2019-05

Working paper. Economics

ISSN 2340-5031

Score-driven time series models with dynamic shape: An application to the Standard & Poor's 500 index

Astrid Ayala , Szabolcs Blazsek and Alvaro Escribano

uc3m | Universidad Carlos III de Madrid

Serie disponible en http://hdl.handle.net/10016/11

Web: http://economia.uc3m.es/

Correo electrónico: departamento.economia@eco.uc3m.es

Score-driven time series models with dynamic shape: An application to the Standard & Poor's 500 index

Astrid Ayala[a] , Szabolcs Blazsek[a] and Alvaro Escribano[*,b]

[a] *School of Business, Universidad Francisco MarroquÍn, Guatemala*

[b] *Department of Economics, Universidad Carlos III de Madrid, Spain*

Abstract: We introduce new dynamic conditional score (DCS) volatility models with dynamic scale and shape parameters for the effective measurement of volatility. In the new models, we use the EGB2 (exponential generalized beta of the second kind), NIG (normal-inverse Gaussian) and Skew-Gen-t (skewed generalized-t) probability distributions. Those distributions involve several shape parameters that control the dynamic skewness, tail shape and peakedness of financial returns. We use daily return data from the Standard & Poor's 500 (S&P 500) index for the period of January 4, 1950 to December 30, 2017. We estimate all models by using the maximum likelihood (ML) method, and we present the conditions of consistency and asymptotic normality of the ML estimates. We study those conditions for the S&P 500 and we also perform diagnostic tests for the residuals. The statistical performances of several DCS specifications with dynamic shape are superior to the statistical performance of the DCS specification with constant shape. Outliers in the shape parameters are associated with important announcements that affected the United States (US) stock market. Our results motivate the application of the new DCS models to volatility measurement, pricing financial derivatives, or estimation of the value-at-risk (VaR) and expected shortfall (ES) metrics.

Keywords: Dynamic conditional score (DCS) models, score-driven shape parameters
JEL classification codes: C22, C52, C58
Compiled: March 4, 2019

* Corresponding author. UC3M-Chair for Internationalization. Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903, Getafe (Madrid), Spain. Telephone: +34-916249854.
E-mail address: alvaroe@eco.uc3m.es.

1. Introduction

The precise measurement of the probability distribution of financial returns and, more specifically, the precise measurement of volatility, are important concerns of practitioners for the effective management of financial portfolios. When probability distributions that include scale and shape parameters are used to model financial returns, then both parameters will influence volatility. In classical financial time series models (e.g. Engle, 1982; Bollerslev, 1986, 1987; Nelson, 1991; Harvey et al., 1994; Harvey and Shephard, 1996; Kim et al., 1998; Barndorff-Nielsen and Shephard, 2002), the scale parameter is dynamic and the shape parameter (if it is specified) is constant over time. In this paper, we introduce new dynamic conditional score (DCS) models (Creal et al., 2011, 2013; Harvey, 2013) for the measurement of conditional volatility, in which the shape parameters are dynamic.

Harvey (2013) provides the following motivation for using DCS models: "The asymptotic distribution theory for a wide range of dynamic conditional score models is of crucial importance in showing their viability. The information matrix can be obtained explicitly, and the proof of the asymptotic normality of the maximum likelihood estimators is relatively straightforward. This contrasts with the situation for most other cases of nonlinear dynamic models. For example, no explicit information matrix is available for the most commonly used GARCH models, whereas for EGARCH models there is virtually no asymptotic theory for ML estimation" (Harvey, 2013, p. 19). The DCS models of the present paper extend the previous financial time series models with constant shape parameters from the literature, since: (i) they have a superior likelihood-based statistical performance; (ii) they estimate the dynamics of both scale and shape parameters effectively; (iii) news on asset value updates volatility not only through scale, but also through shape; (iv) they use different dynamic tail shape for the left and right tails of the return distribution; (v) they identify extreme events effectively; (vi) we present the conditions for the asymptotic properties of the maximum likelihood (ML) estimator for DCS models with several score-driven dynamic variables.

We introduce new DCS models for the EGB2 (exponential generalized beta of the second kind) (e.g. Caivano and Harvey, 2014), NIG (normal-inverse Gaussian) (Barndorff-Nielsen and Halgreen, 1977), and Skew-Gen-t (skewed generalized-t) (e.g. McDonald and Michelfelder, 2017) distributions, for which the error term includes several shape parameters. DCS models are robust to extreme observations, since the score function that updates the dynamic equations discounts the effects of those observations. We show that this property is also true for the new DCS models with dynamic shape parameters. Those models are extensions of the DCS models with constant shape parameters from the works of Harvey (2013), Caivano and Harvey (2014), Harvey and Sucarrat (2014), Harvey and Lange (2017) and Blazsek et al. (2018). We also refer to some recent studies of the authors of the present paper that contribute to the DCS literature: Blazsek and Villatoro (2015), Ayala et al. (2015, 2016), Blazsek and Escribano (2016), Blazsek and Mendoza (2016), Blazsek and Ho (2017), Blazsek and Monteros (2017a,b), Ayala and Blazsek (2018a,b,c), Blazsek and Hernandez (2018), and Ayala and Blazsek (2019).

In the body of literature relevant to this paper, different econometric methods are used to investigate dynamic tail shape for financial returns. Quintos et al. (2001) construct tests of tail shape constancy that allow for an unknown breakpoint, and present applications of those tests for stock price data. Galbraith and Zernov (2004) present applications of the same tests for the Dow Jones Industrial Average (DJIA) and Standard & Poor's 500 (S&P 500) indexes. More recently, Bollerslev and Todorov (2011) suggest a flexible nonparametric method of dynamic tail shape, which is used by the same authors for high-frequency data from the S&P 500. There are several methods in the body of literature that use options data to estimate dynamic tail shape for financial returns (e.g. Bakshi et al., 2003; Bollerslev et al., 2009; Backus et al., 2011). In relation to options data and dynamic tail shape, we also refer to the recent works of Bollerslev and Todorov (2014) and Bollerslev et al. (2015). Furthermore, by using panel data models, Kelly and Jiang (2014) identify a common variation in the tail shape of United States (US) stock returns. In the present paper, (i) we use a new flexible parametric approach to estimate dynamic tail shape; (ii) the proposed econometric models are not only for the dynamic modeling of tail shape, but also for the dynamic modeling of skewness and peakedness of the distribution.

We use daily log-return time series data from the adjusted S&P 500 index for the period of January 4, 1950 to December 30, 2017. The application of S&P 500 data to the new DCS models is useful, for example, for investors of (i) well-diversified US equity portfolios; (ii) S&P 500 futures and options contracts traded at the Chicago Mercantile Exchange (CME); (iii) exchange traded funds (ETFs) related to the S&P 500. For practitioners, the new DCS models with dynamic shape parameters may provide precise estimates and forecasts of (i) stock market volatility for pricing financial derivatives (Hull 2018), and (ii) other classical risk measurement metrics, such as value-at-risk (VaR) (Jorion, 2006) and expected shortfall (ES) (Acharya et al., 2012, 2017).

We estimate all DCS models by using the ML method, and we present the conditions for the asymptotic properties of the ML estimator. For all DCS specifications, we also perform diagnostic tests with respect to the probability distribution of the error term up to the fourth moment. For those diagnostic tests, we define outlier-robust standardized auxiliary variables that are martingale difference sequences (MDSs) under correct specification, and we use the MDS test from the work of Escanciano and Lobato (2009). We compare the in-sample statistical performances of DCS specifications with dynamic and constant shape parameters. We find that the score-driven dynamics of shape parameters are significant for the new DCS models, and we show that the performances of several DCS specifications with dynamic shape parameters are superior to the performance of the DCS specification with constant shape parameters. Outliers in the shape parameters are associated with important announcements that affected the US stock market. For the S&P 500, we find that the likelihood-based performance of Skew-Gen-t-DCS is superior to that of EGB2-DCS and NIG-DCS. These results indicate that the new DCS models with dynamic shape parameters may provide a more precise measurement of volatility dynamics than the DCS models with constant shape parameters from the body of literature, and they may motivate practical applications of the new DCS models with dynamic shape parameters.

In the remainder of this paper, Section 2 presents the new DCS models, Section 3 reviews the statistical inference, Section 4 presents the empirical results, and Section 5 concludes.

2. Econometric methods

2.1. DCS models with dynamic location, scale and shape parameters

We model the daily log-return on the S&P 500 index, $y_t = \ln(p_t/p_{t-1})$ for $t = 1, \ldots, T$, where p_t is closing price, adjusted for dividends and stock splits (for p_0, we use pre-sample data). The general form of all DCS models in this paper is

$$y_t = \mu_t + v_t = \mu_t + \exp(\lambda_t)\epsilon_t \tag{2.1}$$

where μ_t and $\exp(\lambda_t)$ are the dynamic location and scale parameters, respectively. For ϵ_t, we use the EGB2, NIG and Skew-Gen-t distributions (see Appendix A), which are all asymmetric probability distributions with several shape parameters (it is noteworthy that $E(y_t|y_1, \ldots, y_{t-1}) \neq \mu_t$, since $E(\epsilon_t|y_1, \ldots, y_{t-1}) \neq 0$). The k-th shape parameter is determined by a nonlinear transformation of the dynamic parameter $\rho_{k,t}$. To illustrate the flexibility of EGB2, NIG and Skew-Gen-t distributions, we present the density function for each model in Fig. 1, where we consider different shape parameters for each distribution and compare each probability distribution with the standard normal distribution.

The EGB2, NIG and Skew-Gen-t distributions use different dynamic tail shape for the left and right tails, in a similar manner to the works of Bollerslev and Todorov (2014) and Bollerslev et al. (2015).

The parameters μ_t, λ_t and $\rho_{k,t}$ of the DCS model are specified as follows. First, for the dynamic location parameter, we use DCS-QAR(1) (quasi-autoregressive) model (Harvey, 2013):

$$\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1} \tag{2.2}$$

where $|\phi| < 1$ and $u_{\mu,t}$ is the scaled score function of the log-likelihood (LL) with respect to μ_t (we present $u_{\mu,t}$ for the EGB2, NIG and Skew-Gen-t distributions in Appendix A). As an extension of this model, contemporaneous values and lags of exogenous explanatory variables may also be included on the right side of Equation (2.2). This location model can be related to the unobserved components models (UCMs) (Harvey, 1989), because a UCM is obtained by replacing the scaled score function by a Gaussian i.i.d. (independent and identically distributed) error term. The location equation is jointly estimated with the scale and shape equations, because in that way we control for possible dynamics of the expected return and also improve the measurement of volatility dynamics.

Second, the dynamic log-scale parameter is specified as:

$$\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} + \alpha^*\mathrm{sgn}(-\epsilon_{t-1})(u_{\lambda,t-1} + 1) \tag{2.3}$$

where $|\beta| < 1$, $u_{\lambda,t}$ is the score function of the LL with respect to λ_t (see $u_{\lambda,t}$ for the EGB2, NIG and Skew-Gen-t distributions in Appendix A), and $\mathrm{sgn}(x)$ is the signum function. Contemporaneous values and lags of exogenous explanatory variables may also be included on the right side of Equation (2.3). This specification measures leverage effects (i.e. effects of negative unexpected returns), by using parameter α^* in the DCS-EGARCH (exponential generalized autoregressive conditional heteroscedasticity) model (Harvey and Chakravarty, 2008). The DCS-EGARCH models with constant shape parameters that use EGB2, NIG and Skew-Gen-t distributions for ϵ_t are named EGB2-EGARCH (Caivano and Harvey, 2014), NIG-EGARCH (Blazsek et al., 2018), and Beta-Skew-Gen-t-EGARCH (Harvey and Lange, 2017), respectively.

Third, for the k-th dynamic parameter $\rho_{k,t}$ that determines the k-th dynamic shape parameter, we use the following DCS-QAR(1) model:

$$\rho_{k,t} = \delta_k + \gamma_k \rho_{k,t-1} + \kappa_k u_{\rho,k,t-1} \tag{2.4}$$

where $|\gamma_k| < 1$, and $u_{\rho,k,t}$ is the score function of the LL with respect to $\rho_{k,t}$ (we present $u_{\rho,k,t}$ for the EGB2, NIG and Skew-Gen-t distributions in Appendix A). As an extension of this model, contemporaneous values and lags of exogenous explanatory variables may also be included on the right side of Equation (2.4). For the EGB2 distribution, the two dynamic parameters that influence shape are denoted as $\rho_{1,t} = \xi_t$ and $\rho_{2,t} = \zeta_t$ (Appendix A). For the NIG distribution, the two dynamic parameters that influence shape are denoted as $\rho_{1,t} = \nu_t$ and $\rho_{2,t} = \eta_t$ (Appendix A). For the Skew-Gen-t distribution, the three dynamic parameters that influence shape are denoted as $\rho_{1,t} = \tau_t$, $\rho_{2,t} = \nu_t$ and $\rho_{3,t} = \eta_t$ (Appendix A). For each distribution we also use the constant shape parameter model as benchmark, for which $\rho_{k,t} = \delta_k$.

We consider different ways of initialization for each dynamic equation. For the results reported in this paper, we initialize μ_t by using pre-sample data, λ_t by parameter λ_0, and $\rho_{k,t}$ by using its unconditional mean $\delta_k/(1-\gamma_k)$. Nevertheless, our results are also robust to other ways of initialization. For example, we also use parameters μ_0 and $\rho_{k,0}$ for μ_t and $\rho_{k,t}$, respectively, and the corresponding results are similar to the results reported in this paper.

2.2. Specification tests

We perform specification tests for each probability distribution of the error term ϵ_t up to the fourth moment. For the EGB2 and NIG distributions, the first four conditional moments exist. For the Skew-Gen-t distribution, the degrees of freedom parameter specification ensures that the first four conditional moments exist (i.e. the degrees of freedom parameter is > 4).

For the EGB2-DCS model, $\epsilon_t \sim \text{EGB2}[0, 1, \exp(\xi_t), \exp(\zeta_t)]$, where both shape parameters are positive as required for the EGB2 distribution (Appendix A). The conditional mean, conditional variance, conditional skewness and conditional kurtosis of ϵ_t are given by:

$$E(\epsilon_t | y_1, \ldots, y_{t-1}; \Theta) = \Psi^{(0)}[\exp(\xi_t)] - \Psi^{(0)}[\exp(\zeta_t)] \tag{2.5}$$

$$\text{Var}(\epsilon_t | y_1, \ldots, y_{t-1}; \Theta) = \Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)] \tag{2.6}$$

$$\text{Skew}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \Psi^{(2)}[\exp(\xi_t)] - \Psi^{(2)}[\exp(\zeta_t)] \tag{2.7}$$

$$\text{Kurt}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \Psi^{(3)}[\exp(\xi_t)] + \Psi^{(3)}[\exp(\zeta_t)] \tag{2.8}$$

respectively; Θ is the vector of parameters and $\Psi^{(i)}(x)$ is the polygamma function of order i. For the NIG-DCS model, $\epsilon_t \sim \text{NIG}[0, 1, \exp(\nu_t), \exp(\nu_t)\tanh(\eta_t)]$, where $\tanh(x)$ is the hyperbolic tangent function, and the absolute value of parameter $\exp(\nu_t)\tanh(\eta_t)$ is less than parameter $\exp(\nu_t)$ as required for the NIG distribution (Appendix A). The conditional mean, conditional variance, conditional skewness and conditional kurtosis of ϵ_t are given by:

$$E(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \frac{\tanh(\eta_t)}{[1-\tanh^2(\eta_t)]^{1/2}} \tag{2.9}$$

$$\text{Var}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \frac{\exp(-\nu_t)}{[1-\tanh^2(\eta_t)]^{3/2}} \tag{2.10}$$

$$\text{Skew}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \frac{3\tanh(\eta_t)}{\exp(\nu_t/2)\left[1-\tanh^2(\eta_t)\right]^{1/4}} \tag{2.11}$$

$$\text{Kurt}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = 3 + \frac{3\left[1+4\tanh^2(\eta_t)\right]}{\exp(\nu_t)\left[1-\tanh^2(\eta_t)\right]^{1/2}} \tag{2.12}$$

respectively. For the Skew-Gen-t-DCS model, $\epsilon_t \sim$ Skew-Gen-$t[0, 1, \tanh(\tau_t), \exp(\nu_t){+}4, \exp(\eta_t)]$, where shape parameter $\tanh(\tau_t)$ is in the interval $(-1, 1)$ as required for the Skew-Gen-t distribution, degrees of freedom parameter $\exp(\nu_t){+}4$ is higher than four, and shape parameter $\exp(\eta_t)$ is positive as required for the Skew-Gen-t distribution (Appendix A). The conditional mean, conditional variance, conditional skewness and conditional kurtosis of ϵ_t, respectively, are:

$$E(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \frac{2\tanh(\tau_t)[\exp(\nu_t)+4]^{\exp(-\eta_t)} B\left\{\frac{2}{\exp(\eta_t)}, \frac{\exp(\nu_t)+3}{exp(\eta_t)}\right\}}{B\left\{\frac{1}{\exp(\eta_t)}, \frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right\}} \tag{2.13}$$

$$\text{Var}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = [\exp(\nu_t)+4]^{2\exp(-\eta_t)} \times \tag{2.14}$$

$$\times \left\{ \frac{[3\tanh^2(\tau_t)+1] B\left[\frac{3}{\exp(\eta_t)}, \frac{\exp(\nu_t)+2}{\exp(\eta_t)}\right]}{B\left[\frac{1}{\exp(\eta_t)}, \frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]} - \frac{4\tanh^2(\tau_t) B^2\left[\frac{2}{\exp(\eta_t)}, \frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]}{B^2\left[\frac{1}{\exp(\eta_t)}, \frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]} \right\}$$

$$\text{Skew}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta) = \frac{2\tanh(\tau_t)[\exp(\nu_t)+4]^{3\exp(-\eta_t)}}{B^3\left[\frac{1}{\exp(\eta_t)}, \frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]} \times \tag{2.15}$$

$$\times\Bigg\{8\tanh^2(\tau_t)B^3\left[\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]-3\left[1+3\tanh^2(\tau_t)\right]B\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]\times$$

$$\times B\left[\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]B\left[\frac{3}{\exp(\eta_t)},\frac{\exp(\nu_t)+2}{\exp(\eta_t)}\right]$$

$$+2\left[1+\tanh^2(\tau_t)\right]B^2\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]B\left[\frac{4}{\exp(\eta_t)},\frac{\exp(\nu_t)+1}{\exp(\eta_t)}\right]\Bigg\}$$

$$\mathrm{Kurt}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta)=\frac{[\exp(\nu_t)+4]^{4\exp(-\eta_t)}}{B^4\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]}\times \tag{2.16}$$

$$\times\Bigg\{-48\tanh^4(\tau_t)B^4\left[\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]$$

$$+24\tanh^2(\tau_t)\left[1+3\tanh^2(\tau_t)\right]B\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]B^2\left[\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]\times$$

$$\times B\left[\frac{3}{\exp(\eta_t)},\frac{\exp(\nu_t)+2}{\exp(\eta_t)}\right]-32\tanh^2(\tau_t)\left[1+\tanh^2(\tau_t)\right]B^2\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]\times$$

$$\times B\left[\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]B\left[\frac{4}{\exp(\eta_t)},\frac{\exp(\nu_t)+1}{\exp(\eta_t)}\right]$$

$$+\left[1+10\tanh^2(\tau_t)+5\tanh^4(\tau_t)\right]B^3\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]B\left[\frac{5}{\exp(\eta_t)},\frac{\exp(\nu_t)}{\exp(\eta_t)}\right]\Bigg\}$$

respectively; $B(x,y)=\Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the beta function and $\Gamma(x)$ is the gamma function.

