatory variable. The second is to take $\tau_i = i \cdot \sigma$ for $i \in \{-m_-, -(m_- - 1), \ldots, m_+\}$, where σ is the unconditional standard deviation of the explaining variable. The first method is somewhat problematic in the case of modeling conditional skewness because the standardized returns z_t are unobservable, and thus the order statistics of z_t are unavailable; hence, we use the second approach. As z_t's have a standard deviation equal to one, we take $\tau_i = i$ for $i \in \{-m_-, -(m_- - 1), \ldots, m_+\}$.

2.4 Data

Our analysis is focused on stock indexes for the following reason: our estimation sample covers a few decades, during which different factors within a particular company may have changed. Such factors may influence stock return behavior and may not be captured by the model. For example, over the course of two decades, significant changes in the company's capital structure may happen. Aggregate indexes smooth out such changes, and it is more likely that such samples will reveal time-varying skewness in indexes than in individual stocks.

In this Section, we use a few decades of daily logarithmic total returns on the S&P500 and FTSE100 downloaded from `finance.yahoo.com`. Log returns are calculated as $r_t = 100 \log(P_t/P_{t-1})$, where P_t is an index close price at date t adjusted for dividends and splits. The first two columns of Table 1 present time coverage and summary statistics. The series start on different dates, but both end on the last trading day of 2010. Both series have negative sample skewness and very large excess kurtosis, demonstrating two established stylized facts related to higher order moments – stock returns tend to have, in unconditional terms, negative skewness and positive excess kurtosis (see, for example Harvey & Siddique, 1999; Peiro, 1999; Premaratne & Bera, 2000). The top two panels of Figure 3 show the graphs of dynamics of log returns. Both series exhibit volatility clustering and have points that lie very far from typical values.

2.5 Results

Figure 4 presents graphs of the skewness NIC for three semiparametric specifications with zero, one and three knots, conditional on average $\tilde{\lambda}_{t-1}$, for S&P500 (left side) and FTSE100 (right side). Note that the first two are equivalent to a linear and asymmetric linear NIC corresponding to constrained and unconstrained skewness dynamics from Feunou et al. (2014), shown in equation (6), while the last one is more flexible.[5]

It can be seen from the top panels that skewness is positively related to past returns, but the relation for S&P500 is twice as steep as that for FTSE100. When we allow for asymmetric response to negative and positive shocks, we find that the responses are indeed different, but only slightly. In the case of the S&P500, the NIC estimate is practically indistinguishable from a straight line, and the kink at the only zero knot is hardly noticeable; the results are similar to Feunou et al. (2014), where the authors obtain significant positive coefficient estimates of a similar magnitude for $\kappa_{0,-}$ and $\kappa_{0,+}$. In the case of the FTSE100, the kink is more pronounced. However, when more flexibility is allowed, the skewness NIC dramatically differs from that estimated with the 'one knot' model, which fails to capture the nonlinearity, and more concretely, the sharp non-monotonicity of the reaction of skewness to news.[6] Note that at the same time, the kink at zero becomes even less noticeable.

So, the semiparametric estimates bring about a possibility of non-monotonic skewness NIC, particularly negative dependence for bigger standardized returns, especially pronounced for S&P500. This evidence – the abrupt nonlinearity and, especially, non-

[5] We also computed semiparametric estimates with five knots; the results are qualitatively similar to the case of three knots.

[6] For the S&P500, the deviation of the non-monotonic (i.e. with three knots) skewness NIC from linear (no knots) and asymmetric linear (one knot) forms is highly statistically significant: the likelihood ratio statistics (p-values) are equal to 14.6 (0.002) and 12.9 (0.002) respectively. For the FTSE100, the corresponding values are 8.9 (0.031) and 3.9 (0.140).

monotonicity of the skewness NIC – suggests that the previous literature has underestimated the complexity of its shape. We devote the next Section to the development of parametric specifications that would be consistent with the above evidence and allow for richer possibilities than the previous literature has assumed.

3 Parametric analysis of skewness NIC

In this Section, we leave the conditional mean, variance and density specifications as before, only transforming the semiparametric equation for the skewness into a series of parametric specifications of different degrees of flexibility.

3.1 Skewness dynamics

Again, the standardized return z_t is distributed, conditional on the history I_{t-1}, as SGE with constant η and parameter $\tilde{\lambda}_t$ following (10), with the skewness NIC $\psi(z)$ driving the skewness equation. We consider the following parametric specifications of $\psi(z)$:

0. Constant specification: $\kappa_{-1} = 0$, $\psi(z) = 0$ so that $\tilde{\lambda}_t = \kappa_0$.
1. Linear specification:

$$\psi_1(z) = \kappa_1 z.$$

2. Asymmetric linear specification:

$$\psi_2(z) = \kappa_{2-} z \mathbb{I}_{\{z<0\}} + \kappa_{2+} z \mathbb{I}_{\{z>0\}},$$

where $\kappa_{2-} \neq \kappa_{2+}$.

3. Transition specification:

$$\psi_3(z) = \kappa_3 \left(1 + \nu_\lambda |z|\right) z.$$

If $\nu_\lambda < 0$, the transition is able to generate non-monotonic NIC.

4. Flexible specification:

$$\psi_4(z) = \kappa_4 \left(1 + \nu_\lambda |z|\right) \operatorname{sgn}(z) |z|^{\zeta_\lambda},$$

where $\zeta_\lambda \geq 0$. This adds more curvature to the 'transition' specification.

5. Partially asymmetric transition specification:

$$\psi_5(z) = \kappa_{5-} \left(1 + \nu_\lambda |z|\right) z \mathbb{I}_{\{z<0\}} + \kappa_{5+} \left(1 + \nu_\lambda |z|\right) z \mathbb{I}_{\{z>0\}},$$

where $\kappa_{5-} \neq \kappa_{5+}$. This relaxes both the 'asymmetric linear' and 'transition' specifications. For reference, we also estimate the following cubic specification:

$$\psi_{\sim z^3}(z) = \kappa_{\sim z^3} z^3.$$

While the 'constant' specification '0' sets the skewness to be time-invariant, the 'linear' specification '1' and 'cubic' specification '$\sim z^3$' with their GARCH-like dynamics can be found in the previous literature, and so can the 'asymmetric linear' specification '2' that allows the impact to be different for positive and negative past returns, cf. (6). More complex shapes of the skewness NIC are designed to capture its possible non-monotonicity, with varying degree of flexibility, that was detected by our semiparametric analysis (see Section 2). The 'transition' specification '3' posits that the coefficient of a linear relationship depends on the size of the standardized return, making the skewness NIC, if $\nu_\lambda \neq 0$, nonlinear, and, if $\kappa_3 > 0$ and $\nu_\lambda < 0$, non-monotone as well. The 'flexible' specification '4' makes the relationship even more curved, even if $\nu_\lambda = 0$, while the 'partially asymmetric transition' specification '5' allows the impact coefficient to be different for positive and negative past returns. We have also tried a 'fully asymmetric transition' specification where the parameter ν_λ differs for positive and negative returns, but no improvement in terms of AIC was reached for any of the indexes we analyzed compared to the 'partially asymmetric transition' specification '5'.

3.2 Data

In this Section, in addition to data on S&P500 and FTSE100, we use daily log returns on three other stock indexes – NIKKEI225, DAX and CAC40. These data were also downloaded from `finance.yahoo.com` and start on different dates but end on the last trading day of 2010. The rest of Table 1 and Figure 3 confirm the typical properties of stock returns, though the unconditional skewness and kurtosis features for these indexes are milder than those for S&P500 and FTSE100. For out-of-sample tests, we use the data on S&P500 returns spanning from 01/01/2011 to 01/31/2016.

3.3 Results for S&P500

Table 2 displays estimation results for the S&P500 returns. In addition to those from the SGE distribution with various skewness specifications, results from a conditionally normal model are also given (column 'N'). The table shows, in addition to point estimates and robust standard errors,[7] the rankings by the Akaike and Bayesian information criteria, as well as the likelihood ratio tests $LR_{0|j}$ based on comparison of a dynamic skewness model 'j' and the model with constant skewness ('constant' specification '0').