In the remainder of this section, we use the previous mean, variance, skewness and kurtosis formulas for each equation. We define the auxiliary error term as:

$$\epsilon_t^*=\frac{\epsilon_t-E(\epsilon_t|y_1,\ldots,y_{t-1};\Theta)}{\mathrm{Var}^{1/2}(\epsilon_t|y_1,\ldots,y_{t-1};\Theta)}=\frac{\epsilon_t-E(\epsilon_t|\epsilon_1,\ldots,\epsilon_{t-1};\Theta)}{\mathrm{Var}^{1/2}(\epsilon_t|\epsilon_1,\ldots,\epsilon_{t-1};\Theta)} \tag{2.17}$$

This transformation reduces the importance of those outliers that appear within ϵ_t. For DCS models, outliers frequently appear within ϵ_t instead of within the updating terms of the dynamic equations; this is due to the outlier-discounting property of the score-functions. According to the work of Li (2004, Chapter 4), the robustness of specification tests is increased when residuals are standardized according to Equation (2.17). The conditional mean and conditional variance of ϵ_t^* are 0 and 1, respectively.

Therefore, we have:

$$E(\epsilon_t^*|y_1,\ldots,y_{t-1};\Theta)=E(\epsilon_t^*|\epsilon_1^*,\ldots,\epsilon_{t-1}^*;\Theta)=0 \tag{2.18}$$

$$E[(\epsilon_t^*)^2-1|y_1,\ldots,y_{t-1};\Theta]=E[(\epsilon_t^*)^2-1|\epsilon_1^*,\ldots,\epsilon_{t-1}^*;\Theta]=0 \tag{2.19}$$

By using the conditional skewness and conditional kurtosis formulas, we also have

$$E[(\epsilon_t^*)^3-\text{Skew}(\epsilon_t|y_1,\ldots,y_{t-1})|y_1,\ldots,y_{t-1};\Theta]= \tag{2.20}$$

$$=E[(\epsilon_t^*)^3-\text{Skew}(\epsilon_t|y_1,\ldots,y_{t-1})|\epsilon_1^*,\ldots,\epsilon_{t-1}^*;\Theta]=0$$

$$E[(\epsilon_t^*)^4-\text{Kurt}(\epsilon_t|y_1,\ldots,y_{t-1})|y_1,\ldots,y_{t-1};\Theta]= \tag{2.21}$$

$$=E[(\epsilon_t^*)^4-\text{Kurt}(\epsilon_t|y_1,\ldots,y_{t-1})|\epsilon_1^*,\ldots,\epsilon_{t-1}^*;\Theta]=0$$

Within the expectations of Equations (2.18) to (2.21), variables with MDS property appear. We use the MDS test with optimal lag-order selection from the work of Escanciano and Lobato (2009), to verify the correct specification for each probability distribution.

3. Statistical inference

We estimate all models of this paper by using the ML method. With respect to DCS models, Blasques et al. (2017, 2018) present the conditions for the asymptotic properties of ML for DCS models with a single score-driven dynamic variable. In this paper, we present those conditions for DCS models with several score-driven dynamic variables. The first representation of the ML estimator is

$$\hat{\Theta}_{\text{ML}}=\arg\max_{\Theta}\text{LL}(y_1,\ldots,y_T;\Theta)=\arg\max_{\Theta}\frac{1}{T}\sum_{t=1}^{T}\ln f(y_t|y_1,\ldots,y_{t-1};\Theta) \tag{3.1}$$

where $\Theta=(\Theta_1,\ldots,\Theta_K)'$ is the vector of parameters and $\ln f(y_t|y_1,\ldots,y_{t-1};\Theta)$ for the EGB2, NIG and Skew-Gen-t distributions is presented in Appendix A. We assume: (A1) The LL function is correctly specified for each DCS model. (A2) The vector of true values of parameters Θ_0 is an interior point within a compact parameter set in $\mathbb{R}^K$. (A3) $\hat{\Theta}_{\text{ML}}$ is a unique solution to the maximization problem of Equation (3.1). (A4) $\text{LL}(\cdot;\Theta)$ is a Borel measurable function on $\mathbb{R}^T$. (A5) For each $(y_1,\ldots,y_T)\in\mathbb{R}^T$, $\text{LL}(y_1,\ldots,y_T;\cdot)$ is a continuous function on the parameter set. (A6) $\text{LL}(y_1,\ldots,y_T;\Theta)$ is twice continuously differentiable on all of the interior points of the parameter set.

We define the $T \times K$ matrix of contributions to the gradient $G(y_1, \ldots, y_T, \Theta)$ by its elements:

$$G_{ti}(\Theta) = \frac{\partial \ln f(y_t|y_1, \ldots, y_{t-1}; \Theta)}{\partial \Theta_i} \tag{3.2}$$

for period $t = 1, \ldots, T$ and parameter $i = 1, \ldots, K$. We denote the t-th row of $G(y_1, \ldots, y_T, \Theta)$ by using $G_t(\Theta)$, which is the score vector for the t-th observation. Under (A1) to (A6), the first representation of the ML estimator is equivalent to the following second representation:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta}_{\text{ML}})' = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} G_{t1}(\hat{\Theta}_{\text{ML}}) \\ \vdots \\ G_{tK}(\hat{\Theta}_{\text{ML}}) \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};p_0,\hat{\Theta}_{\text{ML}})}{\partial \Theta_1} \\ \vdots \\ \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};p_0,\hat{\Theta}_{\text{ML}})}{\partial \Theta_K} \end{bmatrix} = 0_{K\times 1} \tag{3.3}$$

We write the left side of Equation (3.3) according to the mean-value expansion about the true values of parameters Θ_0, as follows:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta}_{\text{ML}})' = \frac{1}{T}\sum_{t=1}^{T} G_t(\Theta_0)' + \frac{1}{T}\left[\sum_{t=1}^{T} H_t(\bar{\Theta})\right](\hat{\Theta}_{\text{ML}} - \Theta_0) \tag{3.4}$$

where each row of the $K \times K$ Hessian matrix

$$H_t(\Theta) = \frac{\partial^2 \ln f(y_t|y_1, \ldots, y_{t-1}; \Theta)}{\partial \Theta \Theta'} \tag{3.5}$$

of the t-th observation is evaluated at K different mean values, indicated by $\bar{\Theta}$. Each $\bar{\Theta}$ is located between Θ_0 and $\hat{\Theta}_{\text{ML}}$, which can be more formally expressed as: $||\bar{\Theta} - \Theta_0|| \leq ||\hat{\Theta}_{\text{ML}} - \Theta_0||$, where $||\cdot||$ is the Euclidean norm. We define the $K \times K$ contribution to the information matrix for period t that is evaluated at the true values of parameters, as follows:

$$I_t(\Theta_0) = -E[H_t(\Theta_0)|y_1, \ldots, y_{t-1}] = E[G_t(\Theta_0)'G_t(\Theta_0)|y_1, \ldots, y_{t-1}] \tag{3.6}$$

The second equality in Equation (3.6) is based on the conditional information matrix equality, which holds under assumption (A1). From Equations (3.3) and (3.4), we express:

$$\sqrt{T}(\hat{\Theta}_{\mathrm{ML}} - \Theta_0) = \left[-\frac{1}{T}\sum_{t=1}^{T} H_t(\bar{\Theta})\right]^{-1} \left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right] \tag{3.7}$$

We study conditions for the following asymptotic result:

$$\sqrt{T}(\hat{\Theta}_{\mathrm{ML}} - \Theta_0) \to_d N_K\left[0_{K\times 1}, \mathcal{I}^{-1}(\Theta_0)\right] \quad \text{as} \quad T \to \infty \tag{3.8}$$

where $\mathcal{I}(\Theta_0) \equiv E[I_t(\Theta_0)]$ and $\mathcal{I}^{-1}(\Theta_0)$ is positive definite. The asymptotic covariance matrix of $\hat{\Theta}_{\mathrm{ML}}$ is $\mathcal{I}^{-1}(\Theta_0)/T$, which we estimate by using $[\sum_{t=1}^{T} G_t(\hat{\Theta}_{\mathrm{ML}})' G_t(\hat{\Theta}_{\mathrm{ML}})]^{-1}$.

3.1. Information matrix

We study the conditions of the finiteness of all of the elements of $\mathcal{I}(\Theta_0)$. For ease of notation, we start with a DCS model with score-driven μ_t, score-driven λ_t and known constant shape parameters for ϵ_t:

$$y_t = \mu_t + \exp(\lambda_t)\epsilon_t \tag{3.9}$$

$$\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1} \tag{3.10}$$

$$\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} \tag{3.11}$$

We start with the conditions of covariance stationarity of y_t. We re-parameterize μ_t and λ_t, by using the unconditional means $E(\mu_t) = \tilde{c} = c/(1-\phi)$ and $E(\lambda_t) = \tilde{\omega} = \omega/(1-\beta)$, as follows:

$$\mu_t = \tilde{c}(1-\phi) + \phi\mu_{t-1} + \theta u_{\mu,t-1} \tag{3.12}$$

$$\lambda_t = \tilde{\omega}(1-\beta) + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} \tag{3.13}$$

for which $\Theta = (\tilde{c}, \phi, \theta, \tilde{\omega}, \beta, \alpha)$ and $K = 6$. We use these alternative forms, because the information matrix is simpler under these forms. The conditions for the covariance stationarity of y_t are $|\phi| < 1$ and $|\beta| < 1$. We name these conditions as **Condition 1**.

We study the conditions of the finiteness of all of the elements of $\mathcal{I}(\Theta_0)$. We express:

$$\mathcal{I}(\Theta_0) = E[I_t(\Theta_0)] = E\left\{E[G_t(\Theta_0)' G_t(\Theta_0)|y_1, \ldots, y_{t-1}]\right\} = E[G_t(\Theta_0)' G_t(\Theta_0)] \tag{3.14}$$

In the following, we present the conditions under which all of the elements of $E[G_t(\Theta_0)'G_t(\Theta_0)]$ are finite. We express the elements of $G_t(\Theta_0)'$, according to the chain rule, as follows:

$$G_t(\Theta_0)' = \begin{bmatrix} \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\theta} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\phi} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\tilde{c}} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\alpha} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\beta} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\tilde{\omega}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial\mu_t}{\partial\theta} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial\mu_t}{\partial\phi} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial\mu_t}{\partial\tilde{c}} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\lambda_t} \times \frac{\partial\lambda_t}{\partial\alpha} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\lambda_t} \times \frac{\partial\lambda_t}{\partial\beta} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\lambda_t} \times \frac{\partial\lambda_t}{\partial\tilde{\omega}} \end{bmatrix} \tag{3.15}$$

We define four panels within the contribution to the information matrix:

$$I_t(\Theta_0) = E_{t-1}[G_t(\Theta_0)'G_t(\Theta_0)] = E_{t-1}\begin{bmatrix} A_{(3\times3)} & C_{(3\times3)} \\ C_{(3\times3)} & B_{(3\times3)} \end{bmatrix} \tag{3.16}$$

where within panel A elements that involve only the derivatives of μ_t appear, in panel B elements that involve only the derivatives of λ_t appear, and in panel C elements that involve the derivatives of both μ_t and λ_t appear. We factorize out four scalars from A, B and each of the two C panels, respectively, within Equation (3.16), and we define the following 2×2 matrix:

$$I = \begin{Bmatrix} \left[\frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\mu_t}\right]^2 & \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\lambda_t} \\ \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\lambda_t} & \left[\frac{\partial \ln f(y_t|y_1,...,y_{t-1};\Theta_0)}{\partial\lambda_t}\right]^2 \end{Bmatrix} \tag{3.17}$$

Given the factorization with respect to I, we can also write Equation (3.16) as:

$$I_t(\Theta_0) = E_{t-1}(I) \circ D(\Theta_0) = E_{t-1}(I) \circ \begin{bmatrix} \tilde{A}_{(3\times3)} & \tilde{C}_{(3\times3)} \\ \tilde{C}_{(3\times3)} & \tilde{B}_{(3\times3)} \end{bmatrix} \tag{3.18}$$

where $\circ$ denotes the Hadamard product and the panels of the matrix within $D(\Theta_0)$ are given by the outer product of the 1×6 vector

$$\tilde{D} = [(\partial\mu_t/\partial\theta), (\partial\mu_t/\partial\phi), (\partial\mu_t/\partial\tilde{c}), (\partial\lambda_t/\partial\alpha), (\partial\lambda_t/\partial\beta), (\partial\lambda_t/\partial\tilde{\omega})] \tag{3.19}$$

with itself, i.e. $D(\Theta_0) = \tilde{D}'\tilde{D}$. In Equation (3.18), the expectation no longer appears for the matrix formed by $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$, because μ_t and λ_t are not randomly assigned conditional on $(y_1, \ldots, y_{t-1})$. By using the scaled score function and the score function, we write I as:

$$I = \begin{bmatrix} u_{\mu,t}^2/k_t^2 & u_{\mu,t} \times u_{\lambda,t}/k_t \\ u_{\mu,t} \times u_{\lambda,t}/k_t & u_{\lambda,t}^2 \end{bmatrix} \tag{3.20}$$

where k_t is the time-varying scaling parameter. The form of k_t is different for different DCS models. For example, from Equation (A.4) of Appendix A, for EGB2-DCS we have:

$$\frac{\partial \ln f(y_t|y_1, \ldots, y_{t-1}; \Theta_0)}{\partial \mu_t} = u_{\mu,t} \times \{\Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)]\} \exp(2\lambda_t) = \frac{u_{\mu,t}}{k_t} \tag{3.21}$$

For k_t of NIG-DCS and Skew-Gen-t-DCS, see Equations (A.14) and (A.22) of Appendix A, respectively. Based on Equations (3.18) and (3.20), we need the unconditional means of $(u_{\mu,t}^2/k_t^2)$, $u_{\lambda,t}^2$ and $(u_{\mu,t} \times u_{\lambda,t}/k_t)$ to be finite. We name these conditions as **Condition 2**.

We study $E[D(\Theta_0)] < \infty$, by using the following dynamic equations:

$$\tilde{D}' = \begin{bmatrix} \frac{\partial \mu_t}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \phi} \\ \frac{\partial \mu_t}{\partial \tilde{c}} \\ \frac{\partial \lambda_t}{\partial \alpha} \\ \frac{\partial \lambda_t}{\partial \beta} \\ \frac{\partial \lambda_t}{\partial \tilde{\omega}} \end{bmatrix} = \begin{bmatrix} X_{\mu,t-1} \times \frac{\partial \mu_{t-1}}{\partial \theta} + u_{\mu,t-1} \\ X_{\mu,t-1} \times \frac{\partial \mu_{t-1}}{\partial \phi} + \mu_{t-1} - \tilde{c} \\ X_{\mu,t-1} \times \frac{\partial \mu_{t-1}}{\partial \tilde{c}} + 1 - \phi \\ X_{\lambda,t-1} \times \frac{\partial \lambda_{t-1}}{\partial \alpha} + u_{\lambda,t-1} \\ X_{\lambda,t-1} \times \frac{\partial \lambda_{t-1}}{\partial \beta} + \lambda_{t-1} - \tilde{\omega} \\ X_{\lambda,t-1} \times \frac{\partial \lambda_{t-1}}{\partial \tilde{\omega}} + 1 - \beta \end{bmatrix} \tag{3.22}$$

where $X_{\mu,t} = \phi + \theta(\partial u_{\mu,t}/\partial \mu_t)$ and $X_{\lambda,t} = \beta + \alpha(\partial u_{\lambda,t}/\partial \lambda_t)$. With respect to $D(\Theta_0)$, in panel $\tilde{A}$ elements that involve only the derivatives of μ_t appear, in panel $\tilde{B}$ elements that involve only the derivatives of λ_t appear, and in panel $\tilde{C}$ elements that involve the derivatives of both μ_t and λ_t appear. Equation (3.22) provides the following conditions for $E[D(\Theta_0)] < \infty$: For panel $\tilde{A}$ it is necessary that $E(X_{\mu,t}^2) < 1$ and for panel $\tilde{B}$ it is necessary that $E(X_{\lambda,t}^2) < 1$ (for these results, see Harvey, 2013). With respect to panel $\tilde{C}$, it is necessary that $|E(X_{\mu,t}X_{\lambda,t})| < 1$ (we present the proof for panel $\tilde{C}$ in Appendix B). Moreover, it is also necessary that the unconditional means of $X_{\mu,t}$, $X_{\lambda,t}$, $u_{\mu,t}$ and

$u_{\lambda,t}$ are finite, and that the unconditional mean of each product formed by all possible pairs of those variables are also finite (Appendix B). We name these conditions as **Condition 3**. As a consequence, all of the elements of the information matrix are finite under Conditions 1 to 3.

3.2. Central limit theorem (CLT) for the score vector

We focus on the asymptotic properties of following term from Equation (3.6):

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)' \quad \text{as} \quad T \to \infty. \tag{3.23}$$

We use the CLT from the work of White (1984, Theorem 5.15): If (E1) $G_t(\Theta_0)'$ is strictly stationary and ergodic, (E2) all of the elements of the unconditional covariance matrix of $G_t(\Theta_0)'$ are finite, and (E3) $E[G_1(\Theta_0)'|y_0, y_{-1}, \ldots, y_{-t}] \to_{\text{q.m.}} 0$ as $t \to \infty$ (q.m. is quadratic mean), then Equation (3.23) converges in distribution to the normal distribution.

(E1) With respect to stationarity and ergodicity, we refer to the works of Brandt (1986) and Diaconis and Freedman (1999). First, we write Equation (3.22) as:

$$\begin{bmatrix} \frac{\partial \mu_t}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \phi} \\ \frac{\partial \mu_t}{\partial \tilde{c}} \\ \frac{\partial \lambda_t}{\partial \alpha} \\ \frac{\partial \lambda_t}{\partial \beta} \\ \frac{\partial \lambda_t}{\partial \tilde{\omega}} \end{bmatrix} = \begin{bmatrix} X_{\mu,t-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & X_{\mu,t-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & X_{\mu,t-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & X_{\lambda,t-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & X_{\lambda,t-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & X_{\lambda,t-1} \end{bmatrix} \begin{bmatrix} \frac{\partial \mu_{t-1}}{\partial \theta} \\ \frac{\partial \mu_{t-1}}{\partial \phi} \\ \frac{\partial \mu_{t-1}}{\partial \tilde{c}} \\ \frac{\partial \lambda_{t-1}}{\partial \alpha} \\ \frac{\partial \lambda_{t-1}}{\partial \beta} \\ \frac{\partial \lambda_{t-1}}{\partial \tilde{\omega}} \end{bmatrix} + \begin{bmatrix} u_{\mu,t-1} \\ \mu_{t-1} - \tilde{c} \\ 1 - \phi \\ u_{\lambda,t-1} \\ \lambda_{t-1} - \tilde{\omega} \\ 1 - \beta \end{bmatrix} \tag{3.24}$$

We use the following compact notation for the previous equation:

$$Y_t = A^*_{t-1} Y_{t-1} + B^*_{t-1} \tag{3.25}$$

which is a stochastic recurrence equation (SRE). **Condition 4** is that ϵ_t is strictly stationary and ergodic. We refer to the work of White (1984, Theorem 3.35), in which a possibly nonlinear measurable function transforms strictly stationary and ergodic variables to new strictly stationary and ergodic variables. Conditions 1 to 4 imply that $u_{\mu,t}$, $u_{\lambda,t}$, $\partial u_{\mu,t}/\partial \mu_t$, $\partial u_{\lambda,t}/\partial \lambda_t$, μ_t and λ_t are strictly stationary and ergodic, because they are transformations of ϵ_t. As a consequence, A^*_t and B^*_t are strictly stationary and ergodic (see, for example, Harvey, 2013). From $E(X_{\mu,t}^2) < 1$ and $E(X_{\lambda,t}^2) < 1$ of Condition 3 and by using the Cauchy–Schwarz inequality, we have $|E(X_{\mu,t})| < 1$ and $|E(X_{\lambda,t})| < 1$. Based on the results of Brandt (1986) and Diaconis and Freedman (1999), we have that Y_t is strictly stationary and ergodic (i.e. $\tilde{D}'$ is strictly stationary and ergodic). Second, we write Equation (3.15), as follows:

$$G_t(\Theta_0)' = \begin{bmatrix} \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta_0)}{\partial \mu_t} \\ \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta_0)}{\partial \lambda_t} \end{bmatrix} \circ \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix} = \begin{bmatrix} u_{\mu,t}/k_t \\ u_{\lambda,t} \end{bmatrix} \circ \begin{bmatrix} A^\dagger \\ B^\dagger \end{bmatrix} \tag{3.26}$$

where in panel $A^\dagger$ elements that involve only the derivatives of μ_t appear from $\tilde{D}'$ of Equation (3.22), and in panel $B^\dagger$ elements that involve only the derivatives of λ_t appear from $\tilde{D}'$ of Equation (3.22). Under Conditions 1 to 4, $u_{\mu,t}/k_t$ and $u_{\lambda,t}$ are strictly stationary and ergodic. In Equation (3.26), $G_t(\Theta_0)'$ is represented as a product of strictly stationary and ergodic variables. Thus, according to the theorem of White (1984, Theorem 3.35), $G_t(\Theta_0)'$ is also strictly stationary and ergodic.

(E2) The unconditional covariance matrix of $G_t(\Theta_0)'$ is the information matrix, for which the conditions of finiteness were shown in the previous section.

(E3) With respect to $E[G_1(\Theta_0)'|y_0, y_{-1}, \ldots, y_{-t}] \to_{\text{q.m.}} 0$ as $t \to \infty$: this property holds under Conditions 1 to 3 and due to $|E(X_{\mu,t})| < 1$ and $|E(X_{\lambda,t})| < 1$ in Equation (3.22). As a consequence, the CLT 5.15 of White (2004) holds for the score vector under Conditions 1 to 4.

Theorem 1: For the model of Equations (3.9) to (3.11), with a likelihood function (3.1) satisfying the regularity conditions (A1) to (A6), $\sqrt{T}(\hat{\Theta}_{\text{ML}} - \Theta_0) \to_d N[0_{K\times 1}, \mathcal{I}^{-1}(\Theta_0)]$ as $T \to \infty$ under the following conditions: Condition 1 is that $|\phi| < 1$ and $|\beta| < 1$. Condition 2 is that the unconditional means of $(u_{\mu,t}^2/k_t^2)$, $u_{\lambda,t}^2$ and $(u_{\mu,t} \times u_{\lambda,t}/k_t)$ are finite. Condition 3 is that $E(X_{\mu,t}^2) < 1$, $E(X_{\lambda,t}^2) < 1$ and $|E(X_{\mu,t}X_{\lambda,t})| < 1$, where $X_{\mu,t} = \phi + \theta(\partial u_{\mu,t}/\partial \mu_t)$ and $X_{\lambda,t} = \beta + \alpha(\partial u_{\lambda,t}/\partial \lambda_t)$. Under Condition 3, the unconditional means of $X_{\mu,t}$, $X_{\lambda,t}$, $u_{\mu,t}$ and $u_{\lambda,t}$ are finite, and the unconditional mean of each product formed by all possible pairs of those variables is finite. Condition 4 is that ϵ_t is strictly stationary and ergodic.

Conditions 1 to 4 can be extended to models with several score-driven parameters. We present the ML conditions for the EGB2-DCS model with $\epsilon_t \sim \text{EGB2}[0, 1, \exp(\xi_t), \exp(\zeta_t)]$:

$$y_t = \mu_t + \exp(\lambda_t)\epsilon_t \tag{3.27}$$

$$\mu_t = c + \phi \mu_{t-1} + \theta u_{\mu,t-1} \tag{3.28}$$

$$\lambda_t = \omega + \beta \lambda_{t-1} + \alpha u_{\lambda,t-1} \tag{3.29}$$

$$\xi_t = \delta_1 + \gamma_1 \xi_{t-1} + \kappa_1 u_{\xi,t-1} \tag{3.30}$$

$$\zeta_t = \delta_2 + \gamma_2 \zeta_{t-1} + \kappa_2 u_{\zeta,t-1} \tag{3.31}$$

Theorem 2: For the model of Equations (3.27) to (3.31), with a likelihood function (3.1) satisfying the regularity conditions (A1) to (A6), $\sqrt{T}(\hat{\Theta}_{\mathrm{ML}} - \Theta_0) \rightarrow_d N[0_{K\times 1}, \mathcal{I}^{-1}(\Theta_0)]$ as $T \rightarrow \infty$ under the following conditions: Condition 1 is that $|\phi| < 1$, $|\beta| < 1$, $|\gamma_1| < 1$ and $|\gamma_2| < 1$. Condition 2 is that the unconditional means of $(u_{\mu,t}^2/k_t^2)$, $u_{\lambda,t}^2$, $u_{\xi,t}^2$, $u_{\zeta,t}^2$, $(u_{\mu,t} \times u_{\lambda,t}/k_t)$, $(u_{\mu,t} \times u_{\xi,t}/k_t)$, $(u_{\mu,t} \times u_{\zeta,t}/k_t)$, $(u_{\lambda,t} \times u_{\xi,t})$, $(u_{\lambda,t} \times u_{\zeta,t})$, and $(u_{\xi,t} \times u_{\zeta,t})$ are finite. Condition 3 is that $E(X_{\mu,t}^2) < 1$, $E(X_{\lambda,t}^2) < 1$, $E(X_{\xi,t}^2) < 1$, $E(X_{\zeta,t}^2) < 1$, $|E(X_{\mu,t}X_{\lambda,t})| < 1$, $|E(X_{\mu,t}X_{\xi,t})| < 1$, $|E(X_{\mu,t}X_{\zeta,t})| < 1$, $|E(X_{\lambda,t}X_{\xi,t})| < 1$, $|E(X_{\lambda,t}X_{\zeta,t})| < 1$ and $|E(X_{\xi,t}X_{\zeta,t})| < 1$, where $X_{\mu,t} = \phi + \theta(\partial u_{\mu,t}/\partial \mu_t)$, $X_{\lambda,t} = \beta + \alpha(\partial u_{\lambda,t}/\partial \lambda_t)$, $X_{\xi,t} = \gamma_1 + \kappa_1(\partial u_{\xi,t}/\partial \xi_t)$ and $X_{\zeta,t} = \gamma_2 + \kappa_2(\partial u_{\zeta,t}/\partial \zeta_t)$. Under Condition 3, the unconditional means of $X_{\mu,t}$, $X_{\lambda,t}$, $X_{\xi,t}$, $X_{\zeta,t}$, $u_{\mu,t}$, $u_{\lambda,t}$, $u_{\xi,t}$ and $u_{\zeta,t}$ are finite, and the unconditional mean of each product formed by all possible pairs of those variables is also finite. Condition 4 is that ϵ_t is strictly stationary and ergodic.

4. Empirical results

We use daily log-return data from the adjusted S&P 500 index p_t for the period of January 4, 1950 to December 30, 2017 (source: Bloomberg). Descriptive statistics of y_t are presented in Table 1. The negative skewness estimate indicates that the mass of the distribution of y_t is concentrated on the right side, and the high excess kurtosis estimate suggests heavy tails of y_t. The negative correlation coefficient Corr(y_t^2,y_{t-1}) suggests that high volatility often follows significant negative returns, which motivates the consideration of leverage effects within λ_t.

In the remainder of this section, we present the ML results and model diagnostics for the EGB2-DCS (Table 2), NIG-DCS (Table 3) and Skew-Gen-t-DCS (Table 4(a) and Table 4(b)) models. We compare the LL-based performance of those models in Table 5. We present the evolution of $\rho_{k,t}$ for all k and the evolution of λ_t for all DCS specifications in Figs. 2, 3, 4(a) and 4(b). We highlight the dates of extreme events identified by one of the DCS specifications in Fig. 5. We study the outlier-discounting properties for all score functions in Fig. 6.

With respect to the parameters of location, scale and shape dynamics and the parameters of the corresponding updating terms, in Tables 2 to 4 we present the following results. For most of the specifications, we find that ϕ parameter which measures the dynamics of conditional location is significantly different from zero. The scaling parameter of the score function with respect to location θ is positive and significant for all models. For all of the specifications, we find highly significant ω, α, α^* and β parameters for conditional scale. For most of the specifications, we find that the dynamic parameters of shape (i.e. γ_1, γ_2 and γ_3) are significant and positive. We also find for all of the specifications that the scaling parameter of the updating term for shape (i.e. κ_1, κ_2 and κ_3) is significantly different from zero. For all DCS equations we find a significant parameter for the updating term, i.e. the DCS models are identified (Harvey, 2013).

In Tables 2 to 4, we report the estimates of ϕ, β, γ_1, γ_2 and γ_3. All of those estimates are less than one in absolute value, thus, Condition 1 is supported for the ML estimates. Moreover, in Tables 2 to 4, we report the estimates of $C_\mu = E(X^2_{\mu,t})$, $C_\lambda = E(X^2_{\lambda,t})$, $C_{\rho,1} = E(X^2_{\rho,1,t})$, $C_{\rho,2} = E(X^2_{\rho,2,t})$, $C_{\rho,3} = E(X^2_{\rho,3,t})$, $C_{\mu,\lambda} = |E(X_{\mu,t}X_{\lambda,t})|$, $C_{\mu,\rho,1} = |E(X_{\mu,t}X_{\rho,1,t})|$, $C_{\mu,\rho,2} = |E(X_{\mu,t}X_{\rho,2,t})|$, $C_{\mu,\rho,3} = |E(X_{\mu,t}X_{\rho,3,t})|$, $C_{\lambda,\rho,1} = |E(X_{\lambda,t}X_{\rho,1,t})|$, $C_{\lambda,\rho,2} = |E(X_{\lambda,t}X_{\rho,2,t})|$, $C_{\lambda,\rho,3} = |E(X_{\lambda,t}X_{\rho,3,t})|$, $C_{\rho,1,\rho,2} = |E(X_{\rho,1,t}X_{\rho,2,t})|$, $C_{\rho,1,\rho,3} = |E(X_{\rho,1,t}X_{\rho,3,t})|$ and $C_{\rho,2,\rho,3} = |E(X_{\rho,2,t}X_{\rho,3,t})|$. All of those estimates are less than one, thus, the corresponding formulas of Condition 3 are supported for the ML estimates. For the variables of Conditions 2 and 3, we perform the augmented Dickey–Fuller (1979) (hereinafter, ADF) unit root test with constant. We find that Conditions 2 and 3 are supported for the ML estimates. We do not report the ADF test results in this paper, but those are available from the authors on request. Condition 4 is a maintained assumption for all DCS models of this paper.

In Tables 2 to 4, we report diagnostic test results with respect to the correct specification of the error term up to the fourth moment (see Section 2.2). For EGB2-DCS, we consider four different specifications with respect to dynamic versus constant shape parameters, and we find that all of those specifications fail the MDS test with respect to skewness and kurtosis. For NIG-DCS, we consider four different specifications with respect to dynamic versus constant shape parameters, and we find that all of those specifications pass the MDS test up to the fourth moment. For Skew-Gen-t-DCS, we consider eight different specifications with respect to dynamic versus constant shape parameters, and we find that four out of eight specifications pass the MDS test up to the fourth moment.

We compare model performance by using the following model performance metrics: LL, Akaike information criterion (AIC), Bayesian information criterion (BIC), and Hannan-Quinn criterion (HQC). We find that the AIC-, BIC- and HQC-based statistical performances of the DCS model with dynamic shape are superior to the statistical performance of the corresponding DCS model with constant shape (Tables 2 to 5). We also compare model performance by using the likelihood-ratio (LR) test for non-nested models (Vuong, 1989). We denote the conditional density functions of y_t of the DCS models with dynamic and constant shape by using $f(y_t|y_1,\ldots,y_{t-1})$ and $g(y_t|y_1,\ldots,y_{t-1})$, respectively. We define $d_t = \ln f(y_t|y_1,\ldots,y_{t-1}) - \ln g(y_t|y_1,\ldots,y_{t-1})$ for $t = 1,\ldots,T$. We test whether LL of DCS with dynamic shape is superior to that of DCS with constant shape by estimating $d_t = c + \epsilon_t$ with the OLS-HAC (ordinary least squares-heteroskedasticity and autocorrelation consistent; Newey and West, 1987) estimator. If c is significantly positive then DCS with dynamic shape is superior to DCS with constant shape. We always find that at least one of the DCS specifications with dynamic shape parameters is superior to the DCS specification with constant shape parameters (Tables 2 to 5). We find that the statistical performance of the Skew-Gen-t-DCS model is superior to the statistical performances of the EGB2-DCS and NIG-DCS models (Tables 2 to 5).

We present the evolution of $\rho_{k,t}$ for all k and the evolution of λ_t in Figs. 2 to 4. Those figures indicate the following: (i) the shape parameters are time-varying for all DCS models; (ii) for the DCS models with dynamic shape parameters, the shape parameters identify the dates of several extreme events. We analyze the identification of extreme events for one of the best performing DCS specification with dynamic shape parameters (Table 5), and we present evolution of ν_t for the DCS-Skew-Gen-t specification with constant τ_t, dynamic ν_t and constant η_t in Fig. 5. In Fig. 5, we number several days for which ν_t is relatively low, indicating that the probability of an extreme observation is high for those days. We investigated what extreme events caused the high tail thickness of the S&P 500. Interestingly, we were able to associate all of those days with important events that significantly impacted the US stock market (see Appendix C). As aforementioned, a general property of all DCS models is that outliers are discounted by the score functions in the dynamic equations. We investigate whether this property is also true for the new DCS models with score-driven shape parameters. In Fig. 6 we present the score functions, which indicate different ways of discounting outliers for different DCS models. Extreme observations are never accentuated by score functions in the new DCS models.

5. Conclusions

We have suggested new DCS models of conditional volatility, for which both the scale and shape parameters are dynamic. Our models extend previous volatility models with constant shape parameters from the literature, since: (i) they have a superior likelihood-based statistical performance; (ii) they estimate the dynamics of both scale and shape parameters effectively; (iii) news on asset value updates the distribution of financial return not only through scale, but also through shape; (iv) they use different dynamic tail shape for the left and right tails of the return distribution; (v) they identify extreme events effectively; (vi) we have provided the conditions for the asymptotic properties of ML for the new DCS models.

We have introduced new DCS volatility models for the EGB2, NIG, and Skew-Gen-t distributions with dynamic shape parameters. We have used return time series data from the adjusted S&P 500 index for the period of January 4, 1950 to December 30, 2017. We have estimated all DCS models by using the ML method, and we have presented the conditions of the asymptotic properties of the ML estimator. We have found that the likelihood-based performance of Skew-Gen-t-DCS is superior to the likelihood-based performances of EGB2-DCS and NIG-DCS. We have also found that the score-driven dynamics of shape parameters are significant, and we have shown that the likelihood-based performance of the new DCS models is superior to that of the DCS models with constant shape parameters. Our results may motivate practical applications of the new DCS models with dynamic shape parameters, for example, for the estimation of stock market volatility, for pricing financial derivatives, or for the estimation of risk measurement metrics, such as VaR and ES.

Appendix A

In this appendix, for each error specification, we present the conditional distribution of y_t, the conditional mean of y_t, the conditional volatility of y_t, the log of the conditional density of y_t, the scaled score function for location $u_{\mu,t}$, and the score functions for scale $u_{\lambda,t}$ and shape $u_{\rho,k,t}$.