The variance coefficients are stable across different models for skewness, with large persistence and a pronounced leverage effect. The 'thickness-of-tails' coefficient is also consistent throughout all SGE models. The sign of average skewness is stably negative. The persistence coefficient is moderately large in all dynamic skewness models except the '$\sim z^3$' specification where it is negative and big in absolute value though statistically insignificant. The '$\sim z^3$' specification, although is a bit better by likelihood, is not statistically different from the 'constant' skewness specification '0'; evidently, the cubic form is too steep for the actual NIC. The LR test for all other dynamic skewness specifications yields highly statistically significant differences. Most preferable by the value of likelihood is the 'flexible' specification '4', but the additional power parameter

[7] These standard errors are robust to density misspecification if the conditional density belongs to the exponential family; they are also valid with other conditional densities under the correct density specification.

is not significantly different from unity. Both information criteria deem this flexibility not worth an increase in likelihood, just as they do regarding additional asymmetry. The 'transition' specification '3' is considered an optimal degree of parsimony by both AIC and BIC. In Appendix A, we present the effects of perturbations of conditional mean, variance, density and other specifications on the shape of the 'transition' skewness NIC, which prove that this shape is highly robust to such perturbations.

Note that the 'average' slope of the impact of a standardized return on future skewness is positive and is slightly smaller for negative returns than for positive returns (specification '2'). When the non-monotonicity effect is taken into account (specification '5'), this difference is tiny and statistically insignificant ($LR_{0|5} - LR_{0|3} \approx 0.0$). The skewness NIC is indeed non-monotone, as the parameter ν_λ is negative and large in absolute value. Figure 5 depicts the skewness NIC for the five dynamic specifications '1'–'5'. One can see that the linear and piecewise linear specifications provide but crude approximation of the skewness NIC which is sharply different for smaller and larger shocks, while the other specifications imply very similar non-monotonic shapes. Note that about 72% of standardized returns do not exceed 1.0 in absolute value, for which the skewness NIC is approximately linear though steeper than the linear or asymmetric linear approximations. The slope becomes negative for standardized returns larger than about 2.5. This corresponds to 1.8% of the sample or about 140 observations, i.e. it is not the case that only a few outliers drive the result. Note that BIC is too conservative in the sense that its second-best choice is the asymmetric linear specification, which misses the phenomenon of non-monotonicity, while AIC prefers specifications with non-monotonicity to the asymmetric linear one.

In all specifications (except the 'cubic' one) the persistence coefficient κ_{-1} is close to 0.6 and is highly statistically significant. This is an evidence of moderate persistence in conditional skewness, compared to that in GARCH models for volatility. Jondeau & Rockinger (2003) also find that the conditional skewness of the S&P500 exhibits relatively high persistence; Feunou et al. (2014) obtain estimates for such coefficients close to 0.6 for the S&P500 for three subsamples that cover various periods from 1980 to 2009.

Importantly, not only is the persistence much lower for skewness than for volatility, but also the whole skewness NIC is much harder to identify from a long series of observable returns than it is to pin down the volatility NIC from much shorter periods.

Figure 6 presents the 4-year fragments of the series of returns of the S&P500, the conditional variance from the EGARCH model and the conditional skewness computed using the 'transition' specification and formula (9). Returns and volatilities exhibit familiar patterns; the skewness series is less persistent than volatility. It does not appear that volatility and skewness are related; indeed, the correlation between them is only -0.047.

3.4 Conditional density forecasting

An important dimension that helps compare different models of skewness dynamics from a practical viewpoint is conditional density forecasting. To evaluate the models from this perspective, we use the test proposed by Diebold & Mariano (1995) and extended by Amisano & Giacomini (2007). To extend the analysis to multiperiod forecasts, we follow the methodology proposed by Maheu & McCurdy (2011). The method is described in detail in Maheu & McCurdy (2011), so here we just summarize the main points and present the results. The test statistic comparing models i and j is

$$\tau_{i,j} = \frac{\sqrt{T}(D_i - D_j)}{\hat{\sigma}_{i,j}},$$

where

$$D_i = \frac{1}{T}\sum_{t=t_0}^{t_0+T} \log f_{i,k}(r_{t+k}|I_t),$$

the return density $f_{i,k}(r_{t+k}|I_t)$ is implied by model i in period $t+k$ conditional on information available in period t and evaluated at the realized return r_{t+k}, and $\hat{\sigma}_{i,j}$ is a HAC estimator of the long run standard deviation of the log density differential. Under the null hypothesis that models i and j have equal predictive properties, $\tau_{i,j}$ is asymptotically standard normal. A large positive value of $\tau_{i,j}$ is evidence of model i's better performance, a negative one, of model j's.

We compute the statistic $\tau_{j,0}$, $j \in \{1,2,3\}$ for forecast horizons of $k = 1, ..., 60$ days for both in-sample (01/02/1980–12/31/2010) and out-of-sample (01/01/2011–01/31/2016) periods for S&P500. Figure 7 shows plots of statistics $\tau_{1,0}$, $\tau_{2,0}$ and $\tau_{3,0}$ as functions of k. In-sample, there is strong evidence in favor of time-varying skewness. 'Linear' and 'asymmetric linear' specifications significantly outperform static specification '0' in short and long run horizons. 'Transition' specification shows the best in-sample performance: it is significantly better than the static specification in almost all forecasting horizons considered. Out-of-sample, 'linear' and 'asymmetric linear' specifications seem to fare best overall, though the differences tend to be statistically insignificant. In out-of-sample density forecasting, dynamic structures of skewness specification show more advantage for short and long horizons and less for medium horizons; at the same time, simpler models of dynamic skewness seem to perform better. These results are in line with the common wisdom that higher model sophistication, though leading to a better performance in-sample, may not show an advantage out-of-sample.

3.5 Results of direct models for skewness

In addition to models within our 'indirect' framework, we estimate two existing 'direct' models for dynamic skewness for the S&P500. The models of León et al. (2005) and Harvey & Siddique (1999) are described in detail in Appendix B. Table 3 contains skewness equations, estimates of their parameters with corresponding standard errors, and LR test statistics for the constancy of conditional skewness with corresponding p-values.

On top of these models' practical shortcomings discussed in subsection 2.1, empirically, the parameters of skewness dynamics κ_{-1} and κ_1 are small compared to those in Table 2 and, statistically, tend to be marginally significant at most. The LR tests show statistically insignificant deviations from static conditional skewness.

3.6 Results for other indexes

We also apply our parametric model with various skewness specifications to four major European indexes: FTSE100, NIKKEI225, DAX and CAC40. The purpose is to verify whether the shape of the skewness NIC in other liquid markets differs from that for S&P500.

The Monte Carlo analysis (see the next subsection) reveals that AIC may be more precise relative to BIC in choosing the correct specification. Table 4 reports estimates for AIC-selected specifications for the four indexes, and Figure 8 depicts the skewness NIC for all five. While the volatility estimates do not fall far apart, one observes wide diversity in estimates of skewness NIC across the markets. The skewness NIC for DAX is qualitatively similar to that for S&P500, but bigger returns cause a sharper reaction of skewness. Non-monotonicity is not found important while asymmetry is for FTSE100 (cf. Figure 4) and CAC40, and, as in the case of S&P500 ('asymmetric linear' specification), positive returns exert higher impact on skewness than negative returns. Finally, the NIKKEI225 index exhibits a monotone, albeit peculiarly nonlinear, skewness NIC. The persistence also varies significantly across the markets.

This evidence shows that while the volatility characteristics are quite similar in different stock markets, the skewness equation exhibits high variability across them.

3.7 Monte Carlo study

To confirm the reliability of the obtained results we conduct a small Monte Carlo study. We investigate the performance of the parameter estimates of the preferred skewness specification, rejection rates of the likelihood ratio test for its staticness, and frequencies of selection of particular specifications by the two information criteria.