(1) For the EGB2-DCS model, the conditional distribution of y_t is EGB2$[\mu_t, \exp(-\lambda_t), \exp(\xi_t), \exp(\zeta_t)]$. The conditional mean and the conditional standard deviation (SD) (i.e. conditional volatility) of y_t are

$$E(y_t|y_1, \ldots, y_{t-1}; \Theta) = \mu_t + \exp(\lambda_t) \left\{ \Psi^{(0)}[\exp(\xi_t)] - \Psi^{(0)}[\exp(\zeta_t)] \right\} \quad \text{(A.1)}$$

$$SD(y_t|y_1, \ldots, y_{t-1}; \Theta) = \exp(\lambda_t)\{\Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)]\}^{1/2} \quad \text{(A.2)}$$

respectively. The log of the conditional density of y_t is

$$\ln f(y_t|y_1,\ldots,y_{t-1};\Theta) = \exp(\xi_t)\epsilon_t - \lambda_t - \ln\Gamma[\exp(\xi_t)] - \ln\Gamma[\exp(\zeta_t)] \tag{A.3}$$

$$+\ln\Gamma[\exp(\xi_t)+\exp(\zeta_t)] - [\exp(\xi_t)+\exp(\zeta_t)]\ln[1+\exp(\epsilon_t)]$$

The score functions with respect to μ_t, λ_t, ξ_t and ζ_t are as follows. First, the score function with respect to μ_t is

$$\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\mu_t} = u_{\mu,t}\times\{\Psi^{(1)}[\exp(\xi_t)]+\Psi^{(1)}[\exp(\zeta_t)]\}\exp(2\lambda_t) = \frac{u_{\mu,t}}{k_t} \tag{A.4}$$

where

$$u_{\mu,t} = \{\Psi^{(1)}[\exp(\xi_t)]+\Psi^{(1)}[\exp(\zeta_t)]\}\exp(\lambda_t)\left\{[\exp(\xi_t)+\exp(\zeta_t)]\frac{\exp(\epsilon_t)}{\exp(\epsilon_t)+1} - \exp(\xi_t)\right\} \tag{A.5}$$

is the scaled score function. Second, the score function with respect to λ_t is

$$u_{\lambda,t} = \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\lambda_t} = [\exp(\xi_t)+\exp(\zeta_t)]\frac{\epsilon_t\exp(\epsilon_t)}{\exp(\epsilon_t)+1} - \exp(\xi_t)\epsilon_t - 1 \tag{A.6}$$

Third, the score function with respect to ξ_t is

$$u_{\xi,t} = \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\xi_t} = \exp(\xi_t)\epsilon_t - \exp(\xi_t)\Psi^{(0)}[\exp(\xi_t)] \tag{A.7}$$

$$+\exp(\xi_t)\Psi^{(0)}[\exp(\xi_t)+\exp(\zeta_t)] - \exp(\xi_t)\ln[1+\exp(\epsilon_t)]$$

Fourth, the score function with respect to ζ_t is

$$u_{\zeta,t} = \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\zeta_t} = -\exp(\zeta_t)\Psi^{(0)}[\exp(\zeta_t)] \tag{A.8}$$

$$+\exp(\zeta_t)\Psi^{(0)}[\exp(\xi_t)+\exp(\zeta_t)] - \exp(\zeta_t)\ln[1+\exp(\epsilon_t)]$$

(2) For the NIG-DCS model, the conditional distribution of y_t is

$$y_t|(y_1,\ldots,y_{t-1}) \sim \text{NIG}[\mu_t, \exp(\lambda_t), \exp(\nu_t-\lambda_t), \exp(\nu_t-\lambda_t)\tanh(\eta_t)] \tag{A.9}$$

The conditional mean and the conditional volatility of y_t are

$$E(y_t|y_1,\ldots,y_{t-1};\Theta) = \mu_t + \frac{\exp(\lambda_t)\tanh(\eta_t)}{[1-\tanh^2(\eta_t)]^{1/2}} \tag{A.10}$$

$$SD(y_t|y_1,\ldots,y_{t-1};\Theta) = \left\{\frac{\exp(2\lambda_t-\nu_t)}{[1-\tanh^2(\eta_t)]^{3/2}}\right\}^{1/2} \tag{A.11}$$

respectively. The log of the conditional density of y_t is

$$\ln f(y_t|y_1,\ldots,y_{t-1};\Theta) = \nu_t - \lambda_t - \ln(\pi) + \exp(\nu_t)[1-\tanh^2(\eta_t)]^{1/2} \tag{A.12}$$

$$+\exp(\nu_t)\tanh(\eta_t)\epsilon_t+\ln K^{(1)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]-\frac{1}{2}\ln(1+\epsilon_t^2)$$

where $K^{(1)}(x)$ is the modified Bessel function of the second kind of order 1. The score functions with respect to μ_t, λ_t, ν_t and η_t are as follows. First, the score function with respect to μ_t is

$$\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \mu_t}=-\exp(\nu_t-\lambda_t)\tanh(\eta_t)+\frac{\epsilon_t}{\exp(\lambda_t)(1+\epsilon_t^2)} \tag{A.13}$$

$$+\frac{\exp(\nu_t-\lambda_t)\epsilon_t}{\sqrt{1+\epsilon_t^2}}\times\frac{K^{(0)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]+K^{(2)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]}{2K^{(1)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]}$$

where $K^{(0)}(x)$ and $K^{(2)}(x)$ are the modified Bessel functions of the second kind of orders 0 and 2, respectively. We define the scaled score function with respect to μ_t as

$$u_{\mu,t}=\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \mu_t}\times\exp(2\lambda_t)=\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \mu_t}\times k_t \tag{A.14}$$

Second, the score function with respect to λ_t is

$$u_{\lambda,t}=\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \lambda_t}=-1-\exp(\nu_t)\tanh(\eta_t)\epsilon_t+\frac{\epsilon_t^2}{1+\epsilon_t^2} \tag{A.15}$$

$$+\frac{\exp(\nu_t)\epsilon_t^2}{\sqrt{1+\epsilon_t^2}}\times\frac{K^{(0)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]+K^{(2)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]}{2K^{(1)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]}$$

Third, the score function with respect to ν_t is

$$u_{\nu,t}=\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \nu_t}=1+\exp(\nu_t)[1-\tanh^2(\eta_t)]^{1/2}+\exp(\nu_t)\tanh(\eta_t)\epsilon_t \tag{A.16}$$

$$-\exp(\nu_t)\sqrt{1+\epsilon_t^2}\times\frac{K^{(0)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]+K^{(2)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]}{2K^{(1)}\left[\exp(\nu_t)\sqrt{1+\epsilon_t^2}\right]}$$

Fourth, the score function with respect to η_t is

$$u_{\eta,t}=\frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \eta_t}=\exp(\nu_t)\mathrm{sech}^2(\eta_t)\epsilon_t-\exp(\nu_t)\tanh(\eta_t)\mathrm{sech}(\eta_t) \tag{A.17}$$

where $\mathrm{sech}(x)$ is the hyperbolic secant function.

(3) For the Skew-Gen-t-DCS model, the conditional distribution of y_t is

$$y_t|(y_1,\ldots,y_{t-1})\sim \text{Skew-Gen-}t[\mu_t,\exp(\lambda_t),\tanh(\tau_t),\exp(\nu_t)+4,\exp(\eta_t)] \tag{A.18}$$

The conditional mean of y_t is

$$E(y_t|y_1,\ldots,y_{t-1};\Theta)=\mu_t+2\exp(\lambda_t)\tanh(\tau_t)[\exp(\nu_t)+4]^{\exp(-\eta_t)}\times\frac{B\left\{\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right\}}{B\left\{\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right\}} \tag{A.19}$$

The conditional volatility of y_t is

$$SD(y_t|y_1,\ldots,y_{t-1};\Theta) = \exp(\lambda_t)[\exp(\nu_t)+4]^{\exp(-\eta_t)}\times \tag{A.20}$$

$$\times\left\{\frac{[3\tanh^2(\tau_t)+1]B\left[\frac{3}{\exp(\eta_t)},\frac{\exp(\nu_t)+2}{\exp(\eta_t)}\right]}{B\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]}-\frac{4\tanh^2(\tau_t)B^2\left[\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right]}{B^2\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]}\right\}^{1/2}$$

The log of the conditional density of y_t is

$$\ln f(y_t|y_1,\ldots,y_{t-1};\Theta) = \eta_t-\lambda_t-\ln(2)-\frac{\ln[\exp(\nu_t)+4]}{\exp(\eta_t)}-\ln\Gamma\left[\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right] \tag{A.21}$$

$$-\ln\Gamma[\exp(-\eta_t)]+\ln\Gamma\left[\frac{\exp(\nu_t)+5}{\exp(\eta_t)}\right]$$

$$-\frac{\exp(\nu_t)+5}{\exp(\eta_t)}\ln\left\{1+\frac{|\epsilon_t|^{\exp(\eta_t)}}{[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}\times[\exp(\nu_t)+4]}\right\}$$

First, the score function with respect to μ_t is

$$\frac{\partial\ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\mu_t}= \tag{A.22}$$

$$=\frac{[\exp(\nu_t)+4]\exp(\lambda_t)\epsilon_t|\epsilon_t|^{\exp(\eta_t)-2}}{|\epsilon_t|^{\exp(\eta_t)}+[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}[\exp(\nu_t)+4]}\times\frac{\exp(\nu_t)+5}{[\exp(\nu_t)+4]\exp(2\lambda_t)}=$$

$$=u_{\mu,t}\times\frac{\exp(\nu_t)+5}{[\exp(\nu_t)+4]\exp(2\lambda_t)}=\frac{u_{\mu,t}}{k_t}$$

where $u_{\mu,t}$ is the scaled score function. Second, the score function with respect to λ_t is

$$u_{\lambda,t}=\frac{\partial\ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\lambda_t}=\frac{|\epsilon_t|^{\exp(\eta_t)}[\exp(\nu_t)+5]}{|\epsilon_t|^{\exp(\eta_t)}+[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}[\exp(\nu_t)+4]}-1 \tag{A.23}$$

Third, the score function with respect to τ_t is

$$u_{\tau,t}=\frac{\partial\ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\tau_t}=\frac{[\exp(\nu_t)+5]|\epsilon_t|^{\exp(\eta_t)}\text{sgn}(\epsilon_t)\text{sech}(\tau_t)}{[\text{sgn}(\epsilon_t)\sinh(\tau_t)+\cosh(\tau_t)]}\times \tag{A.24}$$

$$\times\left\{|\epsilon_t|^{\exp(\eta_t)}+[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}[\exp(\nu_t)+4]\right\}^{-1}$$

Fourth, the score function with respect to ν_t is

$$u_{\nu,t}=\frac{\partial\ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial\nu_t}=-\frac{\exp(\nu_t-\eta_t)}{\exp(\nu_t)+4}-\exp(\nu_t-\eta_t)\Psi^{(0)}\left[\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right] \tag{A.25}$$

$$+\exp(\nu_t-\eta_t)\Psi^{(0)}\left[\frac{\exp(\nu_t)+5}{\exp(\eta_t)}\right]$$

$$+\frac{\exp(\nu_t-\eta_t)[\exp(\nu_t)+5]|\epsilon_t|^{\exp(\eta_t)}}{[\exp(\nu_t)+4]\left\{|\epsilon_t|^{\exp(\eta_t)}+[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}[\exp(\nu_t)+4]\right\}}$$

$$-\exp(\nu_t-\eta_t)\ln\left\{1+\frac{|\epsilon_t|^{\exp(\eta_t)}}{[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}[\exp(\nu_t)+4]}\right\}$$

Fifth, the score function with respect to η_t is

$$u_{\eta,t} = \frac{\partial \ln f(y_t|y_1,\ldots,y_{t-1};\Theta)}{\partial \eta_t} = 1 + \frac{\ln[\exp(\nu_t)+4]}{\exp(\eta_t)} + \frac{\exp(\nu_t)+4}{\exp(\eta_t)}\Psi^{(0)}\left[\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right] \tag{A.26}$$

$$+\frac{1}{\exp(\eta_t)}\Psi^{(0)}\left[\frac{1}{\exp(\eta_t)}\right] - \frac{\exp(\nu_t)+5}{\exp(\eta_t)}\Psi^{(0)}\left[\frac{\exp(\nu_t)+5}{\exp(\eta_t)}\right]$$

$$+\frac{\exp(\nu_t)+5}{\exp(\eta_t)}\ln\left\{1+\frac{|\epsilon_t|^{\exp(\eta_t)}[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{-\exp(\eta_t)}}{\exp(\nu_t)+4}\right\}$$

$$+\frac{[\exp(\nu_t)+5]|\epsilon_t|^{\exp(\eta_t)}\ln[1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]}{|\epsilon_t|^{\exp(\eta_t)}+[\exp(\nu_t)+4][1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}}$$

$$-\frac{[\exp(\nu_t)+5]|\epsilon_t|^{\exp(\eta_t)}\ln(|\epsilon_t|)}{|\epsilon_t|^{\exp(\eta_t)}+[\exp(\nu_t)+4][1+\tanh(\tau_t)\text{sgn}(\epsilon_t)]^{\exp(\eta_t)}}$$

Appendix B

$\tilde{C}$ is given by the outer product of $[(\partial\mu_t/\partial\theta),(\partial\mu_t/\partial\phi),(\partial\mu_t/\partial\tilde{c})]'$ and $[(\partial\lambda_t/\partial\alpha),(\partial\lambda_t/\partial\beta),(\partial\lambda_t/\partial\tilde{\omega})]'$ with itself. In this appendix we study the conditions under which the expected value of each of the nine elements of $\tilde{C}$ is finite. We study the dynamics for each element by using Equation (3.22), as follows:

With respect to $(\partial\mu_t/\partial\theta)\times(\partial\lambda_t/\partial\alpha)$, its expectation that is conditional on $(y_1,\ldots,y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\tfrac{\partial\mu_t}{\partial\theta}\tfrac{\partial\lambda_t}{\partial\alpha}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\tfrac{\partial\mu_{t-1}}{\partial\theta}\tfrac{\partial\lambda_{t-1}}{\partial\alpha} + E_{t-2}(X_{\mu,t-1}u_{\lambda,t-1})\tfrac{\partial\mu_{t-1}}{\partial\theta} + \\ & E_{t-2}(X_{\lambda,t-1}u_{\mu,t-1})\tfrac{\partial\lambda_{t-1}}{\partial\alpha} + E_{t-2}(u_{\mu,t-1}u_{\lambda,t-1}) \end{aligned} \tag{B.1}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.1), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the second and third terms on the right side of Equation (B.1), we use Condition 3 and Harvey (2013, p. 36, Lemma 6). According to Harvey (2013), $E(\partial\mu_t/\partial\theta) = E(\partial\lambda_t/\partial\alpha) = 0$, hence the second and third terms are zero. The fourth term on the right side of Equation (B.1) is constant under Condition 3. Thus, $E[(\partial\mu_t/\partial\theta)\times(\partial\lambda_t/\partial\alpha)]$ is finite.

With respect to $(\partial\mu_t/\partial\theta)\times(\partial\lambda_t/\partial\beta)$, its expectation that is conditional on $(y_1,\ldots,y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\tfrac{\partial\mu_t}{\partial\theta}\tfrac{\partial\lambda_t}{\partial\beta}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\tfrac{\partial\mu_{t-1}}{\partial\theta}\tfrac{\partial\lambda_{t-1}}{\partial\beta} + E_{t-2}(X_{\mu,t-1})(\lambda_{t-1}-\tilde{\omega})\tfrac{\partial\mu_{t-1}}{\partial\theta} + \\ & E_{t-2}(X_{\lambda,t-1}u_{\mu,t-1})\tfrac{\partial\lambda_{t-1}}{\partial\beta} + E_{t-2}(u_{\mu,t-1})(\lambda_{t-1}-\tilde{\omega}) \end{aligned} \tag{B.2}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.2), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the third term on the right side of Equation (B.2), we use Condition 3 and Harvey (2013, p. 36, Lemma 6). According to Harvey (2013), $E(\partial\lambda_t/\partial\beta) = 0$, hence the third term is zero. The fourth term on the right side of Equation (B.2) is zero, since $E(\lambda_t - \tilde{\omega}) = 0$. For the second term on the right side of Equation (B.2), we write the expectation:

$$\begin{aligned} E_{t-3}\left[(\lambda_{t-1}-\tilde{\omega})\tfrac{\partial\mu_{t-1}}{\partial\theta}\right] = \quad & E_{t-3}\left\{[\beta(\lambda_{t-2}-\tilde{\omega})+\alpha u_{\lambda,t-2}]\times\left[X_{\mu,t-2}\tfrac{\partial\mu_{t-2}}{\partial\theta}+u_{\mu,t-2}\right]\right\} = \\ & E_{t-3}(X_{\mu,t-2})\beta(\lambda_{t-2}-\tilde{\omega})\tfrac{\partial\mu_{t-2}}{\partial\theta} + E_{t-3}(u_{\mu,t-2})\beta(\lambda_{t-2}-\tilde{\omega}) + \\ & E_{t-3}(X_{\mu,t-2}u_{\lambda,t-2})\alpha\tfrac{\partial\mu_{t-2}}{\partial\theta} + E_{t-3}(u_{\mu,t-2}u_{\lambda,t-2})\alpha \end{aligned} \tag{B.3}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.3) is the first lag of the second term on the right side of Equation (B.2), multiplied by $|\beta| < 1$ (Condition 1). Under Condition 3, the expected value of the first term is finite. The second term on the right side is zero under Condition 3, and since $E(\lambda_t - \tilde{\omega}) = 0$. The third term on the right size is zero under Condition 3, and under $E(\partial\mu_t/\partial\theta) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right size is constant under Condition 3. Thus, $E[(\partial\mu_t/\partial\theta) \times (\partial\lambda_t/\partial\alpha)]$ is finite.

With respect to $(\partial\mu_t/\partial\theta) \times (\partial\lambda_t/\partial\tilde{\omega})$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\frac{\partial\mu_t}{\partial\theta}\frac{\partial\lambda_t}{\partial\tilde{\omega}}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\theta}\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}} + E_{t-2}(X_{\mu,t-1})\frac{\partial\mu_{t-1}}{\partial\theta}(1-\beta) + \\ & E_{t-2}(X_{\lambda,t-1}u_{\mu,t-1})\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}} + E_{t-2}(u_{\mu,t-1})(1-\beta) \end{aligned} \tag{B.4}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.4), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the second term on the right side of Equation (B.4), we use Condition 3 and Harvey (2013, p. 36, Lemma 6). According to Harvey (2013), $E(\partial\mu_t/\partial\theta) = 0$, hence the second term is zero. For the third term on the right side of Equation (B.4), we use Condition 3 and Harvey (2013, p. 36, Lemma 6). According to Harvey (2013), $E(\partial\lambda_t/\partial\tilde{\omega}) = (1-\beta)/[1-E(X_{\lambda,t})]$, hence the third term is constant. For the fourth term on the right side of Equation (B.4), the law of iterated expectations gives zero, because $E(u_{\mu,t}) = 0$. Thus, $E[(\partial\mu_t/\partial\theta) \times (\partial\lambda_t/\partial\tilde{\omega})]$ is finite.

With respect to $(\partial\mu_t/\partial\phi) \times (\partial\lambda_t/\partial\alpha)$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\frac{\partial\mu_t}{\partial\phi}\frac{\partial\lambda_t}{\partial\alpha}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\phi}\frac{\partial\lambda_{t-1}}{\partial\alpha} + E_{t-2}(X_{\mu,t-1}u_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\phi} + \\ & E_{t-2}(X_{\lambda,t-1})(\mu_{t-1}-\tilde{c})\frac{\partial\lambda_{t-1}}{\partial\alpha} + E_{t-2}(u_{\lambda,t-1})(\mu_{t-1}-\tilde{c}) \end{aligned} \tag{B.5}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.5), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the second term on the right side of Equation (B.5), we use Condition 3 and Harvey (2013, p. 36, Lemma 6). According to Harvey (2013), $E(\partial\mu_t/\partial\phi) = 0$, hence the second term is zero. For the fourth term on the right side of Equation (B.5), the law of iterated expectations gives zero, because $E(u_{\lambda,t}) = 0$. For the third term on the right side of Equation (B.5), we write the expectation:

$$\begin{aligned} E_{t-3}\left[(\mu_{t-1}-\tilde{c})\frac{\partial\lambda_{t-1}}{\partial\alpha}\right] = \quad & E_{t-3}\left\{[\phi(\mu_{t-2}-\tilde{c}) + \theta u_{\mu,t-2}] \times \left[X_{\lambda,t-2}\frac{\partial\lambda_{t-2}}{\partial\alpha} + u_{\lambda,t-2}\right]\right\} = \\ & E_{t-3}(X_{\lambda,t-2})\phi(\mu_{t-2}-\tilde{c})\frac{\partial\lambda_{t-2}}{\partial\alpha} + E_{t-3}(u_{\lambda,t-2})\phi(\mu_{t-2}-\tilde{c}) + \\ & E_{t-3}(X_{\lambda,t-2}u_{\mu,t-2})\theta\frac{\partial\lambda_{t-2}}{\partial\alpha} + E_{t-3}(u_{\mu,t-2}u_{\lambda,t-2})\theta \end{aligned} \tag{B.6}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.6) is the first lag of the third term on the right side of Equation (B.5), multiplied by $|\phi| < 1$ (Condition 1). Under Condition 3, the expected value of the first term is finite. The second term on the right side of Equation (B.6) is zero, since $E(\mu_t - \tilde{c}) = 0$. The third term on the right size of Equation (B.6) is zero under Condition 3, and under $E(\partial\lambda_t/\partial\alpha) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right size of Equation (B.6) is constant under Condition 3. Thus, $E[(\partial\mu_t/\partial\phi) \times (\partial\lambda_t/\partial\alpha)]$ is finite.

With respect to $(\partial\mu_t/\partial\phi) \times (\partial\lambda_t/\partial\beta)$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\frac{\partial\mu_t}{\partial\phi}\frac{\partial\lambda_t}{\partial\beta}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\phi}\frac{\partial\lambda_{t-1}}{\partial\beta} + E_{t-2}(X_{\mu,t-1})(\lambda_{t-1}-\tilde{\omega})\frac{\partial\mu_{t-1}}{\partial\phi} + \\ & E_{t-2}(X_{\lambda,t-1})(\mu_{t-1}-\tilde{c})\frac{\partial\lambda_{t-1}}{\partial\beta} + (\mu_{t-1}-\tilde{c})(\lambda_{t-1}-\tilde{\omega}) \end{aligned} \tag{B.7}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.7), the absolute value of the autoregressive parameter is < 1 under Condition 3. In the following we analyze covariance stationarity of the (i) second, (ii) third and (iii) fourth terms of Equation (B.7), respectively:

(i) For the second term on the right side of Equation (B.7), we write the expectation:

$$\begin{aligned} E_{t-3}\left[(\lambda_{t-1}-\tilde{\omega})\frac{\partial\mu_{t-1}}{\partial\phi}\right] = \quad & E_{t-3}\left\{[\beta(\lambda_{t-2}-\tilde{\omega}) + \alpha u_{\lambda,t-2}] \times \left[X_{\mu,t-2}\frac{\partial\mu_{t-2}}{\partial\phi} + \mu_{t-2}-\tilde{c}\right]\right\} = \\ & E_{t-3}(X_{\mu,t-2})\beta(\lambda_{t-2}-\tilde{\omega})\frac{\partial\mu_{t-2}}{\partial\phi} + \beta(\lambda_{t-2}-\tilde{\omega})(\mu_{t-2}-\tilde{c}) + \\ & E_{t-3}(X_{\mu,t-2}u_{\lambda,t-2})\alpha\frac{\partial\mu_{t-2}}{\partial\phi} + E_{t-3}(u_{\lambda,t-2})\alpha(\mu_{t-2}-\tilde{c}) \end{aligned} \tag{B.8}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.8) is the first lag of the second term on the right side of Equation (B.7), multiplied by $|\beta| < 1$ (Condition 1). Under Condition 3, the expected value of the first term is finite. The second term on the right side of Equation (B.8) is the first lag of the fourth term on the right side of Equation (B.7), multiplied by $|\beta| < 1$ (Condition 1). Thus, the expected value of the second term is finite. The third term on the right side of Equation (B.8) is zero under Condition 3, and under $E(\partial\mu_t/\partial\phi) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.8) is zero, because $E(\mu_t - \tilde{c}) = 0$.

(ii) For the third term on the right side of Equation Equation (B.7), we write the expectation:

$$\begin{aligned} E_{t-3}\left[(\mu_{t-1}-\tilde{c})\frac{\partial\lambda_{t-1}}{\partial\beta}\right] = \quad & E_{t-3}\left\{[\phi(\mu_{t-2}-\tilde{c}) + \theta u_{\mu,t-2}] \times \left[X_{\lambda,t-2}\frac{\partial\lambda_{t-2}}{\partial\beta} + \lambda_{t-2}-\tilde{\omega}\right]\right\} = \\ & E_{t-3}(X_{\lambda,t-2})\phi(\mu_{t-2}-\tilde{c})\frac{\partial\lambda_{t-2}}{\partial\beta} + \phi(\mu_{t-2}-\tilde{c})(\lambda_{t-2}-\tilde{\omega}) + \\ & E_{t-3}(X_{\lambda,t-2}u_{\mu,t-2})\theta\frac{\partial\lambda_{t-2}}{\partial\beta} + E_{t-3}(u_{\mu,t-2})(\lambda_{t-2}-\tilde{\omega}) \end{aligned} \tag{B.9}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.9) is the first lag of the third term on the right side of Equation (B.7), multiplied by $|\phi| < 1$ (Condition 1). Under Condition 3, the expected values of the first term is finite. The second term on the right side of Equation (B.9) is the first lag of the fourth term on the right side of Equation (B.7), multiplied by $|\phi| < 1$ (Condition 1). Thus, the expected values of the second term is finite. The third term on the right side of Equation (B.9) is zero under Condition 3, and under $E(\partial\lambda_t/\partial\beta) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.9) is zero, because $E(\lambda_t - \tilde{\omega}) = 0$.

(iii) For the fourth term on the right side of Equation (B.7), we write the expectation:

$$\begin{aligned} E_{t-3}\left[(\mu_{t-1}-\tilde{c})(\lambda_{t-1}-\tilde{\omega})\right] = \quad & E_{t-3}\left\{[\phi(\mu_{t-2}-\tilde{c}) + \theta u_{\mu,t-2}] \times [\beta(\lambda_{t-2}-\tilde{\omega}) + \alpha u_{\lambda,t-2}]\right\} = \\ & \phi\beta(\mu_{t-2}-\tilde{c})(\lambda_{t-2}-\tilde{\omega}) + E_{t-3}(u_{\lambda,t-2})\phi\alpha(\mu_{t-2}-\tilde{c}) + \\ & E_{t-3}(u_{\mu,t-2})\theta\beta(\lambda_{t-2}-\tilde{\omega}) + \theta\alpha E_{t-3}(u_{\mu,t-2}u_{\lambda,t-2}) \end{aligned} \tag{B.10}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.10) is the first lag of the fourth term on the right side of Equation (B.7), multiplied by $|\phi\beta| < 1$. Thus, the expected value of the first term is finite. The second and third terms on the right side of Equation (B.10) are zero, because $E(\mu_t - \tilde{c}) = E(\lambda_t - \tilde{\omega}) = 0$. The fourth term on the right side of Equation (B.10) is constant under Condition 3. Thus, $E[(\partial\mu_t/\partial\phi) \times (\partial\lambda_t/\partial\beta)]$ is finite.

With respect to $(\partial\mu_t/\partial\phi) \times (\partial\lambda_t/\partial\tilde{\omega})$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\frac{\partial\mu_t}{\partial\phi}\frac{\partial\lambda_t}{\partial\tilde{\omega}}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\phi}\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}} + E_{t-2}(X_{\mu,t-1})(1-\beta)\frac{\partial\mu_{t-1}}{\partial\phi} + \\ & E_{t-2}(X_{\lambda,t-1})(\mu_{t-1} - \tilde{c})\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}} + (\mu_{t-1} - \tilde{c})(1-\beta) \end{aligned} \tag{B.11}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.11), the absolute value of the autoregressive parameter is < 1 under Condition 3. The second term on the right side of Equation (B.11) is zero under Condition 3, and under $E(\partial\mu_t/\partial\phi) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.11) is zero, because $E(\mu_t - \tilde{c}) = 0$. For the third term on the right side of Equation (B.11), we write the expectation:

$$\begin{aligned} E_{t-3}\left[(\mu_{t-1} - \tilde{c})\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}}\right] = \quad & E_{t-3}\left\{[\phi(\mu_{t-2} - \tilde{c}) + \theta u_{\mu,t-2}] \times \left[X_{\lambda,t-2}\frac{\partial\lambda_{t-2}}{\partial\tilde{\omega}} + 1 - \beta\right]\right\} = \\ & E_{t-3}(X_{\lambda,t-2})\phi(\mu_{t-2} - \tilde{c})\frac{\partial\lambda_{t-2}}{\partial\tilde{\omega}} + \phi(\mu_{t-2} - \tilde{c})(1-\beta) + \\ & E_{t-3}(X_{\lambda,t-2}u_{\mu,t-2})\theta\frac{\partial\lambda_{t-2}}{\partial\tilde{\omega}} + E_{t-3}(u_{\mu,t-2})\theta(1-\beta) \end{aligned} \tag{B.12}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.12) is the first lag of the third term on the right side of Equation (B.11), multiplied by $|\phi| < 1$ (Condition 1). Under Condition 3, the expected value of the first term is finite. The second term on the right side of Equation (B.12) is zero, because $E(\mu_t - \tilde{c}) = 0$. The third term on the right side of Equation (B.12) is constant under Condition 3, and under $E(\partial\lambda_t/\partial\tilde{\omega}) = (1-\beta)/[1 - E(X_{\lambda,t})]$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.12) is zero, because $E(u_{\mu,t}) = 0$. Thus, $E[(\partial\mu_t/\partial\phi) \times (\partial\lambda_t/\partial\tilde{\omega})]$ is finite.

With respect to $(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\alpha)$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$\begin{aligned} E_{t-2}\left(\frac{\partial\mu_t}{\partial\tilde{c}}\frac{\partial\lambda_t}{\partial\alpha}\right) = \quad & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\tilde{c}}\frac{\partial\lambda_{t-1}}{\partial\alpha} + E_{t-2}(X_{\mu,t-1}u_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\tilde{c}} + \\ & E_{t-2}(X_{\lambda,t-1})(1-\phi)\frac{\partial\lambda_{t-1}}{\partial\alpha} + E_{t-2}(u_{\lambda,t-1})(1-\phi) \end{aligned} \tag{B.13}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.13), the absolute value of the autoregressive parameter is < 1 under Condition 3. The second term on the right side of Equation (B.13) is constant under Condition 3, and under $E(\partial\mu_t/\partial\tilde{c}) = (1-\phi)/[1 - E(X_{\mu,t})]$ in accordance with Harvey (2013, p. 36, Lemma 6). The third term on the right side of Equation (B.13) is zero under Condition 3, and under $E(\partial\lambda_t/\partial\alpha) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.13) is zero, because $E(u_{\lambda,t}) = 0$. Thus, $E[(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\alpha)]$ is finite.

With respect to $(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\beta)$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$E_{t-2}\left(\frac{\partial\mu_t}{\partial\tilde{c}}\frac{\partial\lambda_t}{\partial\beta}\right) = \begin{aligned} & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\tilde{c}}\frac{\partial\lambda_{t-1}}{\partial\beta} + E_{t-2}(X_{\mu,t-1})(\lambda_{t-1}-\tilde{\omega})\frac{\partial\mu_{t-1}}{\partial\tilde{c}} + \\ & E_{t-2}(X_{\lambda,t-1})(1-\phi)\frac{\partial\lambda_{t-1}}{\partial\beta} + (1-\phi)(\lambda_{t-1}-\tilde{\omega}) \end{aligned} \tag{B.14}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.14), the absolute value of the autoregressive parameter is < 1 under Condition 3. The third term on the right side of Equation (B.14) is zero under Condition 3, and under $E(\partial\lambda_t/\partial\beta) = 0$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.14) is zero, because $E(\lambda_t - \tilde{\omega}) = 0$. For the second term on the right side of Equation (B.14), we write the expectation:

$$E_{t-3}\left[(\lambda_{t-1}-\tilde{\omega})\frac{\partial\mu_{t-1}}{\partial\tilde{c}}\right] = \begin{aligned} & E_{t-3}\left\{[\beta(\lambda_{t-2}-\tilde{\omega}) + \alpha u_{\lambda,t-2}] \times \left[X_{\mu,t-2}\frac{\partial\mu_{t-2}}{\partial\tilde{c}} + 1 - \phi\right]\right\} = \\ & E_{t-3}(X_{\mu,t-2})\beta(\lambda_{t-2}-\tilde{\omega})\frac{\partial\mu_{t-2}}{\partial\tilde{c}} + \beta(\lambda_{t-2}-\tilde{\omega})(1-\phi) + \\ & E_{t-3}(X_{\mu,t-2}u_{\lambda,t-2})\alpha\frac{\partial\mu_{t-2}}{\partial\tilde{c}} + E_{t-3}(u_{\lambda,t-2})\alpha(1-\phi) \end{aligned} \tag{B.15}$$

We use the law of iterated expectations for the previous equation. The first term on the right side of Equation (B.15) is the first lag of the second term on the right side of Equation (B.14), multiplied by $|\beta| < 1$ (Condition 1). Under Condition 3, the expected value of the first term is finite. The second term on the right side of Equation (B.15) is zero, because $E(\lambda_{t-2} - \tilde{\omega}) = 0$. The third term on the right side of Equation (B.15) is constant under Condition 3, and under $E(\partial\mu_t/\partial\tilde{c}) = (1-\phi)/[1 - E(X_{\mu,t})]$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.15) is zero, because $E(u_{\lambda,t}) = 0$. Thus, $E[(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\beta)]$ is finite.

With respect to $(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\tilde{\omega})$, its expectation that is conditional on $(y_1, \ldots, y_{t-2})$ is:

$$E_{t-2}\left(\frac{\partial\mu_t}{\partial\tilde{c}}\frac{\partial\lambda_t}{\partial\tilde{\omega}}\right) = \begin{aligned} & E_{t-2}(X_{\mu,t-1}X_{\lambda,t-1})\frac{\partial\mu_{t-1}}{\partial\tilde{c}}\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}} + E_{t-2}(X_{\mu,t-1})(1-\beta)\frac{\partial\mu_{t-1}}{\partial\tilde{c}} + \\ & E_{t-2}(X_{\lambda,t-1})(1-\phi)\frac{\partial\lambda_{t-1}}{\partial\tilde{\omega}} + (1-\phi)(1-\beta) \end{aligned} \tag{B.16}$$

We use the law of iterated expectations for the previous equation. For the first term on the right side of Equation (B.16), the absolute value of the autoregressive parameter is < 1 under Condition 3. The second term on the right side of Equation (B.16) is constant under Condition 3, and under $E(\partial\mu_t/\partial\tilde{c}) = (1-\phi)/[1 - E(X_{\mu,t})]$ in accordance with Harvey (2013, p. 36, Lemma 6). The third term on the right side of Equation (B.16) is constant under Condition 3, and under $E(\partial\lambda_t/\partial\tilde{\omega}) = (1-\beta)/[1 - E(X_{\lambda,t})]$ in accordance with Harvey (2013, p. 36, Lemma 6). The fourth term on the right side of Equation (B.16) is constant. Thus, $E[(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\tilde{\omega})]$ is finite.

Appendix C

In this appendix, we describe the circumstances of the extreme events numbered in Figure 5.

(1) June 27-28, 1950. June 25, 1950: The Korean War began. North Korean (Democratic People's Republic of Korea) troops invaded South Korea (Republic of Korea) and proceeded toward Seoul. June 27, 1950: US President Harry Truman ordered US warships to assist South Korean forces.

(2) February 10, 1953. Egypt and West Germany (Federal Republic of Germany) broke their economic negotiations, due to the contacts established by Egypt with East Germany (German Democratic Republic).

(3) June 7, 1955. Prime Minister of India Jawaharlal Nehru visited the USSR.

(4) September 28-29, 1955. September 24, 1955: US President Dwight D. Eisenhower suffered a heart attack and was hospitalized for 6 weeks.

(5) August 11, 1959. August 9, 1959: The SM-65 Atlas, America's first intercontinental ballistic missile (ICBM) was declared to be operational after successful testing.

(6) April 18-19, 1961. April 17, 1961: The Bay of Pigs military invasion of Cuba undertaken by the Central Intelligence Agency (CIA) failed.

(7) May 29, 1962. On May 28, 1962, the stock exchanges of New York, London, Tokyo, Paris, Frankfurt and Zurich exhibited the largest one-day decline since the Great Depression.

(8) August 17, 1971. August 15, 1971: US President Richard M. Nixon announced the end of the international convertibility of the US dollar to gold.

(9) August 3, 1978. August 2, 1978: President Jimmy Carter declared an unprecedented state emergency and evacuation, immediately following the revelation that Niagara Falls, New York, neighborhood Love Canal was built on a toxic waste dump.

(10) September 5, 1979. The 1979 oil shock was related to events in the Middle East (the Iranian Revolution) and a strong global oil demand. The oil prices more than doubled between April 1979 and April 1980. This event influenced the increase of the inflation in the US to 9% by the end of 1979.

(11) August 18, 1982. Stock market crash of Kuwait's stock market named Souk Al-Manakh. Kuwait's financial sector was badly shaken by the crash, as was the entire economy. The S&P 500 declined 6% during the period of August 3-12, 1982. August 12, 1982: Mexico defaulted on its foreign debt.

(12) October 26, 1982. October 26, 1982: US budget deficit reached more than USD110 trillion for 1982.

(13) December 19, 1984. The Sino-British Joint Declaration, stating that China would resume the exercise of sovereignty over Hong Kong and the United Kingdom would restore Hong Kong to China with effect from July 1, 1997 was signed in Beijing, China by Deng Xiaoping and Margaret Thatcher.

(14) July 8, 1986. July 2, 1986: General strike against Pinochet regime in Chile. July 7, 1986: Supreme Court struck down Gramm-Rudman deficit-reduction law.

(15) September 12, 1986. September 11, 1986: Egyptian President Hosni Mubarak received Israeli Prime Minister Shimon Peres. September 11, 1986: US performed a nuclear test at Nevada Test Site. September 11, 1986: Dow Jones Industrial Average declined 86.61 points to 1,792.89.

(16) October 20, 1987. October 19 1987: Black Monday, stock markets around the world crashed.

(17) January 11, 1988. January 2, 1988: USSR began its program of economic restructuring (perestroika) with legislation initiated by Mikhail Gorbachev.

(18) April 15, 1988. April 3, 1988: USSR performed a nuclear test at Semipalitinsk Test Site. April 7, 1988: Russia announced that it would withdrew its troops from Afghanistan. April 7, 1988: US performed a nuclear test at Nevada Test Site. April 9, 1988: US imposed economic sanctions on Panama.

(19) May 15-16, 1989. May 10, 1989: General Manuel Noriega's Panama government nullified the country's elections, which the opposition had won by a 3-1 margin. May 11, 1989: US President George H. W. Bush ordered nearly 2,000 troops to Panama. May 13, 1989: Approximately 2,000 students began hunger strike in Tiananmen Square, China. May 14, 1989: Demonstration in Beijing's Tiananmen square.

(20) October 16-17, 1989. October 13, 1989: The S&P 500 index declined 6.1% as a result of the junk bond market collapse. On Friday 13 October 1989, there was a stock market mini-crash. The crash was caused by the breakdown of a USD6.75 billion leveraged buyout deal for UAL Corporation, the parent company of United Airlines. It triggered the collapse of the junk bond market.

(21) November 18-19, 1991. November 6, 1991: Russian President Boris Yeltsin outlawed the Communist Party. November 15, 1991: Dow Jones dropped 120.31 points (5th largest dive). November 15, 1991: The NASDAQ composite index declined 4.2%.