In the first experiment, we simulate 500 artificial samples of length 7,500 (roughly corresponding to sample sizes used in our real data analysis) from the EGARCH-SGED model with dynamic 'transition' skewness specification '3' and parameters close to those obtained for S&P 500. Summary statistics of the estimates of skewness dynamics and shape parameters are reported in Table 5. All parameter estimates are nearly mean and median unbiased. Especially precise are estimates of the thickness-of-tails parameter, but the interquartile ranges of the skewness parameters are narrow enough to ensure the estimated 'rotated S'-shape of the skewness NIC if it is present in the data. Standard deviations of estimates for most of the parameters are close to standard errors of estimates obtained from the empirical analysis. Note that the excellent properties of estimates of the conditional skewness NIC are in sharp contrast to the poor precision of estimates of the *unconditional* skewness measures reported in Kim & White (2004).

Next, in Table 6 we present the results of LR testing of the null hypothesis of 'constant' skewness against its 'transition' dynamics. In the upper panel, we simulate the series from specification '0', in the lower panel, from specification '3'. There is moderate overrejection under the null hypothesis of constant skewness: the test rejects about twice as often as the nominal asymptotic size implies.[8] When instead the dynamic skewness specification drives the data, the LR test exhibits unit power for all significance levels.

Finally, Table 7 presents the results of model selection experimentation. We generate samples from the EGARCH-SGED model with static and three dynamic skewness specifications and count model selection scores using AIC and BIC. A number in the i^{th} row and j^{th} column represents the percentage of samples when specification 'i' is selected according to AIC/BIC with the data generated from the model with specification 'j'.

As expected, BIC selects more parsimonious specifications compared to AIC. Perhaps less expected, BIC almost never selects a specification that is more parameterized than the DGP, in a sense providing a 'lower bound' for the true model parsimony. However, most of the time, BIC prefers a linear model for skewness, and neglects the nonlinearity of the skewness specification when in fact it is present. Therefore, AIC seems to be a more suitable selection criterion in this context, even though moderately often it selects relatively simple dynamic specification when it is in fact static. When the 'transition'

[8]If we use size-corrected critical values for the $LR_{0|3}$ statistic in our empirical analysis, the null of constant skewness is still rejected at conventional significance levels.

specification '3' is in effect, AIC is right on target in three out of four cases.

All in all, the Monte Carlo evidence indicates that the empirical results regarding the shape of the skewness NIC obtained before appear genuine.

4 Discussion

Our empirical study has revealed several intriguing patterns observed in the distribution of stock market returns. First, for almost all indexes considered, the skewness NIC has a positive slope for moderate values of past standardized returns. Second, the skewness NIC of S&P500 and DAX are clearly non-monotone: for moderate values of standardized return, the slope of the NIC is positive, but for returns that are large in absolute value, it is negative, and this pattern is robust to specifications of volatility dynamics, density shape, exclusion of extreme events and other perturbations.

The first phenomenon may be explained by the existence of time-varying expected returns. Pindyck (1984), French et al. (1987), and Campbell & Hentschel (1992) use the hypothesis of time-varying expected return to study the relation between stock returns and volatility. Particularly, the 'volatility feedback' effect established in this literature explains the asymmetry of the volatility NIC: an increase in future expected return volatility leads to an increase in the expected return and, consequently, to a price drop and negative return in the current period; a decrease in future volatility leads to a decrease in the expected return and a price increase in the current period. Along these lines, one can explain the positive slope of the skewness NIC for moderate values of standardized returns.

Kraus & Litzenberger (1976) and Harvey & Siddique (2000) provide theoretical justification and empirical evidence that systemic skewness is priced in the stock market. Particularly, they find that investors prefer positively skewed returns and ask for a positive risk premium for negative return skewness in the stock market. It then follows that a positive shock to expected future skewness should reduce the risk premium and consequently lead to a price increase and positive return in the current period. A negative shock to the expected future skewness, on the contrary, will lead to a price drop and negative return. This explains the positive slope of the skewness NIC of stock indexes.

The second empirical phenomenon, non-monotonicity of the skewness NIC for S&P500 and DAX, is inconsistent with the time-varying risk premium hypothesis and is harder to explain. This means there might be other factors that connect skewness and returns, that overpower the 'skewness feedback' effect when returns are large in absolute value. To the best of our knowledge, this paper is the first to find this pattern and there is no off-the-shelf theory that can explain it. Intuitively, one can think about this effect in the following way. A large positive return is followed by a lower skewness in subsequent periods, which means that the probability of large negative returns becomes higher. A large negative return, in contrast, is followed with a higher skewness and higher probability of large positive returns. These informal observations are reminiscent of the market overreaction and market rebound phenomena and can give some hints about possible explanations of the pattern; however a rigorous theory that could explain these observations is yet to be developed. Another important question in this regard is why this effect is observed for some and not for other indexes. We leave these questions to future research.

Our empirical findings are also interesting from a practical point of view. Skewness is a characteristic that largely determines the probability distribution of tail events. Our results suggest that the probability of tail events is time-varying and predictable. Better models for conditional skewness thus provide more precise estimation and control of tail risks, which is crucial in dynamic asset allocation.

5 Concluding remarks

Studying the distributional asymmetry of financial returns may take different approaches: Feunou et al. (2013) work with an asymmetry measure based on the difference between upside and downside realized volatilities; Ghysels et al. (2015) analyze asymmetry using robust skewness measures that are computed from quantiles. We take a conventional route and model the whole conditional return distribution, including the evolution of conditional skewness, with an eye to the skewness NIC. This approach allows one not only to quantify the degree of asymmetry and analyze how it evolves over time but also study the link between current news and future asymmetry. Because our approach involves full parametric specification of the conditional distribution, the results can be straightforwardly used in problems of asset allocation, risk management and option pricing.

We have discovered that even though the skewness equation is much more difficult to identify from return data, the skewness is negative on average, it has a moderate degree of persistence, and its NIC tends to exhibit positive slope, (sometimes) sign asymmetry, (sometimes) non-monotonicity and high diversity across different returns. While there is a broad literature that emphasizes the importance of time-varying skewness and a slim literature on modeling the skewness NIC, there is little research that would provide a theoretical ground for the findings. While we are able to explain some of those, more research is needed to explain the newly discovered phenomena on the theory level.

References

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, 118, 135–167.

Amisano, G. & Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25, 177–190.

Anatolyev, S. & Shakin, D. (2007). Trade intensity in the Russian stock market: dynamics, distribution and determinants. *Applied Financial Economics*, 17, 87–104.

Athayde, G. M. & Flôres, R. G. (2004). Finding a maximum skewness portfolio – a general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, 28, 1335–1352.

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, 269–282.

Bali, T. G. & Theodossiou, P. (2008). Risk measurement performance of alternative distribution functions. *Journal of Risk and Insurance*, 75, 411–437.

Brännäs, K. & Nordman, N. (2003). An alternative conditional asymmetry specification for stock returns. *Applied Financial Economics*, 13, 537–541.

Brooks, C., Burke, S. P., Heravi, S., & Persand, G. (2005). Autoregressive conditional kurtosis. *Journal of Financial Econometrics*, 3, 399–421.

Campbell, J. Y. & Hentschel, L. (1992). No news is good news: an asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31, 281–318.

Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4, 537–572.

DeMiguel, V., Plyakha, Y., Uppal, R., & Vilkov, G. (2013). Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis*, 48, 1813–1845.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13, 253–263.

Duffie, D. & Pan, J. (1997). An overview of value at risk. *Journal of Derivatives*, 4, 7–49.

Engle, R. F. (2011). Long-term skewness and systemic risk. *Journal of Financial Econometrics*, 9, 437–468.

Engle, R. F. & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, 48, 1749–1778.

Feunou, B., Jahan-Parvar, M. R., & Tédongap, R. (2013). Modeling market downside volatility. *Review of Finance*, 17, 443–481.

Feunou, B., Jahan-Parvar, M. R., & Tédongap, R. (2014). Which parametric model for conditional skewness? *European Journal of Finance*, 1, 1–35.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19, 3–29.

Ghysels, E., Plazzi, A., & Valkanov, R. (2015). Why invest in emerging markets? The role of conditional return asymmetry. *Journal of Finance*, forthcoming.

Grigoletto, M. & Lisi, F. (2009). Looking for skewness in financial time series. *Econometrics Journal*, 12, 310–323.

Guidolin, M. & Timmermann, A. (2008). International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies*, 21, 889–935.

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35, 705–730.

Harvey, C. R. & Siddique, A. (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, 34, 465–488.

Harvey, C. R. & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance*, 55, 1263–1295.

Hashmi, A. R. & Tay, A. S. (2007). Global regional sources of risk in equity markets: evidence from factor models with time-varying conditional skewness. *Journal of International Money and Finance*, 26, 430–453.

Jha, R. & Kalimipalli, M. (2010). The economic significance of conditional skewness in index option markets. *Journal of Futures Markets*, 30, 378–406.

Jondeau, E. & Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics and Control*, 27, 1699–1737.

Kane, A. (1982). Skewness preference and portfolio choice. *Journal of Financial and Quantitative Analysis*, 17, 15–25.

Kim, T.-H. & White, H. (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1, 56–73.

Kraus, A. & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *Journal of Finance*, 31, 1085–1100.

León, A., Rubio, G., & Serna, G. (2005). Autoregresive conditional volatility, skewness and kurtosis. *Quarterly Review of Economics and Finance*, 45, 599–618.

Maheu, J. M. & McCurdy, T. H. (2011). Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160, 69–76.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59, 347–370.

Neuberger, A. (2012). Realized skewness. *Review of Financial Studies*, 25, 3423–3455.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, 2, 130–168.

Peiro, A. (1999). Skewness in financial returns. *Journal of Banking and Finance*, 23, 847–862.

Pindyck, R. S. (1984). Risk, inflation, and the stock market. *American Economic Review*, 74, 335–351.

Premaratne, G. & Bera, A. (2000). Modeling asymmetry and excess kurtosis in stock return data. Working paper no. 00–123, University of Illinois.

Rehman, Z. & Vilkov, G. (2012). Risk-neutral skewness: return predictability and its sources. Working paper, Goethe University Frankfurt.

Rodríguez, M. J. & Ruiz, E. (2012). Revisiting several popular GARCH models with leverage effect: differences and similarities. *Journal of Financial Econometrics*, 10, 637–668.

Schneider, P., Wagner, C., & Zechner, J. (2016). Low risk anomalies? SSRN paper no. 2593519.

Simaan, Y. (1993). Portfolio selection and asset pricing: three-parameter framework. *Management Science*, 39, 568–577.

White, H., Kim, T.-H., & Manganelli, S. (2010). Modeling autoregressive conditional skewness and kurtosis with multi-quantile CAViaR. In T. Bollerslev, J. Russell, & M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Oxford: Oxford University Press.

Wilhelmsson, A. (2009). Value at risk with time varying variance, skewness and kurtosis – the NIG-ACD model. *Econometrics Journal*, 12, 82–104.

Yan, J. (2005). Asymmetry, fat-tail, and autoregressive conditional density in financial return data with systems of frequency curves. Working paper no. 355, University of Iowa.

Tables and Figures

Table 1: Summary statistics for log returns series

	S&P500	FTSE100	NIKKEI225	DAX	CAC40
Start	01/02/80	01/02/85	01/02/85	11/26/90	03/01/90
Length	7822	6570	6388	5079	5269
Mean	0.031	0.023	−0.002	0.031	0.014
Median	0.054	0.013	0.028	0.079	0.033
Min	−22.900	−13.029	−16.137	−9.871	−9.472
Max	10.957	9.384	13.235	10.797	10.595
Standard deviation	1.148	1.117	1.487	1.457	1.419
Skewness	−1.214	−0.520	−0.213	−0.091	−0.012
Kurtosis	31.012	13.627	11.045	7.957	7.800

Notes: Log returns are defined as $r_t = 100 \log(P_t/P_{t-1})$, where P_t is close price at date t. Samples end on December 30 or 31, 2010.

Table 2: Estimation results for S&P500

	S&P500							
Model $j \rightarrow$	0	1	2	3	4	5	$\sim z^3$	N
variance equation								
β_{-1}	0.986 (0.001)	0.987 (0.003)	0.987 (0.004)	0.988 (0.003)	0.988 (0.002)	0.988 (0.003)	0.986 (0.003)	0.983 (0.004)
β_1	-0.075 (0.008)	-0.078 (0.010)	-0.077 (0.019)	-0.074 (0.010)	-0.074 (0.006)	-0.074 (0.009)	-0.076 (0.010)	-0.083 (0.017)
$\beta_{\lvert 1\rvert}$	0.119 (0.010)	0.123 (0.013)	0.122 (0.013)	0.117 (0.012)	0.118 (0.009)	0.117 (0.012)	0.120 (0.013)	0.131 (0.023)
skewness equation								
κ_0	-0.138 (0.027)	-0.051 (0.020)	-0.078 (0.025)	-0.050 (0.016)	-0.051 (0.015)	-0.051 (0.054)	-0.179 (0.066)	
κ_{-1}		0.582 (0.113)	0.579 (0.080)	0.606 (0.088)	0.599 (0.083)	0.607 (0.084)	-0.313 (0.400)	
$\kappa_j, \begin{pmatrix} \kappa_{j-} \\ \kappa_{j+} \end{pmatrix}$		0.120 (0.028)	0.097 (0.036) 0.168 (0.026)	0.231 (0.042)	0.237 (0.043)	0.220 (0.080) 0.235 (0.135)	0.001 (0.001)	
ν_λ				-0.212 (0.021)	-0.205 (0.056)	-0.213 (0.033)		
ζ_λ					0.851 (0.221)			
shape parameter								
η	1.36 (0.04)	1.36 (0.04)	1.36 (0.04)	1.35 (0.04)	1.35 (0.03)	1.36 (0.04)	1.36 (0.04)	
diagnostics								
$LR_{0\mid j}$ (p-value)	0	44.5 (0.000)	46.2 (0.000)	57.0 (0.000)	57.3 (0.000)	57.0 (0.000)	2.1 (0.144)	
AIC	7	4	5	1	2	3	6	8
BIC	6	2	5	1	3	4	7	8

Notes: Robust standard errors are in parentheses. Row $LR_{0|j}$ shows loglikelihood ratio tests comparing current specification 'j' with specification '0'. AIC and BIC are Akaike and Bayesian information criteria, and corresponding numbers point at rankings of models. For model specifications, see main text.

Table 3: Estimation results of direct models for S&P500

Source	Skewness specification	κ_0	κ_{-1}	κ_1	LR_0 (p-value)
Leon et al (2005)	$s_t = \kappa_0 + \kappa_{-1} s_{t-1} + \kappa_1 z_{t-1}^3$	-0.058 (0.017)	0.014 (0.153)	0.006 (0.003)	3.13 (0.209)
Harvey & Siddique (1999)	$s_t = \kappa_0 + \kappa_{-1} s_{t-1} + \kappa_1 \sigma_{t-1}^3 z_{t-1}^3$	0.017 (0.000)	-0.088 (0.095)	0.040 (0.027)	4.53 (0.104)

Notes: Robust standard errors are in parentheses. Column LR_0 shows loglikelihood ratio test comparing current specification with analogous constant specification (the null $\kappa_{-1} = \kappa_1 = 0$). Full model specifications are described in Appendix B.

Table 4: Estimation results for other indexes

	FTSE100	NIKKEI225	DAX	CAC40
Model $j \rightarrow$	2	4	3	2
		variance equation		
β_{-1}	0.985 (0.003)	0.977 (0.004)	0.984 (0.003)	0.983 (0.003)
β_1	-0.067 (0.008)	-0.091 (0.012)	-0.071 (0.009)	-0.084 (0.009)
$\beta_{\lvert 1 \rvert}$	0.150 (0.015)	0.179 (0.017)	0.134 (0.014)	0.119 (0.014)
		skewness equation		
κ_0	-0.095 (0.030)	-0.067 (0.031)	-0.236 (0.057)	-0.347 (0.090)
κ_{-1}	0.832 (0.071)	0.302 (0.172)	-0.158 (0.072)	-0.579 (0.268)
κ_1				
$\begin{pmatrix} \kappa_{2-} \\ \kappa_{2+} \end{pmatrix}$	$\begin{pmatrix} 0.011 \\ (0.033) \\ 0.163 \end{pmatrix}$ (0.049)			$\begin{pmatrix} -0.065 \\ (0.047) \\ 0.145 \end{pmatrix}$ (0.059)
κ_3			0.232 (0.073)	
κ_4		0.090 (0.062)		
$\begin{pmatrix} \kappa_{5-} \\ \kappa_{5+} \end{pmatrix}$				
ν_λ		-0.104 (0.028)	-0.358 (0.046)	
ζ_λ		1.161 (0.818)		
		shape parameter		
η	1.69 (0.05)	1.44 (0.05)	1.46 (0.08)	1.62 (0.08)
LR_0 (p-value)	33.2 (0.000)	34.2 (0.000)	11.4 (0.001)	8.5 (0.004)

Notes: Robust standard errors are in parentheses. Row $LR_{0|j}$ shows loglikelihood ratio tests comparing current specification 'j' with specification '0'. For model specifications, see main text.