(22) February 17, 1993. February 5, 1993 Grenade exploded in Sarajevo, killing 63 and injuring 160.

(23) February 7, 1994. February 5, 1994: 68 killed and 200 wounded due to a mortar bomb in Sarajevo.

(24) May 19, 1995. May 1, 1995: Croatian forces launched Operation Flash during the Croatian War of Independence. May 2, 1995: Serbian missiles exploded in the heart of Zagreb, killing six. May 12, 1995: Dow Jones for 5th straight day of the week set a new record (4,430.59).

(25) March 11, 1996. March 7, 1996: The first democratically elected Palestinian parliament formed.

(26) July 8, 1996. July 7, 1996: Nelson Mandela stepped down as President of South Africa.

(27) October 28-29, 1997. October 20, 1997: The US accused Microsoft of violating a pact to stop Microsoft forcing makers of personal computers to include its Internet browser automatically. October 22, 1997: Compaq testified that Microsoft threatened to break Windows 95 agreement if they showcased a Netscape icon. October 27, 1997: Microsoft argued it should be "free from government interference". October 29, 1997: Iraq's Revolution Command Council announced that it would no longer allowed US citizens and US aircraft to serve with UN arms inspection teams.

(28) January 5, 2000. January 4, 2000: Alan Greenspan was nominated as US Federal Reserve Chairman for a fourth term.

(29) April 17, 2000. April 14, 2000: Metallica filed a lawsuit against P2P sharing phenomenon Napster. This law-suit eventually led the movement against file-sharing programs.

(30) February 28 and March 1, 2007. Stock prices in the US declined 3.5%, after a surprising 9% fall in the Shanghai market provoked worries worldwide about the global economy and the valuation of share prices. In the US, markets had already been shrinking due to concerns about deterioration in the mortgage market for people with poor credit, as well as worries about the economy. Alan Greenspan told a conference on February 26, 2007 that a recession in the US was likely.

(31) September 30, 2008. September 21, 2008: Goldman Sachs and Morgan Stanley, the two last remaining independent investment banks on Wall Street, became bank holding companies as a result of the subprime mortgage crisis. September 29, 2008: Dow Jones Industrial Average fell 777.68 points, its largest single-day point loss, following the bankruptcies of Lehman Brothers and Washington Mutual.

(32) February 23, 2011. February 11, 2011: Egyptian Revolution culminated in the resignation of Hosni Mubarak and the transfer of power to the Supreme Military Council after 18 days of protests (Arab Spring). February 14, 2011: The 2011 Bahraini uprising commenced. February 15, 2011: Libyan protests began opposing Colonel Muammar al-Gaddafi's rule.

(33) June 27-28, 2016. June 23, 2016: Brexit referendum: United Kingdom voted to leave the European Union (EU). June 24, 2016: British Prime Minister David Cameron resigned after the UK voted to leave the EU. June 26, 2016: City of Falluja freed from Islamic State (IS) control after a month-long campaign by Iraqi forces. June 28, 2016: Suicide bombings and gun attacks at Istanbul's Ataturk Airport.

(34) September 12, 2016. September 9, 2016: North Korea conducted its fifth nuclear test at the Punggye-ri Nuclear Test Site, at the time of its largest ever test in North Korea at 10 kilotons.

(35) May 18, 2017. May 9, 2017: US President Donald Trump dismissed FBI Director James Comey. May 9, 2017: Moon Jae-in was elected President of South Korea after a snap election to replace Park Geun-hye. May 15, 2017: UN Security Council condemned North Korea missile test.

(36) August 11, 2017. August 2, 2017: US President Donald Trump signed legislation imposing sanctions on Russia. August 5, 2017: UN Security Council voted to impose sanctions on North Korea for its continued missile program. August 9, 2017: North Korea said it planned to fire rockets on the US territory of Guam in the continuing escalation of tension between North Korea and the US.

References

Acharya, V., Engle, R., Richardson, M., 2012. Capital shortfall: a new approach to ranking and regulating systemic risks. American Economic Review 102 (3): 59–64. doi: 10.1257/aer.102.3.59.

Acharya, V.V., Pedersen, L.H., Philippon, T., Richardson, M., 2017. Measuring systemic risk. The Review of Financial Studies 30 (1): 2–47. doi: 10.1093/rfs/hhw088.

Ayala, A., Blazsek, S., Gonzalez, R.B., 2015. Default risk of sovereign debt in Central America. In: Finch, N. (Ed.), Emerging Markets and Sovereign Risk, Palgrave Macmillan UK, London, pp. 18–44.

Ayala, A., Blazsek, S., Cunado, J., Gil-Alana, L.A., 2016. Regime-switching purchasing power parity in Latin America: Monte Carlo unit root tests with dynamic conditional score. Applied Economics 48 (29): 2675–2696. doi: 10.1080/00036846.2015.1128076.

Ayala, A., Blazsek, S., 2018a. Score-driven copula models for portfolios of two risky assets. The European Journal of Finance 24 (18): 1861–1884. doi: 10.1080/1351847X.2018.1464488.

Ayala, A., Blazsek, S., 2018b. Equity market neutral hedge funds and the stock market: an application of score-driven copula models. Applied Economics 50 (37): 4005–4023. doi: 10.1080/00036846.2018.1440062.

Ayala, A., Blazsek, S., 2018c. Score-driven currency exchange rate seasonality as applied to the Guatemalan Quetzal/US Dollar. SERIEs-Journal of the Spanish Economic Association. doi: 10.1007/s13209-018-0186-0.

Ayala, A., Blazsek, S., 2019. Score-driven models of stochastic seasonality in location and scale: an application case study of the Indian rupee to USD exchange rate. Applied Economics (forthcoming).

Backus, D., Chernov, M., Martin, I., 2011. Disasters implied by equity index options. *Journal of Finance* 66 (6): 1969–2012. doi: 10.1111/j.1540-6261.2011.01697.x.

Bakshi, G., Kapadia, N., Madan D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. Review of Financial Studies 16 (1): 101–143. doi: 10.1093/rfs/16.1.0101.

Barndorff-Nielsen, O., Halgreen, C., 1977. Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. Probability Theory and Related Fields 38 (4): 309–311. doi: 10.1007/bf00533162.

Barndorff-Nielsen, O., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64 (2): 253–280. doi: 10.1111/1467-9868.00336.

Blasques, F., Koopman, S.J., Lucas, A., 2017. Maximum likelihood estimation of score-driven models. TI 2014-029/III, Tinbergen Institute Discussion Paper. https://papers.tinbergen.nl/14029.pdf.

Blasques, F., Gorgi, P., Koopman, S.J., Wintenberger, O., 2018. Feasible invertibility conditions and maximum likelihood estimation for observation-driven models. Electronic Journal of Statistics 12 (1): 1019–1052. doi: 10.1214/18-EJS1416.

Blazsek, S., Escribano, A., 2016. Score-driven dynamic patent count panel data models. Economics Letters 149: 116–119. doi: 10.1016/j.econlet.2016.10.026.

Blazsek, S., Mendoza, V., 2016. QARMA-Beta-t-EGARCH versus ARMA-GARCH: an application to S&P 500. Applied Economics 48 (12): 1119–1129. doi: 10.1080/00036846.2015.1093086.

Blazsek, S., Monteros, L.A., 2017a. Event-study analysis by using dynamic conditional score models. Applied Economics 49 (45): 4530–4541. doi: 10.1080/00036846.2017.1284996.

Blazsek, S., Monteros, L.A., 2017b. Dynamic conditional score models of degrees of freedom: filtering with score-driven heavy tails. Applied Economics 49 (53): 5426–5440. doi: 10.1080/00036846.2017.1307935.

Blazsek, S., Ho, H.-C., 2017. Markov regime-switching Beta-t-EGARCH. Applied Economics 49 (47): 47934805. doi: 10.1080/00036846.2017.1293794.

Blazsek, S., Ho, H.-C., Liu, S.-P., 2018. Score-driven Markov-switching EGARCH models: an application to systematic risk analysis. Applied Economics 50 (56): 6047–6060. doi: 10.1080/00036846.2018.1488073.

Blazsek, S., Villatoro, M., 2015. Is Beta-t-EGARCH(1,1) superior to GARCH(1,1)? Applied Economics 47 (17): 1764–1774. doi: 10.1080/00036846.2014.1000536.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3): 307–327. doi:10.1016/0304-4076(86)90063-1.

Bollerslev, T., 1987. A conditionally heteroscedastic time series model for speculative prices and rates of return. The Review of Economics and Statistics 69 (3): 542–547. doi: 10.2307/1925546.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Review of Financial Studies 22 (11): 4463–4492. doi: 10.1093/rfs/hhp008.

Bollerslev, T., Todorov, V., 2011. Estimation of jump tails. Econometrica 79 (6): 1727–1783. doi: 10.3982/ECTA9240.

Bollerslev, T., Todorov, V., 2014. Time-varying jump tails. Journal of Econometrics 183 (2): 168–180. doi: 10.1016/j.jeconom.2014.05.007.

Bollerslev, T., Todorov, V., Xu, L., 2015. Tail risk premia and return predictability. Journal of Financial Economics 118 (1): 113–134. doi: 10.1016/j.jfineco.2015.02.010.

Brandt, A. 1986., The stochastic equation $Y_{n+1} = A_nY_n + B_n$ with stationary coefficients. Advances in Applied Probability 18 (1): 211–220. doi: 10.2307/1427243.

Caivano, M., Harvey, A.C., 2014. Time-series models with an EGB2 conditional distribution. Journal of Time Series Analysis 35 (6): 558–571. doi: 10.1111/jtsa.12081.

Creal, D., Koopman, S.J., Lucas, A., 2011. A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. Journal of Business & Economic Statistics 29 (4): 552–563. doi: 10.1198/jbes.2011.10070.

Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. Journal of Applied Econometrics 28 (5): 777–795. doi: 10.1002/jae.1279.

Diaconis, P., Freedman, D., 1999. Iterated random functions. SIAM Review 41 (1): 45–76. doi: 10.1137/S0036144598338446.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74 (366): 427–431. doi: 10.2307/2286348.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4): 987–1008. doi: 10.2307/1912773.

Escanciano, J.C., Lobato, I.N., 2009. An automatic Portmanteau test for serial correlation. Journal of Econometrics 151 (2): 140–149. doi: 10.1016/j.jeconom.2009.03.001.

Galbraith, J.W., Zernov, S., 2004. Circuit breakers and the tail index of equity returns. Journal of Financial Econometrics 2 (1): 109–129. doi: 10.1093/jjfinec/nbh005.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton.

Harvey, A.C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.

Harvey, A.C., 2013. Dynamic Models for Volatility and Heavy Tails. Cambridge University Press, Cambridge.

Harvey, A.C., Chakravarty, T., 2008. Beta-t-(E)GARCH. Cambridge Working Papers in Economics 0840, Faculty of Economics, University of Cambridge, Cambridge. http://www.econ.cam.ac.uk/research/repec/cam/pdf/cwpe0840.pdf.

Harvey, A.C., Lange, R.J., 2017. Volatility modeling with a generalized t-distribution. Journal of Time Series Analysis 38 (2): 175–190. doi: 10.1111/jtsa.12224.

Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Review of Economic Studies 61 (2): 247–264. doi: 10.2307/2297980.

Harvey, A.C., Shephard, N., 1996. Estimation of an asymmetric stochastic volatility models for asset returns. Journal of Business & Economic Statistics 14 (4): 429–434. doi: 10.1080/07350015.1996.10524672.

Harvey, A.C., Sucarrat, G., 2014. EGARCH models with fat tails, skewness and leverage. Computational Statistics & Data Analysis 76: 320–338. doi: 10.1016/j.csda.2013.09.022.

Hull, J.C., 2018. Options, Futures, and Other Derivatives. Pearson, New York.

Jorion, P., 2006. Value at Risk: The New Benchmark for Managing Financial Risk. McGraw-Hill, New York.

Kelly, B., Jiang, H., 2014. Tail risk and asset prices. Review of Financial Studies 27 (10): 2841–2871. doi: 10.1093/rfs/hhu039.

Kib, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. The Review of Economic Studies 65 (3): 361–393. doi: 10.1111/1467-937X.00050.

Li, W.K., 2004. Diagnostic Checks in Time Series. Chapman & Hall/CRC, Boca Raton.

McDonald, J.B., Michelfelder, R.A., 2017. Partially adaptive and robust estimation of asset models: accommodating skewness and kurtosis in returns. Journal of Mathematical Finance 7: 219–237. doi: 10.4236/jmf.2017.71012.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59 (2): 347–370. doi: 10.2307/2938260.

Newey, K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3): 703–738. doi: 10.2307/1913610.

Quintos, C., Fan, Z., Phillips, P.C.B., 2001. Structural change tests in tail behavior and the Asian crisis. The Review of Economic Studies 68 (3): 633–663. doi: 10.1111/1467-937X.00184.

Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57 (2): 307–333. doi: 10.2307/1912557.

White, H., 1984. Asymptotic Theory for Econometricians. Academic Press, San Diego.

Table 1

Descriptive statistics for daily log-returns on the S&P 500 index, $y_t = \ln(p_t/p_{t-1})$.

Start date	4-Jan-1950
End date	30-Dec-2017
Sample size T	$17,109$
Minimum	-0.2290
Maximum	0.1096
Mean	0.0003
Standard deviation	0.0096
Skewness	-1.0162
Excess kurtosis	27.4010
$\mathrm{Corr}(y_t, y_{t-1})$	0.0269
$\mathrm{Corr}(y_t^2, y_{t-1})$	-0.0877

Source of data: Bloomberg

Table 2
Parameter estimates and model diagnostics, EGB2-DCS.

Constant ξ_t and ζ_t		Dynamic ξ_t and ζ_t		Dynamic ξ_t and constant ζ_t		Constant ξ_t and dynamic ζ_t	
c	0.0010***(0.0002)	c	0.0010***(0.0002)	c	0.0011***(0.0002)	c	0.0009***(0.0001)
ϕ	−0.1702**(0.0816)	ϕ	−0.1851**(0.0935)	ϕ	−0.2699***(0.1013)	ϕ	−0.0780(0.0779)
θ	0.0796***(0.0069)	θ	0.0764***(0.0089)	θ	0.0618***(0.0083)	θ	0.0950***(0.0087)
ω	−0.0633***(0.0076)	ω	−0.0453***(0.0066)	ω	−0.0521***(0.0070)	ω	−0.0606***(0.0074)
α	0.0381***(0.0019)	α	0.0362***(0.0018)	α	0.0363***(0.0019)	α	0.0382***(0.0019)
α^*	0.0242***(0.0014)	α^*	0.0191***(0.0015)	α^*	0.0225***(0.0015)	α^*	0.0225***(0.0014)
β	0.9890***(0.0014)	β	0.9921***(0.0012)	β	0.9909***(0.0013)	β	0.9894***(0.0013)
λ_0	−6.0829***(0.3351)	λ_0	−6.0100***(0.3111)	λ_0	−6.0369***(0.3220)	λ_0	−6.0705***(0.3326)
δ_1	−0.2270***(0.0600)	δ_1	−0.0584***(0.0225)	δ_1	−0.0966**(0.0378)	δ_1	−0.2156***(0.0596)
		γ_1	0.7138***(0.0770)	γ_1	0.5862***(0.1275)		
		κ_1	0.0418***(0.0076)	κ_1	0.0365***(0.0085)		
δ_2	−0.1278**(0.0641)	δ_2	−0.0236(0.0147)	δ_2	−0.1399**(0.0649)	δ_2	−0.0354(0.0236)
		γ_2	0.7992***(0.0583)			γ_2	0.6997***(0.1127)
		κ_2	−0.0337***(0.0071)			κ_2	−0.0283***(0.0089)
C_μ	0.0290	C_μ	0.0343	C_μ	0.0728	C_μ	0.0061
C_λ	0.8816	C_λ	0.8910	C_λ	0.8897	C_λ	0.8817
		C_ξ	0.4557	C_ξ	0.3054		
		C_ζ	0.6904			C_ζ	0.5273
$C_{\mu,\lambda}$	0.0255	$C_{\mu,\lambda}$	0.0305	$C_{\mu,\lambda}$	0.0648	$C_{\mu,\lambda}$	0.0054
		$C_{\mu,\xi}$	0.0156	$C_{\mu,\xi}$	0.0222		
		$C_{\mu,\zeta}$	0.0236			$C_{\mu,\zeta}$	0.0032
		$C_{\lambda,\xi}$	0.4080	$C_{\lambda,\xi}$	0.2731		
		$C_{\lambda,\zeta}$	0.6135			$C_{\lambda,\zeta}$	0.4636
		$C_{\xi,\zeta}$	0.3164				
MDS(mean)	0.7332	MDS(mean)	0.9522	MDS(mean)	0.0001	MDS(mean)	0.3188
MDS(variance)	0.1296	MDS(variance)	0.1353	MDS(variance)	0.1528	MDS(variance)	0.1194
MDS(skewness)	0.0000	MDS(skewness)	0.0000	MDS(skewness)	0.0000	MDS(skewness)	0.0000
MDS(kurtosis)	0.0000	MDS(kurtosis)	0.0000	MDS(kurtosis)	0.0000	MDS(kurtosis)	0.0000
LL	3.4531	LL	**3.4544**	LL	3.4537	LL	3.4534
AIC	−6.9050	AIC	**−6.9072**	AIC	−6.9060	AIC	−6.9055
BIC	−6.9004	BIC	**−6.9009**	BIC	−6.9005	BIC	−6.9000
HQC	−6.9035	HQC	**−6.9051**	HQC	−6.9042	HQC	−6.9037
		LR	0.0019	LR	0.0256	LR	0.1156

Notes: Standard errors are reported in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. p-values are reported for the MDS and LR tests. Bold numbers indicate superior statistical performance. Model specification: $y_t = \mu_t + \exp(\lambda_t)\epsilon_t$, $\epsilon_t \sim \text{EGB2}[0, 1, \exp(\xi_t), \exp(\zeta_t)]$, $\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1}$ and $\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} + \alpha^*\text{sgn}(-\epsilon_{t-1})(u_{\lambda,t-1}+1)$. For constant ξ_t or ζ_t: $\xi_t = \delta_1$ and $\zeta_t = \delta_2$. For dynamic ξ_t or ζ_t: $\xi_t = \delta_1 + \gamma_1\xi_{t-1} + \kappa_1 u_{\xi,t-1}$ and $\zeta_t = \delta_2 + \gamma_2\zeta_{t-1} + \kappa_2 u_{\zeta,t-1}$.

Table 3

Parameter estimates and model diagnostics, NIG-DCS.