Table 5: Monte Carlo study: parameter estimates

	κ_0	κ_{-1}	κ_3	ν	η
True value	-0.050	0.610	0.230	-0.210	1.350
Mean	-0.052	0.603	0.237	-0.206	1.350
Standard deviation	0.017	0.093	0.052	0.077	0.031
$Q_{5\%}$	-0.080	0.453	0.153	-0.300	1.297
$Q_{25\%}$	-0.061	0.546	0.203	-0.253	1.330
$Q_{50\%}$	-0.050	0.610	0.240	-0.219	1.350
$Q_{75\%}$	-0.040	0.666	0.273	-0.176	1.369
$Q_{95\%}$	-0.028	0.741	0.322	-0.078	1.405

Notes: Simulations are based on 500 samples of length 7,500 drawn from EGARCH-SGED model with dynamic skewness specification '3'. $Q_{X\%}$ denotes X-percentile of empirical distribution of estimates.

Table 6: Monte Carlo study: likelihood ratio test

Nominal size	1%	2.5%	5%	10%
Simulated data from model with 'constant' specification '0'				
Rejection rate	2.6%	5.0%	11.0%	18.0%
Simulated data from model with 'transition' specification '3'				
Rejection rate	100.0%	100.0%	100.0%	100.0%

Notes: Simulations are based on 500 samples of length 7,500 drawn from EGARCH-SGED model with static (upper panel) and dynamic (lower panel) skewness specifications. Figures indicate rejection rates by likelihood ratio test $LR_{0|3}$. Critical values are obtained from chi-squared distribution with three degrees of freedom.

Table 7: Monte Carlo study: model selection

DGP $\rightarrow$	0	1	2	3	0	1	2	3
$\downarrow$ Model	AIC				BIC			
0	63%	0%	0%	0%	100%	4%	2%	0%
1	15%	72%	52%	18%	0%	95%	95%	75%
2	14%	12%	33%	6%	0%	1%	3%	0%
3	8%	16%	15%	76%	0%	0%	0%	24%

Notes: Simulations are based on 500 samples of length 7500 drawn from EGARCH-SGED model with each of skewness specifications '0', '1', '2' and '3' with parameter values matching empirical estimates for S&P 500. In i^{th} row and j^{th} column we report a percentage of samples when specification 'i' is selected according to AIC or BIC with data generated from model with specification 'j'.

Figure 1: The SGE density for different values of λ and η

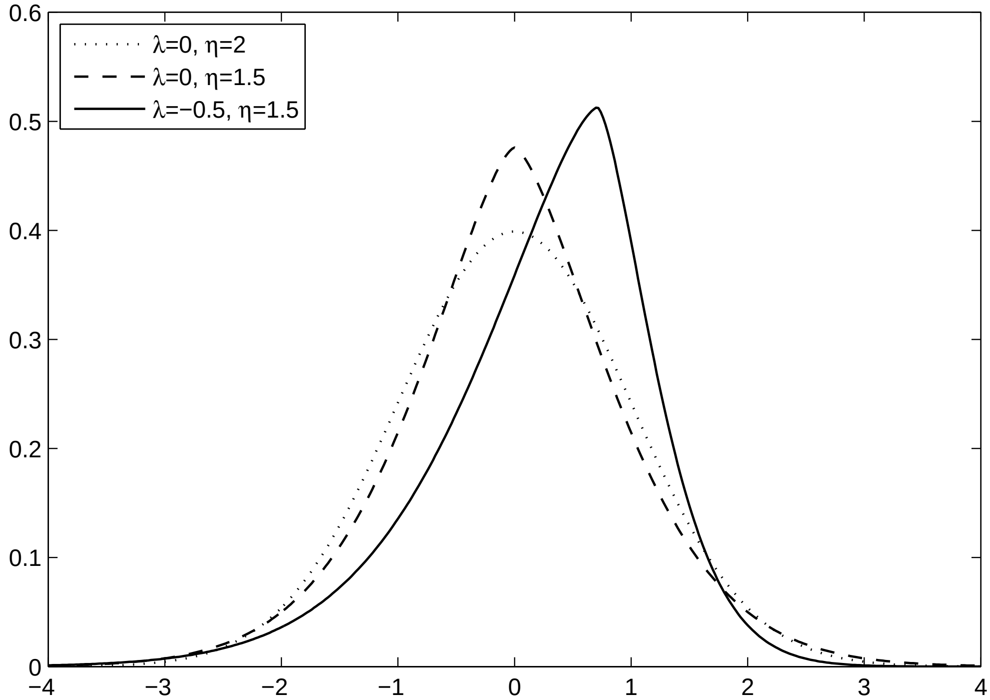


Figure 2: Skewness of SGE distribution for different values of η as a function of λ and $\tilde{\lambda}$

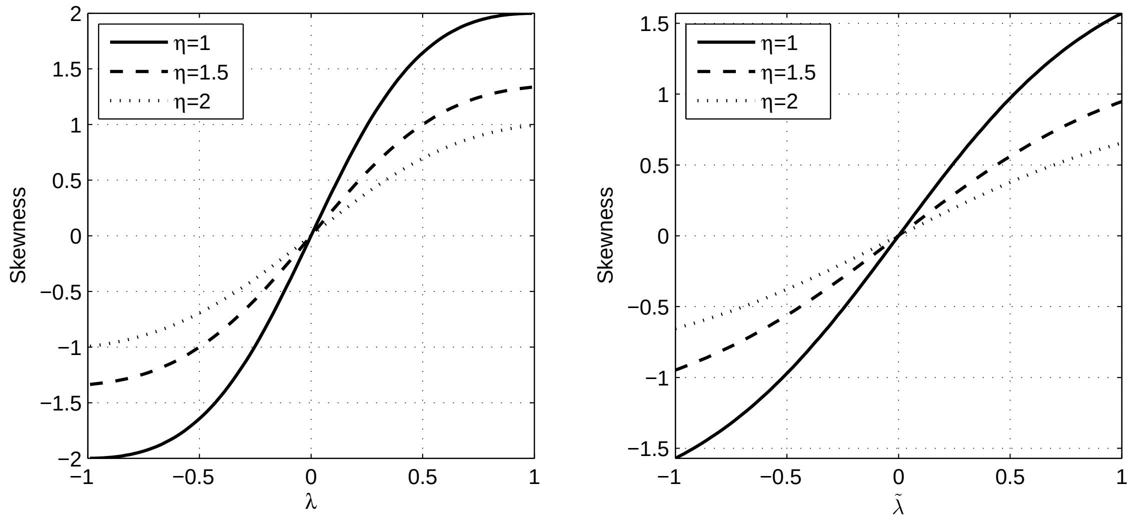
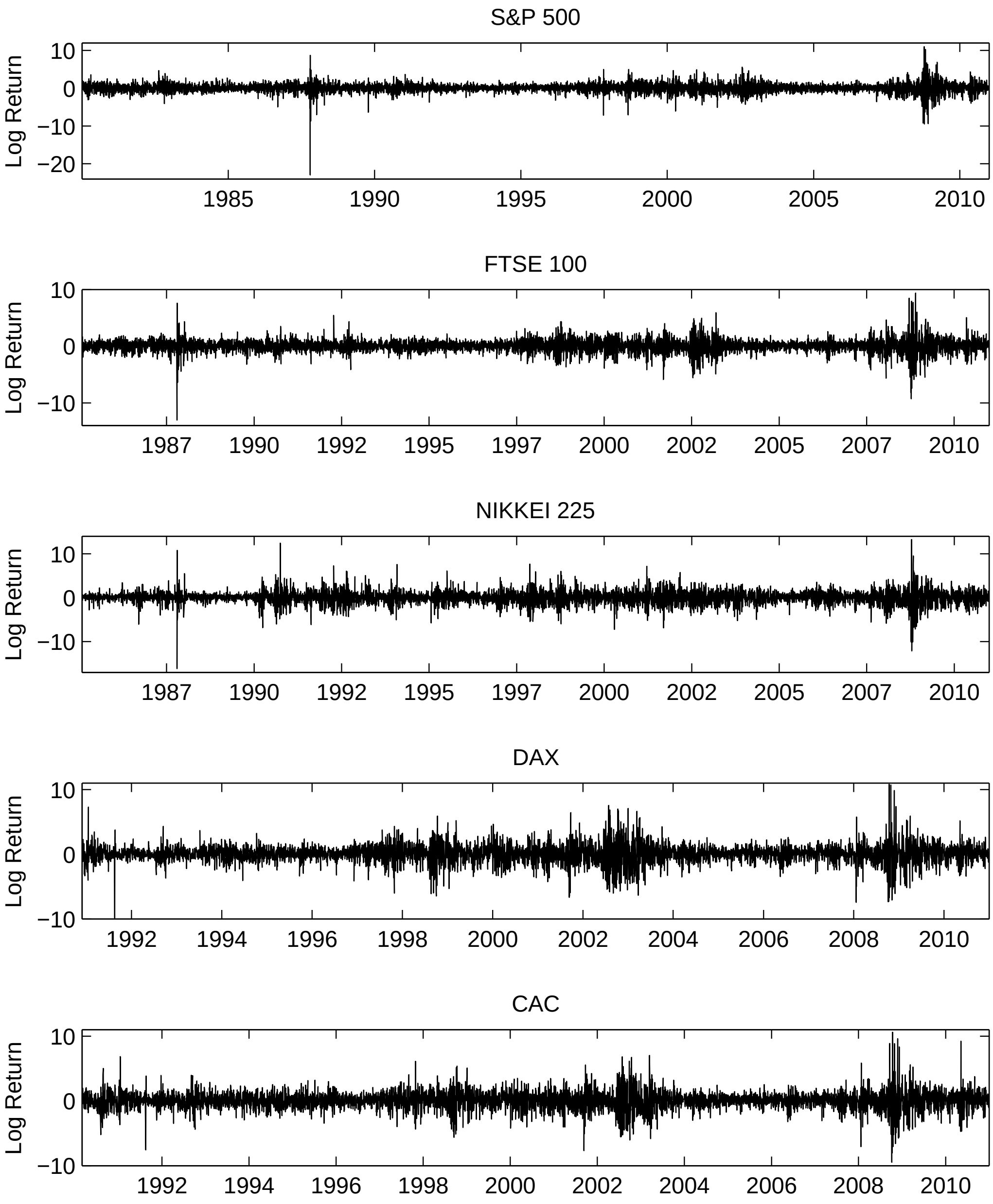
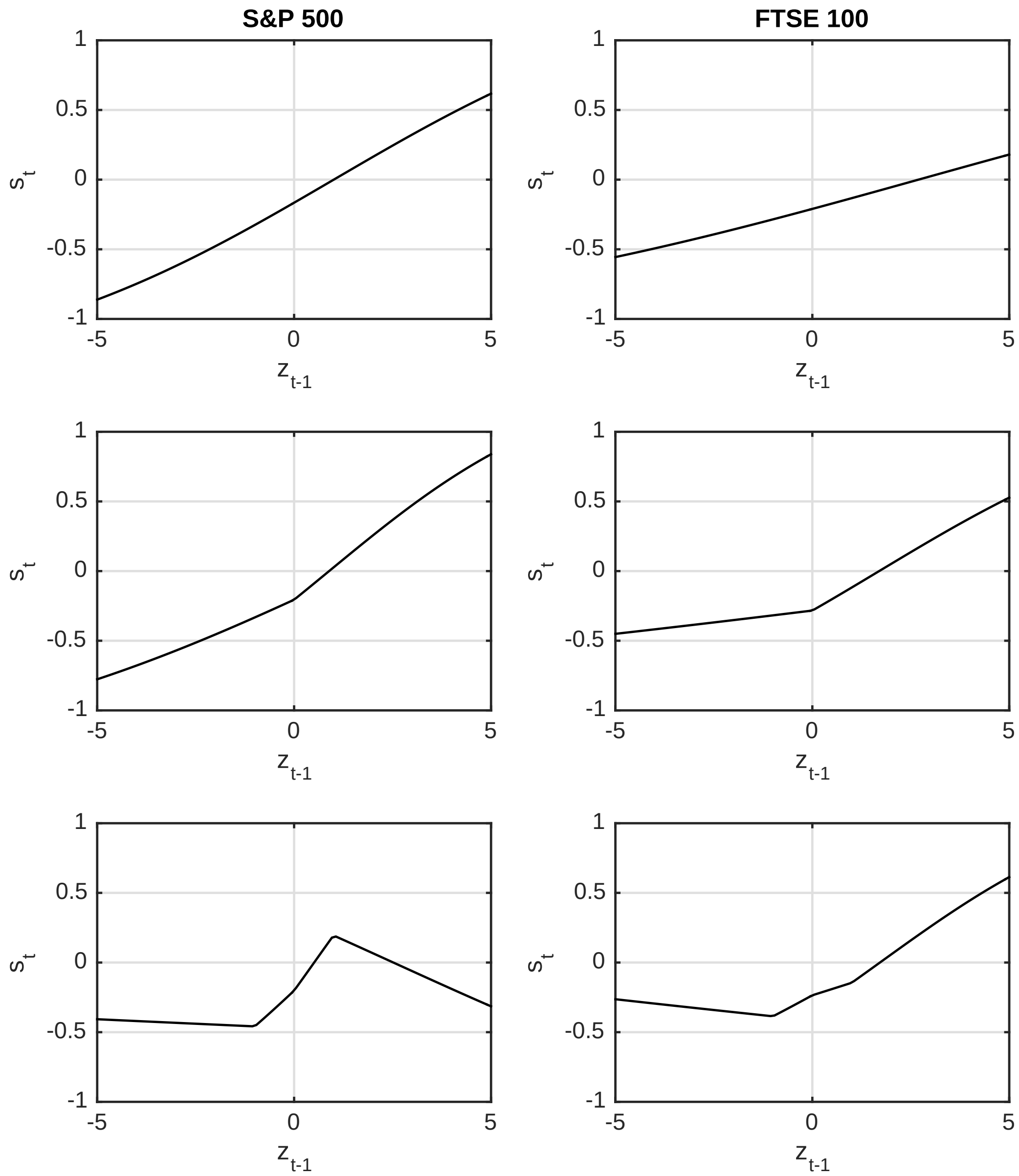