Constant ν_t and η_t		Dynamic ν_t and η_t		Dynamic ν_t and constant η_t		Constant ν_t and dynamic η_t	
c	0.0010***(0.0001)	c	0.0010***(0.0002)	c	0.0010***(0.0001)	c	0.0010***(0.0002)
ϕ	−0.1687**(0.0812)	ϕ	−0.1948(0.1767)	ϕ	−0.1586**(0.0783)	ϕ	−0.1844*(0.0969)
θ	0.0400***(0.0039)	θ	0.0296***(0.0064)	θ	0.0412***(0.0039)	θ	0.0380***(0.0056)
ω	−0.0511***(0.0062)	ω	−0.0344***(0.0053)	ω	−0.0373***(0.0054)	ω	−0.0509***(0.0062)
α	0.0387***(0.0020)	α	0.0314***(0.0025)	α	0.0321***(0.0025)	α	0.0388***(0.0020)
α^*	0.0250***(0.0015)	α^*	0.0224***(0.0015)	α^*	0.0229***(0.0015)	α^*	0.0250***(0.0015)
β	0.9890***(0.0014)	β	0.9927***(0.0012)	β	0.9921***(0.0012)	β	0.9891***(0.0014)
λ_0	−5.0000***(0.3368)	λ_0	−4.9049***(0.3092)	λ_0	−4.9217***(0.3137)	λ_0	−5.0002***(0.3364)
δ_1	0.6897***(0.0534)	δ_1	0.1386***(0.0535)	δ_1	0.1528**(0.0661)	δ_1	0.6876***(0.0536)
		γ_1	0.8015***(0.0735)	γ_1	0.7844***(0.0903)		
		κ_1	0.0884***(0.0156)	κ_1	0.0780***(0.0157)		
δ_2	−0.0598***(0.0117)	δ_2	−0.0559**(0.0266)	δ_2	−0.0541***(0.0119)	δ_2	−0.0405(0.0825)
		γ_2	−0.0272(0.4062)			γ_2	0.3250(1.3515)
		κ_2	0.0164**(0.0069)			κ_2	0.0028(0.0051)
C_μ	0.0689	C_μ	0.0696	C_μ	0.0655	C_μ	0.0747
C_λ	0.9276	C_λ	0.9442	C_λ	0.9422	C_λ	0.9275
		C_ν	0.6033	C_ν	0.5811		
		C_η	0.0036			C_η	0.1020
$C_{\mu,\lambda}$	0.0650	$C_{\mu,\lambda}$	0.0663	$C_{\mu,\lambda}$	0.0626	$C_{\mu,\lambda}$	0.0703
		$C_{\mu,\nu}$	0.0430	$C_{\mu,\nu}$	0.0392		
		$C_{\mu,\eta}$	0.0002			$C_{\mu,\eta}$	0.0076
		$C_{\lambda,\nu}$	0.5720	$C_{\lambda,\nu}$	0.5497		
		$C_{\lambda,\eta}$	0.0034			$C_{\lambda,\eta}$	0.0946
		$C_{\nu,\eta}$	0.0022				
MDS(mean)	**0.7901**	MDS(mean)	**0.4975**	MDS(mean)	**0.7264**	MDS(mean)	**0.9504**
MDS(variance)	**0.1202**	MDS(variance)	**0.2563**	MDS(variance)	**0.2476**	MDS(variance)	**0.1171**
MDS(skewness)	**0.3265**	MDS(skewness)	**0.3181**	MDS(skewness)	**0.4159**	MDS(skewness)	**0.3143**
MDS(kurtosis)	**0.2947**	MDS(kurtosis)	**0.4125**	MDS(kurtosis)	**0.3529**	MDS(kurtosis)	**0.2939**
LL	3.4535	LL	**3.4542**	LL	3.4541	LL	3.4535
AIC	−6.9059	AIC	**−6.9068**	AIC	−6.9068	AIC	−6.9056
BIC	−6.9013	BIC	−6.9005	BIC	**−6.9013**	BIC	−6.9002
HQC	−6.9044	HQC	−6.9047	HQC	**−6.9050**	HQC	−6.9038
		LR	0.0884	LR	0.1201	LR	0.8418

Notes: Standard errors are reported in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. p-values are reported for the MDS and LR tests. For MDS, bold numbers indicate that the NIG distribution is supported. For the LL metrics, bold numbers indicate superior statistical performance. Model specification: $y_t = \mu_t + \exp(\lambda_t)\epsilon_t$, $\epsilon_t \sim \text{NIG}[0, 1, \exp(\nu_t), \exp(\nu_t)\tanh(\eta_t)]$, $\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1}$ and $\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} + \alpha^*\text{sgn}(-\epsilon_{t-1})(u_{\lambda,t-1} + 1)$. For constant ν_t or η_t: $\nu_t = \delta_1$ and $\eta_t = \delta_2$. For dynamic ν_t or η_t: $\nu_t = \delta_1 + \gamma_1\nu_{t-1} + \kappa_1 u_{\nu,t-1}$ and $\eta_t = \delta_2 + \gamma_2\eta_{t-1} + \kappa_2 u_{\eta,t-1}$.

Table 4(a)

Parameter estimates and model diagnostics, Skew-Gen-t-DCS.

Constant τ_t, ν_t and η_t		Dynamic τ_t, ν_t and η_t		Dynamic τ_t, ν_t and constant η_t		Dynamic τ_t, constant ν_t and dynamic η_t	
c	0.0009***(0.0001)	c	0.0009***(0.0001)	c	0.0009***(0.0001)	c	0.0009***(0.0001)
ϕ	−0.1683**(0.0819)	ϕ	−0.2197***(0.0355)	ϕ	−0.2233***(0.0793)	ϕ	−0.2111***(0.0734)
θ	0.1240***(0.0110)	θ	0.0951***(0.0114)	θ	0.0995***(0.0121)	θ	0.0925***(0.0084)
ω	−0.0574***(0.0068)	ω	−0.0455***(0.0006)	ω	−0.0486***(0.0063)	ω	−0.0466***(0.0068)
α	0.0388***(0.0020)	α	0.0356***(0.0014)	α	0.0367***(0.0021)	α	0.0360***(0.0021)
α^*	0.0263***(0.0015)	α^*	0.0246***(0.0010)	α^*	0.0251***(0.0015)	α^*	0.0246***(0.0015)
β	0.9888***(0.0014)	β	0.9912***(0.0001)	β	0.9905***(0.0012)	β	0.9910***(0.0013)
λ_0	−5.5171***(0.3280)	λ_0	−5.4697***(0.1601)	λ_0	−5.4821***(0.2809)	λ_0	−5.4846***(0.0093)
δ_1	−0.0419***(0.0090)	δ_1	−0.0367***(0.0079)	δ_1	−0.0345**(0.0164)	δ_1	−0.0385***(0.0065)
		γ_1	0.1106(0.1087)	γ_1	0.1656(0.3524)	γ_1	0.0739***(0.0202)
		κ_1	0.0133***(0.0047)	κ_1	0.0111***(0.0040)	κ_1	0.0136***(0.0038)
δ_2	1.5619***(0.1333)	δ_2	1.6708***(0.0423)	δ_2	0.9829*(0.5162)	δ_2	1.6022***(0.1216)
		γ_2	−0.0177(0.0219)	γ_2	0.4190(0.2851)		
		κ_2	3.3172***(0.0056)	κ_2	4.3525**(2.0503)		
δ_3	0.6072***(0.0274)	δ_3	0.1146***(0.0035)	δ_3	0.5871***(0.0249)	δ_3	0.2386***(0.0568)
		γ_3	0.8064***(0.0027)			γ_3	0.6010***(0.1030)
		κ_3	0.0492***(0.0100)			κ_3	0.0938***(0.0151)
C_μ	0.0638	C_μ	0.0832	C_μ	0.0842	C_μ	0.0914
C_λ	0.8949	C_λ	0.9068	C_λ	0.9038	C_λ	0.9049
		C_τ	0.0076	C_τ	0.0208	C_τ	0.0027
		C_ν	0.0427	C_ν	0.2110		
		C_η	0.6346			C_η	0.3402
$C_{\mu,\lambda}$	0.0588	$C_{\mu,\lambda}$	0.0770	$C_{\mu,\lambda}$	0.0776	$C_{\mu,\lambda}$	0.0853
		$C_{\mu,\tau}$	0.0007	$C_{\mu,\tau}$	0.0019	$C_{\mu,\tau}$	0.0003
		$C_{\mu,\nu}$	0.0020	$C_{\mu,\nu}$	0.0161		
		$C_{\mu,\eta}$	0.0536			$C_{\mu,\eta}$	0.0339
		$C_{\lambda,\tau}$	0.0071	$C_{\lambda,\tau}$	0.0192	$C_{\lambda,\tau}$	0.0026
		$C_{\lambda,\nu}$	0.0393	$C_{\lambda,\nu}$	0.1929		
		$C_{\lambda,\eta}$	0.5770			$C_{\lambda,\eta}$	0.3098
		$C_{\nu,\tau}$	0.0000	$C_{\nu,\tau}$	0.0039		
		$C_{\nu,\eta}$	0.0308				
		$C_{\eta,\tau}$	0.0049			$C_{\eta,\tau}$	0.0010
MDS(mean)	0.9175	MDS(mean)	0.5212	MDS(mean)	**0.4630**	MDS(mean)	0.5659
MDS(variance)	0.0728	MDS(variance)	0.3766	MDS(variance)	**0.1659**	MDS(variance)	0.2602
MDS(skewness)	0.2162	MDS(skewness)	0.0480	MDS(skewness)	**0.1316**	MDS(skewness)	0.0588
MDS(kurtosis)	0.2373	MDS(kurtosis)	0.4743	MDS(kurtosis)	**0.4796**	MDS(kurtosis)	0.1744
LL	3.4547	LL	**3.4560**	LL	3.4557	LL	3.4558
AIC	−6.9080	AIC	**−6.9099**	AIC	−6.9097	AIC	−6.9099
BIC	−6.9030	BIC	−6.9022	BIC	−6.9029	BIC	−6.9031
HQC	−6.9064	HQC	−6.9074	HQC	−6.9075	HQC	−6.9076
		LR	0.0079	LR	0.0024	LR	0.0150

Notes: Standard errors are reported in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. p-values are reported for the MDS and LR tests. For MDS, bold numbers indicate that the Skew-Gen-t distribution is supported. For the LL metrics, bold numbers indicate superior statistical performance. Model specification: $y_t = \mu_t + \exp(\lambda_t)\epsilon_t$, $\epsilon_t \sim$ Skew-Gen-$t[0, 1, \tanh(\tau_t), \exp(\nu_t) + 4, \exp(\eta_t)]$, $\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1}$ and $\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} + \alpha^*\text{sgn}(-\epsilon_{t-1})(u_{\lambda,t-1} + 1)$. For constant τ_t, ν_t or η_t: $\tau_t = \delta_1$, $\nu_t = \delta_2$ and $\eta_t = \delta_3$ For dynamic τ_t, ν_t or η_t: $\tau_t = \delta_1 + \gamma_1\tau_{t-1} + \kappa_1 u_{\tau,t-1}$, $\nu_t = \delta_2 + \gamma_2\nu_{t-1} + \kappa_2 u_{\nu,t-1}$ and $\eta_t = \delta_3 + \gamma_3\eta_{t-1} + \kappa_3 u_{\eta,t-1}$.

Table 4(b)

Parameter estimates and model diagnostics, Skew-Gen-t-DCS.

Dynamic τ_t and constant ν_t, η_t		Constant τ_t and dynamic ν_t, η_t		Constant τ_t, dynamic ν_t and constant η_t		Constant τ_t, ν_t and dynamic η_t	
c	0.0009***(0.0001)	c	0.0009***(0.0001)	c	0.0009***(0.0001)	c	0.0009***(0.0001)
ϕ	−0.2240***(0.0778)	ϕ	−0.1641***(0.0542)	ϕ	−0.1673***(0.0271)	ϕ	−0.1653***(0.0017)
θ	0.1055***(0.0136)	θ	0.1252***(0.0088)	θ	0.1248***(0.0092)	θ	0.1247***(0.0087)
ω	−0.0573***(0.0073)	ω	−0.0466***(0.0059)	ω	−0.0489***(0.0072)	ω	−0.0477***(0.0062)
α	0.0393***(0.0021)	α	0.0352***(0.0018)	α	0.0361***(0.0021)	α	0.0356***(0.0021)
α^*	0.0266***(0.0015)	α^*	0.0245***(0.0015)	α^*	0.0249***(0.0015)	α^*	0.0245***(0.0015)
β	0.9888***(0.0015)	β	0.9909***(0.0012)	β	0.9905***(0.0014)	β	0.9907***(0.0012)
λ_0	−5.5108***(0.0125)	λ_0	−5.4802***(0.2607)	λ_0	−5.4860***(0.2460)	λ_0	−5.4966***(0.0116)
δ_1	−0.0271(0.0180)	δ_1	−0.0411***(0.0064)	δ_1	−0.0418***(0.0069)	δ_1	−0.0415***(0.0079)
γ_1	0.3398(0.3911)						
κ_1	0.0083*(0.0044)						
δ_2	1.5603***(0.1766)	δ_2	1.5387***(0.3584)	δ_2	0.9429**(0.4619)	δ_2	1.5888***(0.1681)
		γ_2	0.0470(0.2068)	γ_2	0.4338*(0.2589)		
		κ_2	3.5991**(1.6798)	κ_2	4.2927**(1.9030)		
δ_3	0.6040***(0.0293)	δ_3	0.1060(0.0895)	δ_3	0.5955***(0.0293)	δ_3	0.2393***(0.0881)
		γ_3	0.8237***(0.1499)			γ_3	0.6054***(0.1422)
		κ_3	0.0410*(0.0241)			κ_3	0.0887***(0.0194)
C_μ	0.0867	C_μ	0.0633	C_μ	0.0639	C_μ	0.0629
C_λ	0.8941	C_λ	0.9067	C_λ	0.9043	C_λ	0.9050
C_τ	0.1044						
		C_ν	0.0491	C_ν	0.2229		
		C_η	0.6652			C_η	0.3463
$C_{\mu,\lambda}$	0.0792	$C_{\mu,\lambda}$	0.0590	$C_{\mu,\lambda}$	0.0594	$C_{\mu,\lambda}$	0.0584
$C_{\mu,\tau}$	0.0093						
		$C_{\mu,\nu}$	0.0013	$C_{\mu,\nu}$	0.0126		
		$C_{\mu,\eta}$	0.0428			$C_{\mu,\eta}$	0.0227
$C_{\lambda,\tau}$	0.0940						
		$C_{\lambda,\nu}$	0.0453	$C_{\lambda,\nu}$	0.2038		
		$C_{\lambda,\eta}$	0.6044			$C_{\lambda,\eta}$	0.3153
		$C_{\nu,\eta}$	0.0363				
MDS(mean)	0.4935	MDS(mean)	**0.5079**	MDS(mean)	**0.7284**	MDS(mean)	**0.4952**
MDS(variance)	0.0707	MDS(variance)	**0.3413**	MDS(variance)	**0.1757**	MDS(variance)	**0.2443**
MDS(skewness)	0.1908	MDS(skewness)	**0.1744**	MDS(skewness)	**0.2204**	MDS(skewness)	**0.2054**
MDS(kurtosis)	0.2320	MDS(kurtosis)	**0.4739**	MDS(kurtosis)	**0.4802**	MDS(kurtosis)	**0.1502**
LL	3.4548	LL	3.4558	LL	3.4556	LL	3.4556
AIC	−6.9080	AIC	−6.9097	AIC	−6.9096	AIC	−6.9097
BIC	−6.9021	BIC	−6.9029	BIC	−6.9037	BIC	**−6.9038**
HQC	−6.9061	HQC	−6.9075	HQC	−6.9077	HQC	**−6.9077**
LR	0.3323	LR	0.0169	LR	0.0064	LR	0.0341

Notes: Standard errors are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. p-values are reported for the MDS and LR tests. For MDS, bold numbers indicate that the Skew-Gen-t distribution is supported. For the LL metrics, bold numbers indicate superior statistical performance. Model specification: $y_t = \mu_t + \exp(\lambda_t)\epsilon_t$, $\epsilon_t \sim$ Skew-Gen-$t[0, 1, \tanh(\tau_t), \exp(\nu_t) + 4, \exp(\eta_t)]$, $\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1}$ and $\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} + \alpha^*\text{sgn}(-\epsilon_{t-1})(u_{\lambda,t-1} + 1)$. For constant τ_t, ν_t or η_t: $\tau_t = \delta_1$, $\nu_t = \delta_2$ and $\eta_t = \delta_3$ For dynamic τ_t, ν_t or η_t: $\tau_t = \delta_1 + \gamma_1\tau_{t-1} + \kappa_1 u_{\tau,t-1}$, $\nu_t = \delta_2 + \gamma_2\nu_{t-1} + \kappa_2 u_{\nu,t-1}$ and $\eta_t = \delta_3 + \gamma_3\eta_{t-1} + \kappa_3 u_{\eta,t-1}$.

Table 5
Likelihood-based model comparison.

EGB2-DCS	AIC	AIC rank	BIC	BIC rank	HQC	HQC rank
Constant ξ_t and ζ_t	−6.9050	16	−6.9004	14	−6.9004	16
Dynamic ξ_t and ζ_t	−6.9072	9	−6.9009	11	−6.9051	9
Dynamic ξ_t and constant ζ_t	−6.9060	12	−6.9005	12	−6.9042	13
Constant ξ_t and dynamic ζ_t	−6.9055	15	−6.9000	16	−6.9037	15
NIG-DCS	AIC	AIC rank	BIC	BIC rank	HQC	HQC rank
Constant ν_t and η_t	**−6.9059**	13	**−6.9013**	10	**−6.9044**	12
Dynamic ν_t and η_t	**−6.9068**	10	**−6.9005**	13	**−6.9047**	11
Dynamic ν_t and constant η_t	**−6.9068**	11	**−6.9013**	9	**−6.9050**	10
Constant ν_t and dynamic η_t	**−6.9056**	14	**−6.9002**	15	**−6.9038**	14
Skew-Gen-*t*-DCS	AIC	AIC rank	BIC	BIC rank	HQC	HQC rank
Constant τ_t, ν_t and η_t	−6.9080	8	−6.9030	4	−6.9064	7
Dynamic τ_t, ν_t and η_t	−6.9099	1	−6.9022	7	−6.9074	6
Dynamic τ_t, ν_t and constant η_t	**−6.9097**	4	**−6.9029**	6	**−6.9075**	5
Dynamic τ_t, constant ν_t and dynamic η_t	−6.9099	2	−6.9031	3	−6.9076	3
Dynamic τ_t and constant ν_t, η_t	−6.9080	7	−6.9021	8	−6.9061	8
Constant τ_t and dynamic ν_t, η_t	**−6.9097**	3	**−6.9029**	5	**−6.9075**	4
Constant τ_t, dynamic ν_t and constant η_t	**−6.9096**	6	**−6.9037**	2	**−6.9077**	2
Constant τ_t, ν_t and dynamic η_t	**−6.9097**	5	**−6.9038**	1	**−6.9077**	1

Notes: AIC, BIC and HQC with bold font indicate that all MDS specification tests support the model (see Tables 2 to 4).