Figure 3: Dynamics of log returns



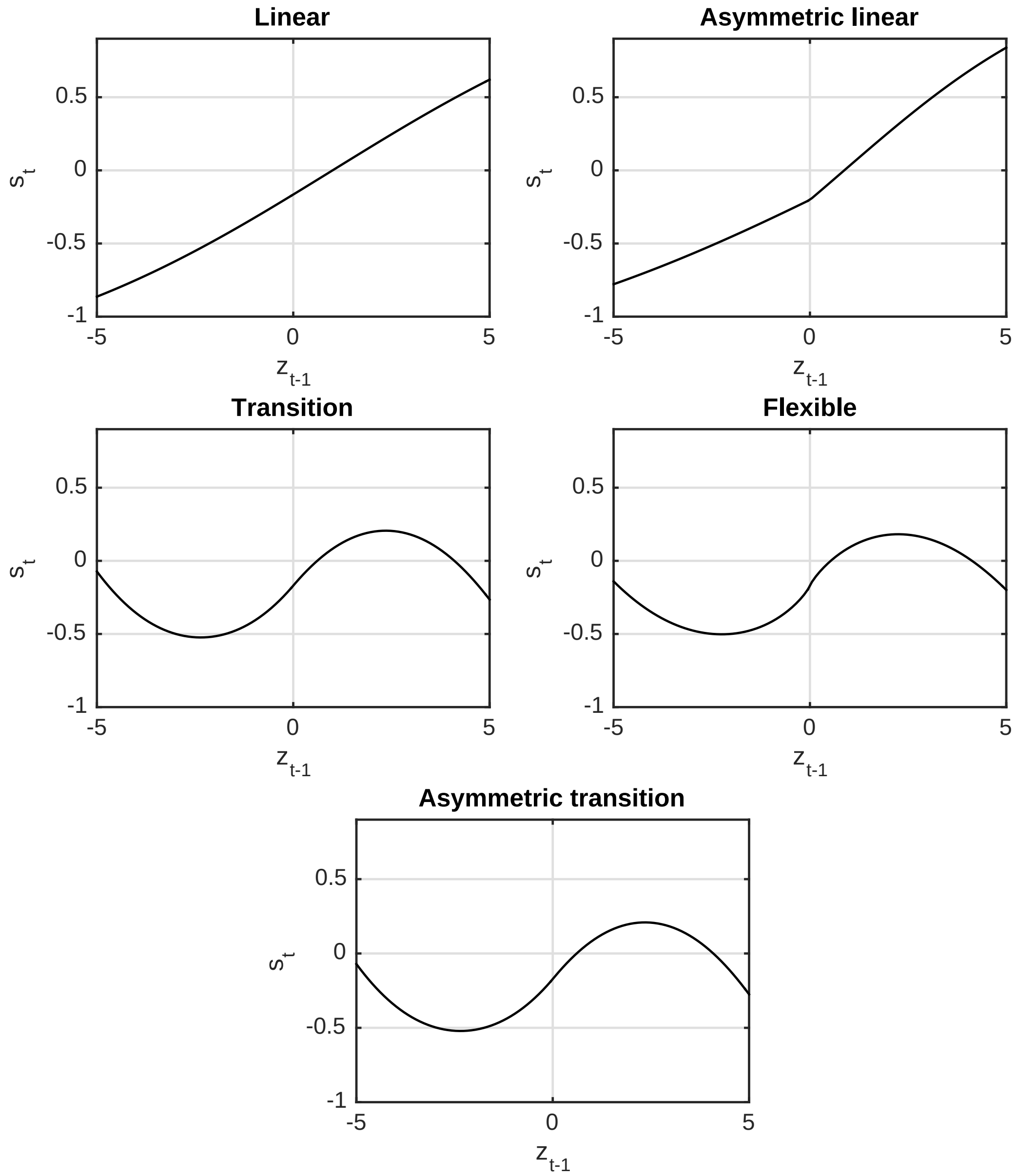
Notes: Log returns are defined as $r_t = 100 \log(P_t/P_{t-1})$, where P_t is an index close price at date t adjusted for dividends and splits.

Figure 4: Semiparametric skewness NIC for S&P500 and FTSE100 with zero, one and three knots



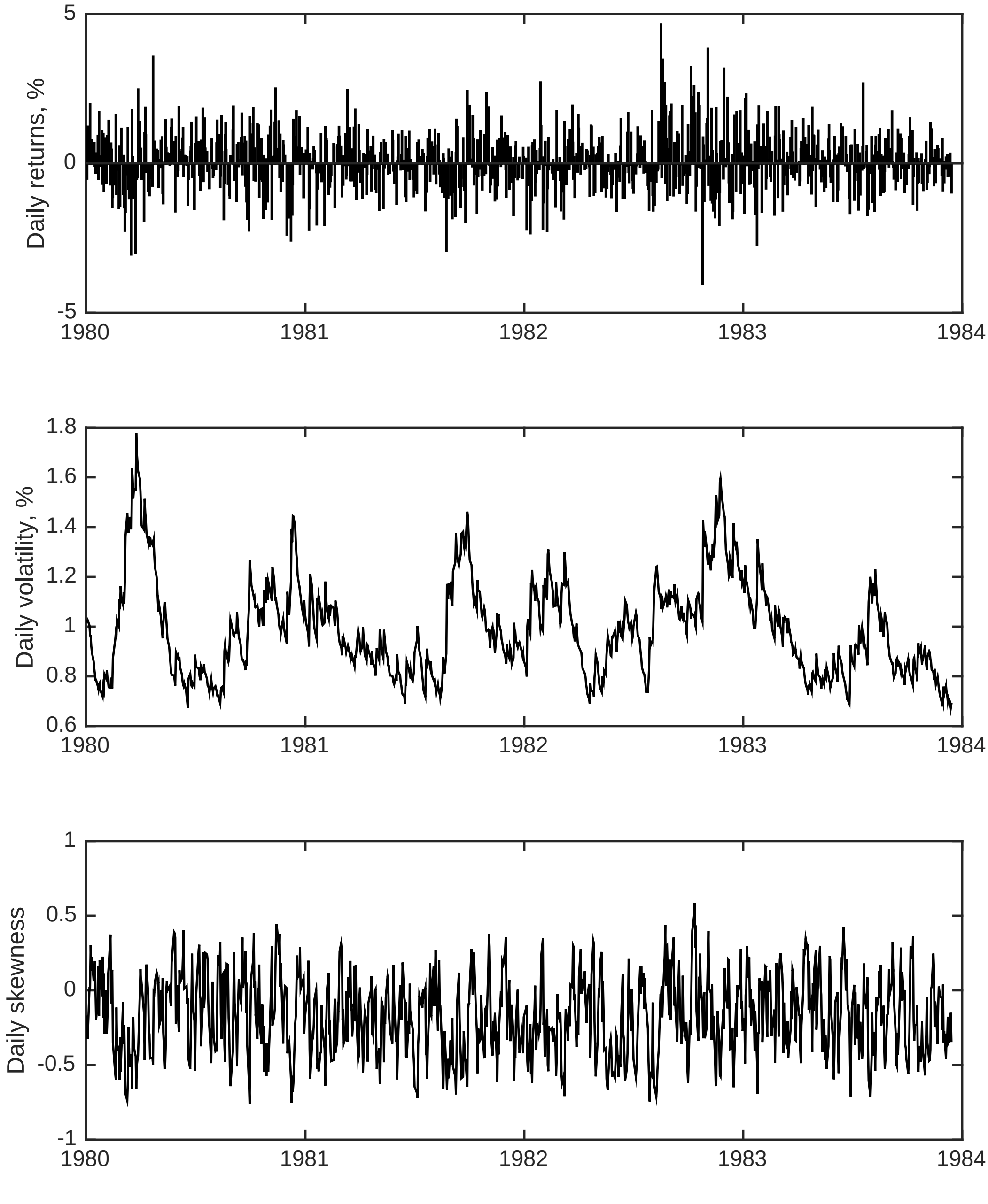
Notes: Skewness NIC is inferred from semiparametric model for conditional skewness s_t for daily S&P500 log returns depending on standardized residual z_{t-1} conditional on average $\tilde{\lambda}_{t-1}$. Top panels illustrate a model with zero knots (linear NIC), middle panels – semiparametric model with 1 knot (asymmetric linear model), bottom panels – semiparametric model with 3 knots.

Figure 5: Parametric skewness NIC for S&P500 with five specifications



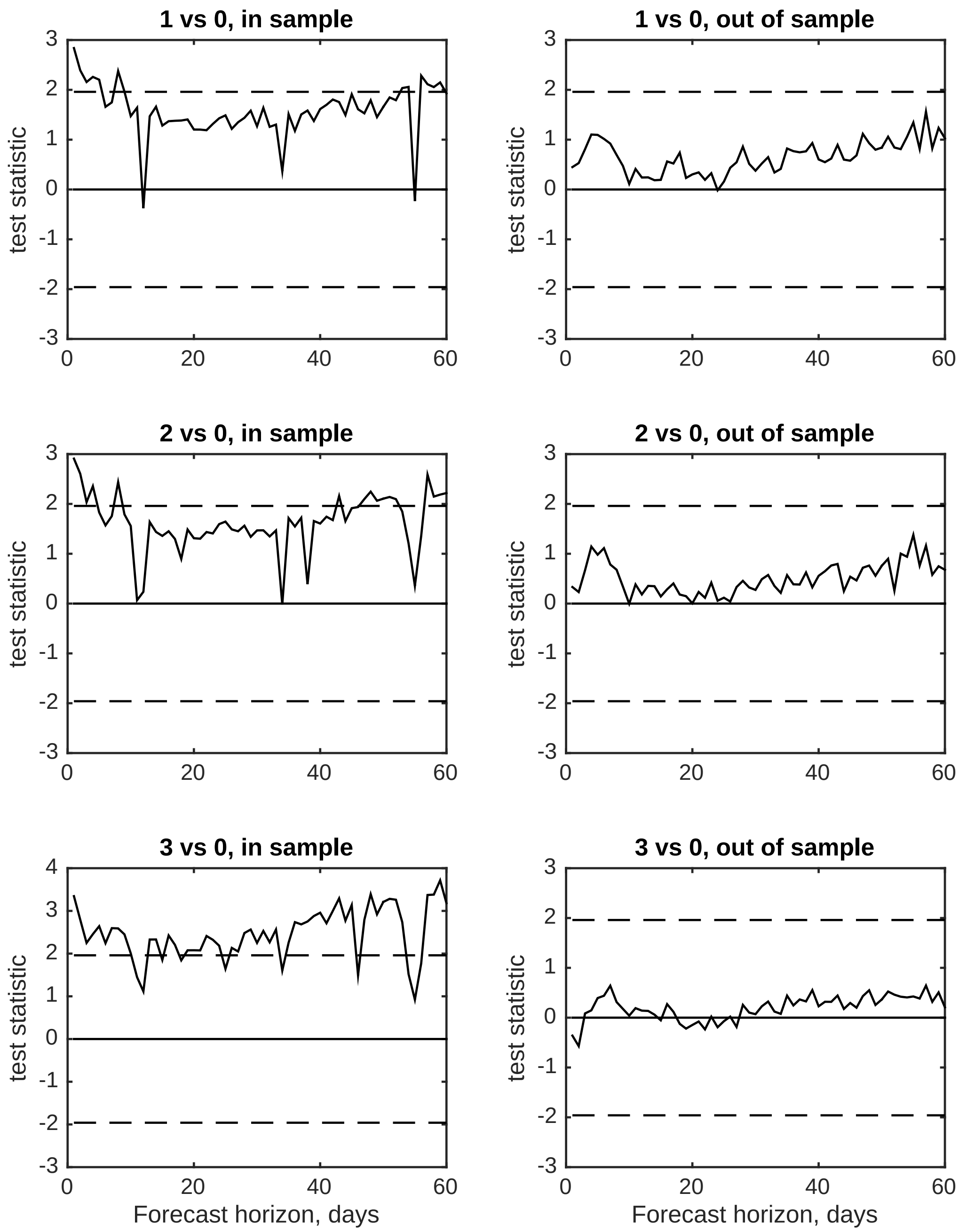
Notes: Skewness NIC is inferred from parametric models for conditional skewness s_t for daily S&P500 log returns depending on standardized residual z_{t-1} conditional on average $\tilde{\lambda}_{t-1}$.

Figure 6: Returns, volatility and skewness



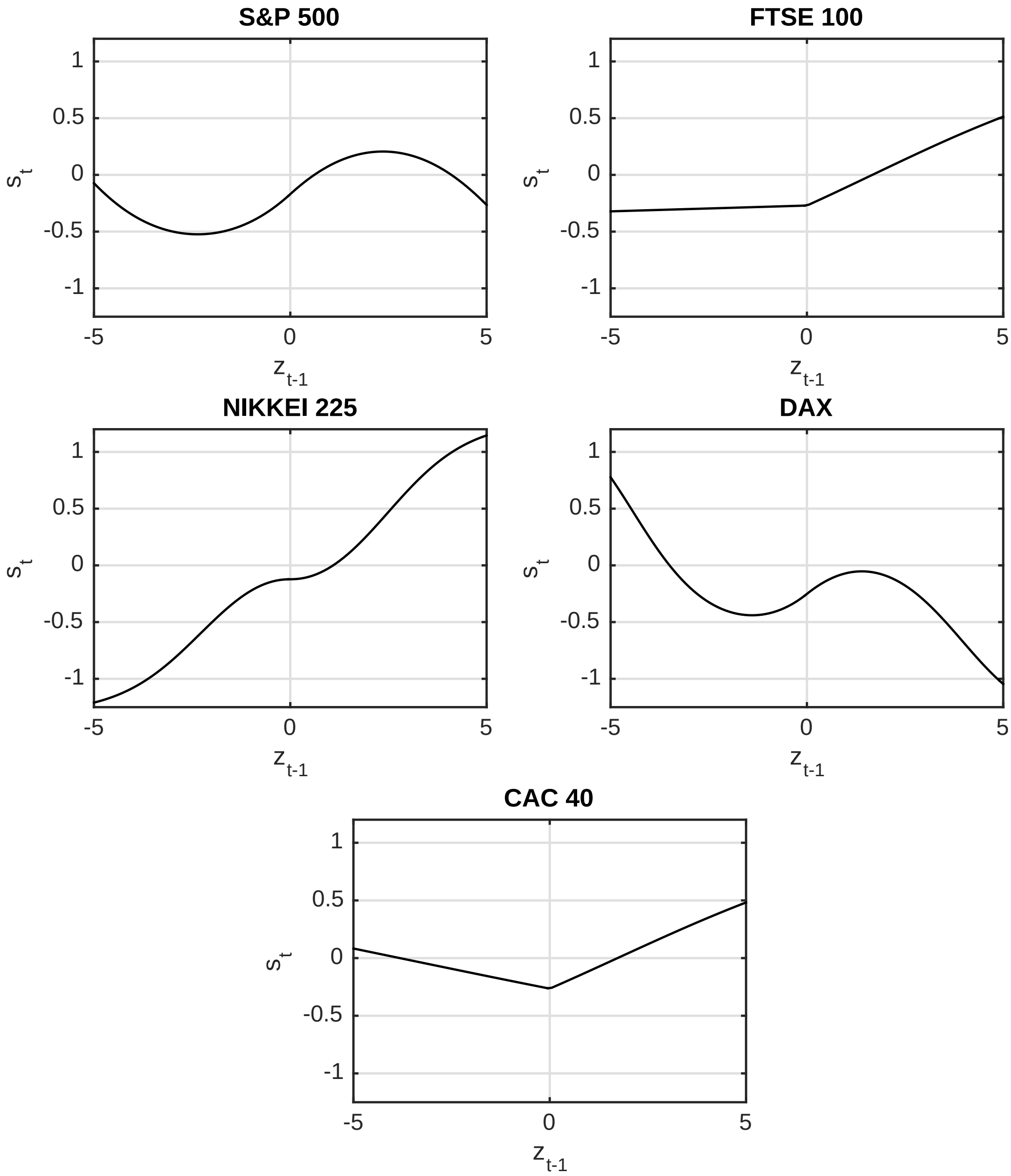
Notes: Top panel shows fragment of series of daily S&P500 log returns, middle panel shows fragment of series of conditional variance, bottom panel shows fragment of series of conditional skewness.

Figure 7: Conditional density forecasting



Notes: Test statistics $\tau_{i,j}$ compare average predictive log densities of models i and j. They are shown for daily S&P500 log returns as functions of forecast horizon k. Positive statistic for 'i vs j' is evidence in favor of model i against model j. Dashed lines correspond to 5% significance levels.

Figure 8: Parametric skewness NIC for five indexes



Notes: Skewness NIC is inferred from parametric models for conditional skewness s_t for daily S&P500, FTSE100, NIKKEI225, DAX and CAC40 log returns depending on standardized residual z_{t-1} conditional on average $\tilde{\lambda}_{t-1}$.