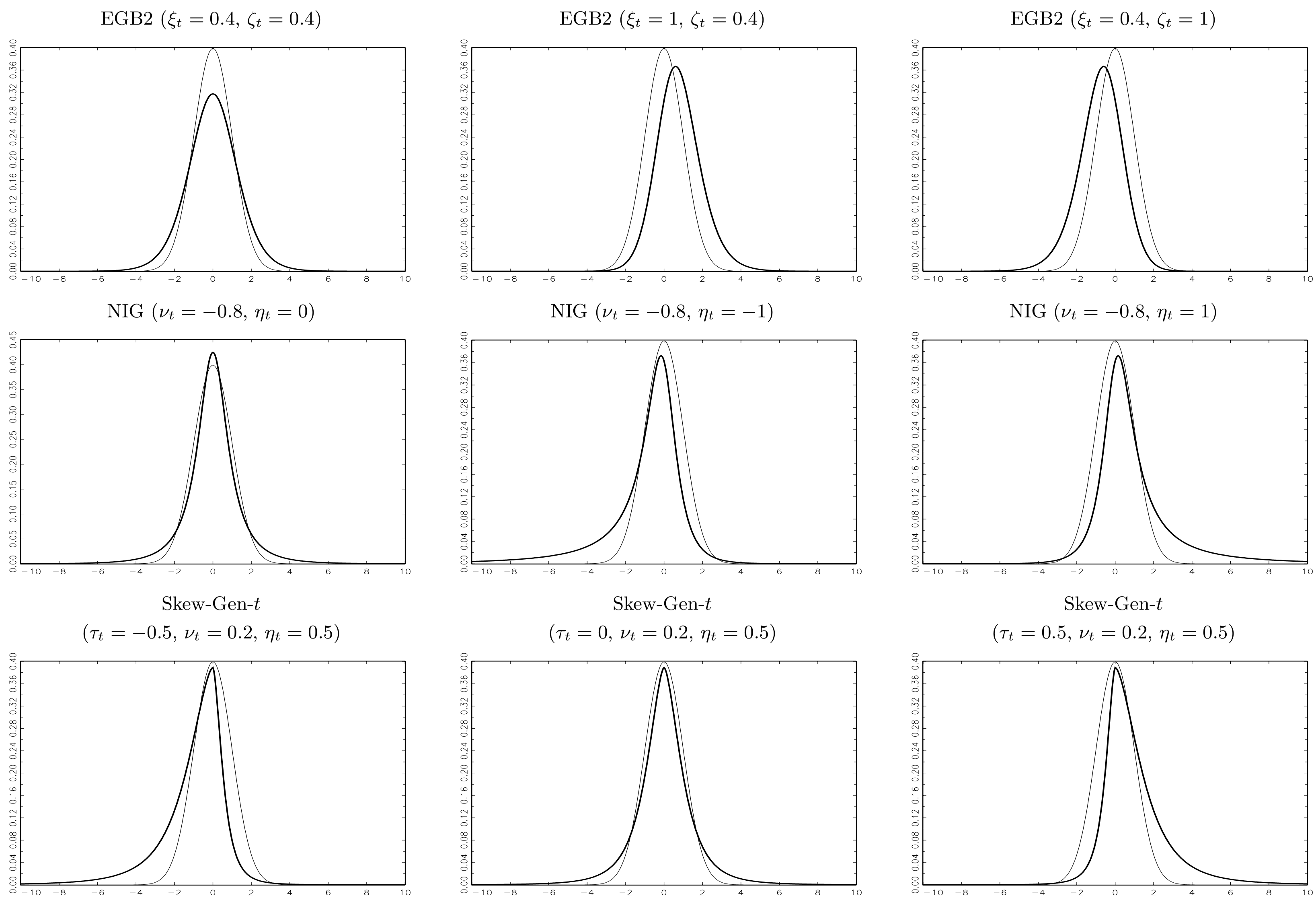

Fig. 1. Alternative density functions of ϵ_t (thick lines), compared to N(0,1) (thin lines).

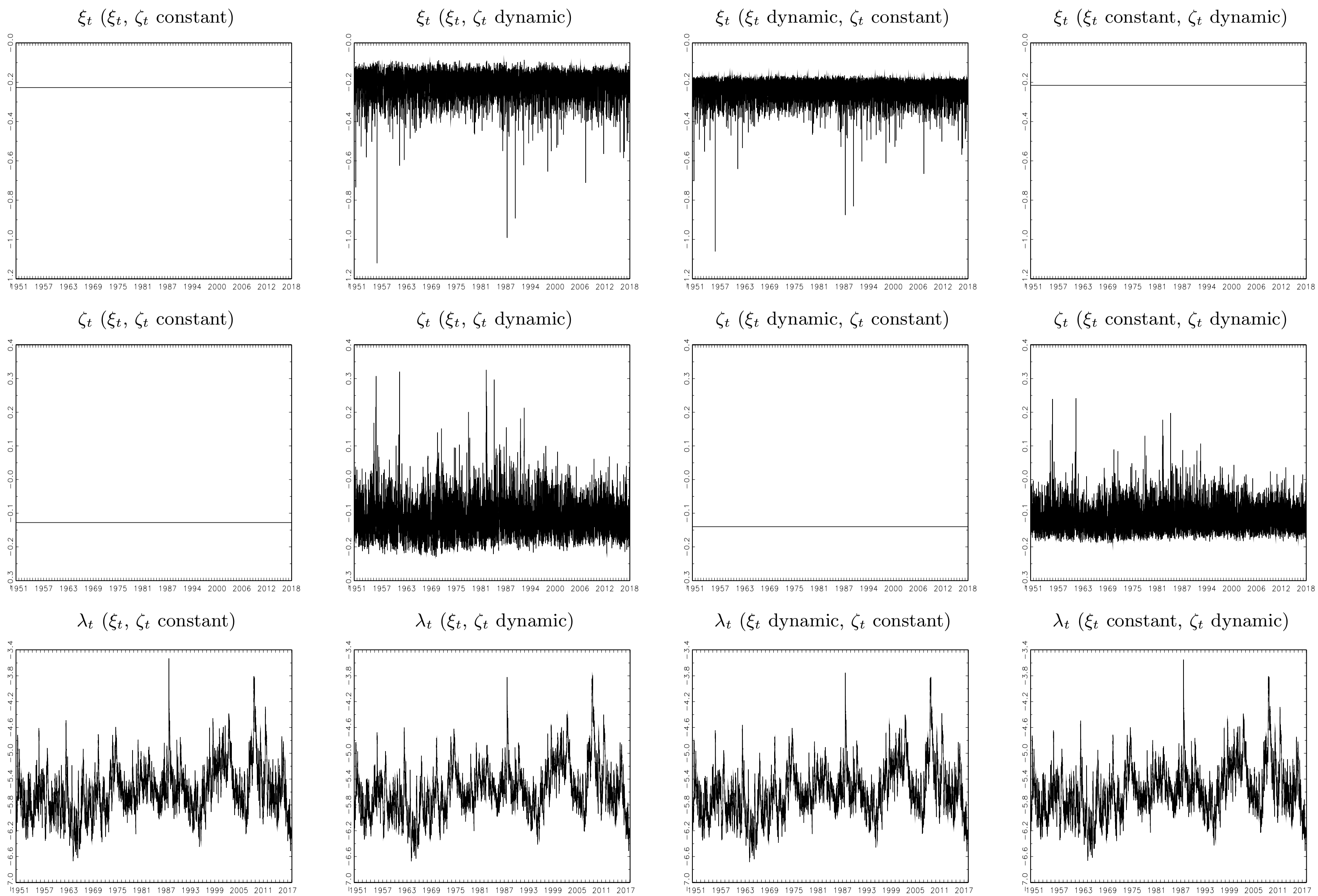

Fig. 2. DCS-EGB2 specifications. *Note*: Each column shows a different shape parameter specification.

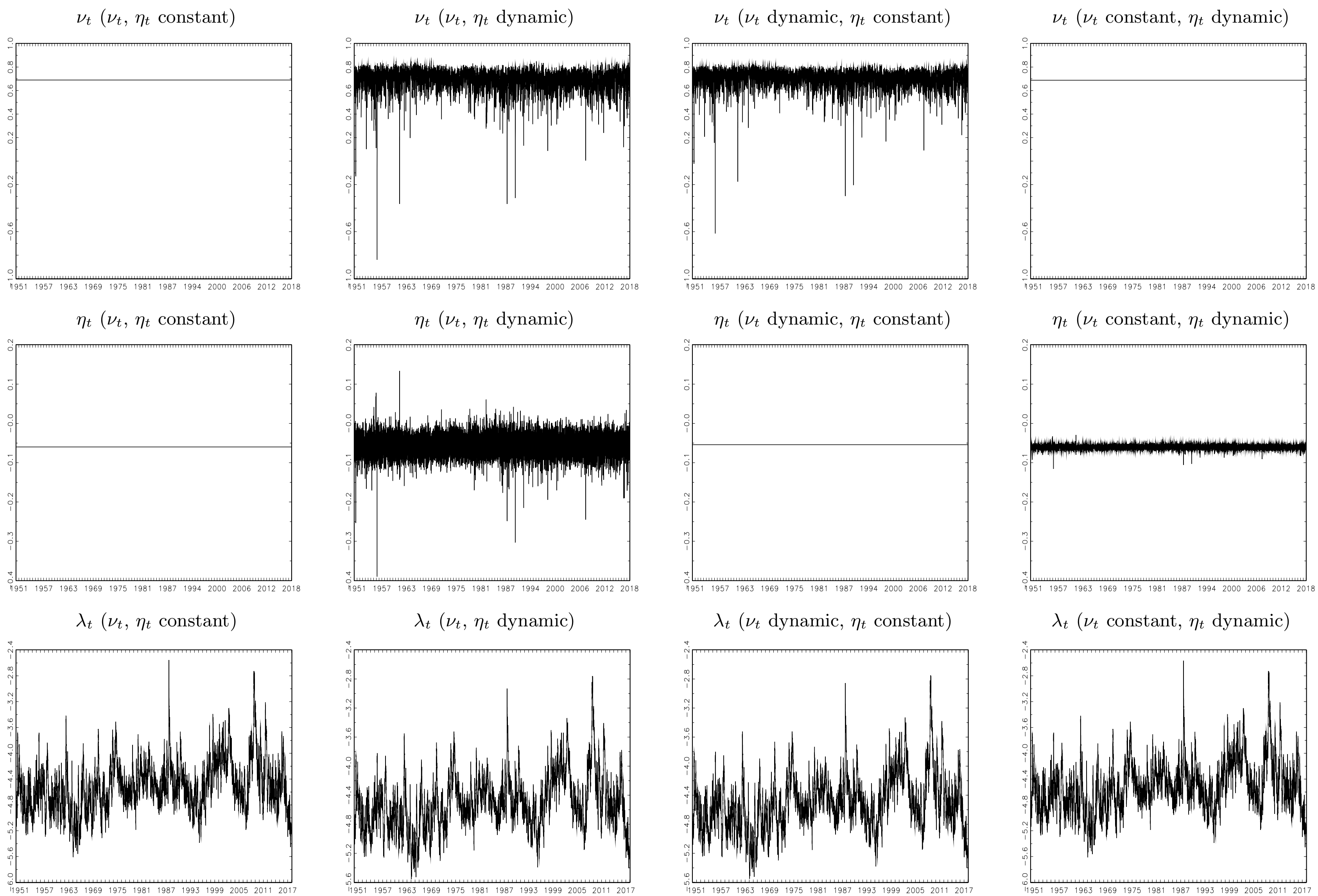

Fig. 3. DCS-NIG specifications. *Note*: Each column shows a different shape parameter specification.

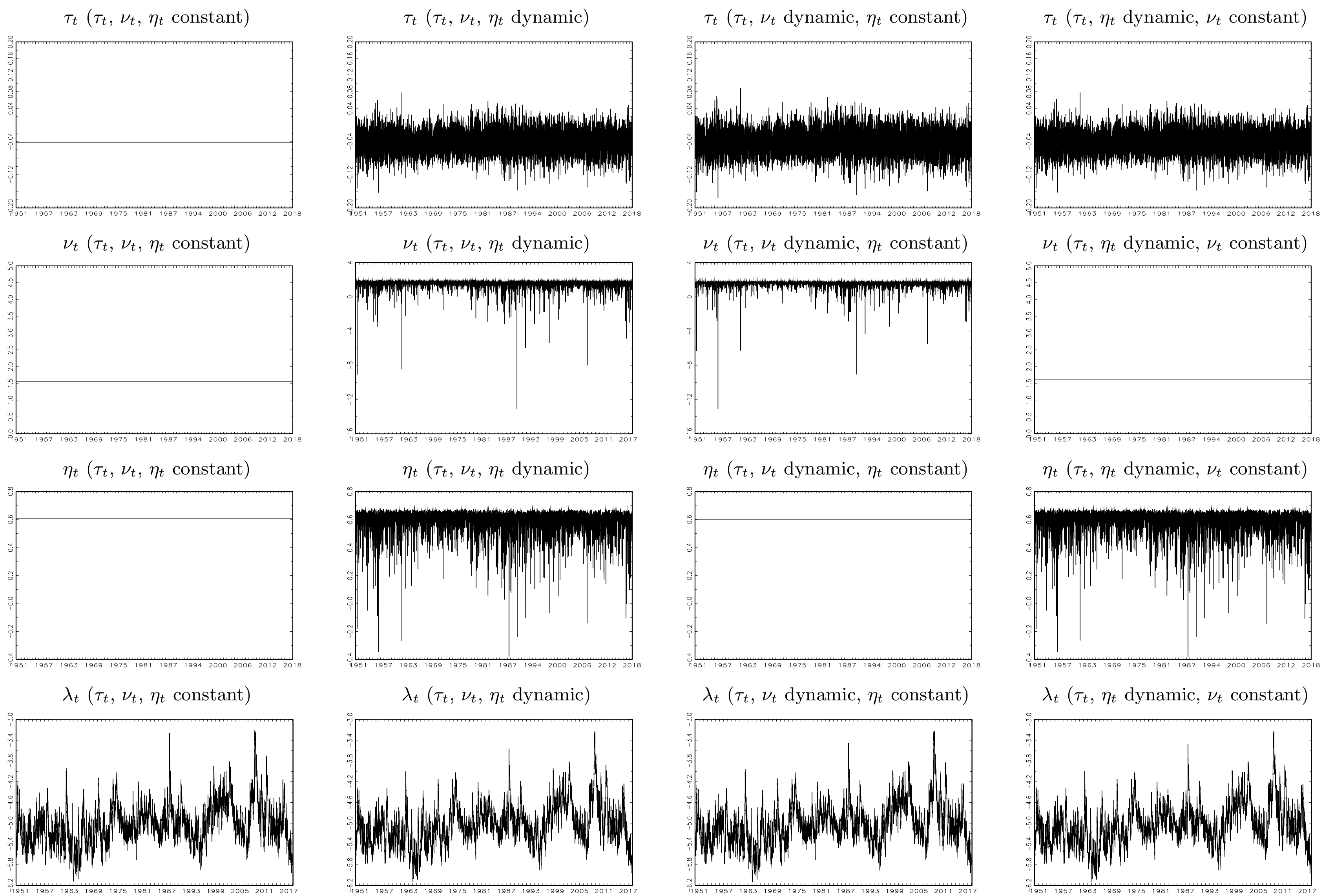

Fig. 4(a). DCS-Skew-Gen-t specifications. *Note*: Each column shows a different shape parameter specification.

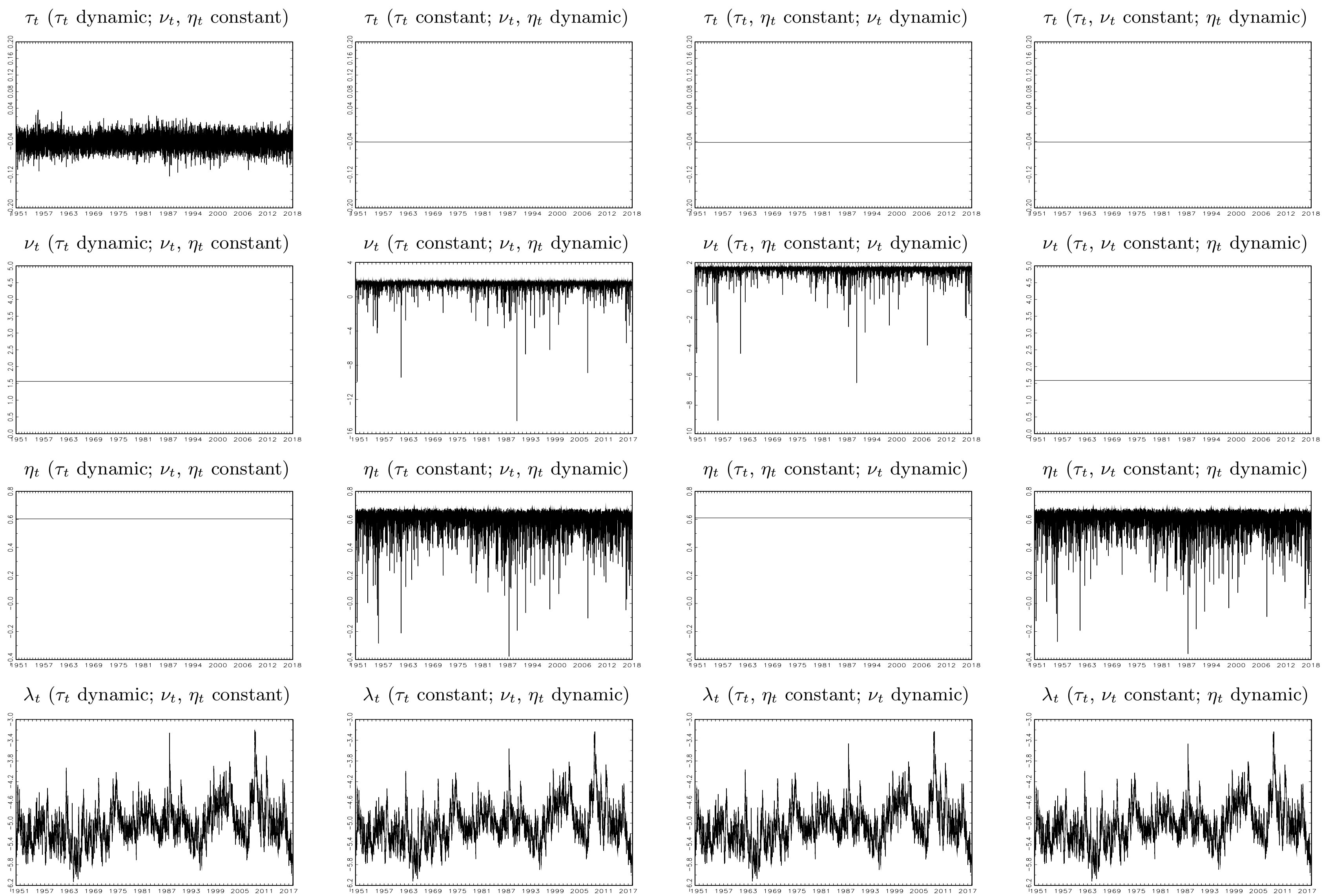

Fig. 4(b). DCS-Skew-Gen-t specifications. *Note*: Each column shows a different shape parameter specification.

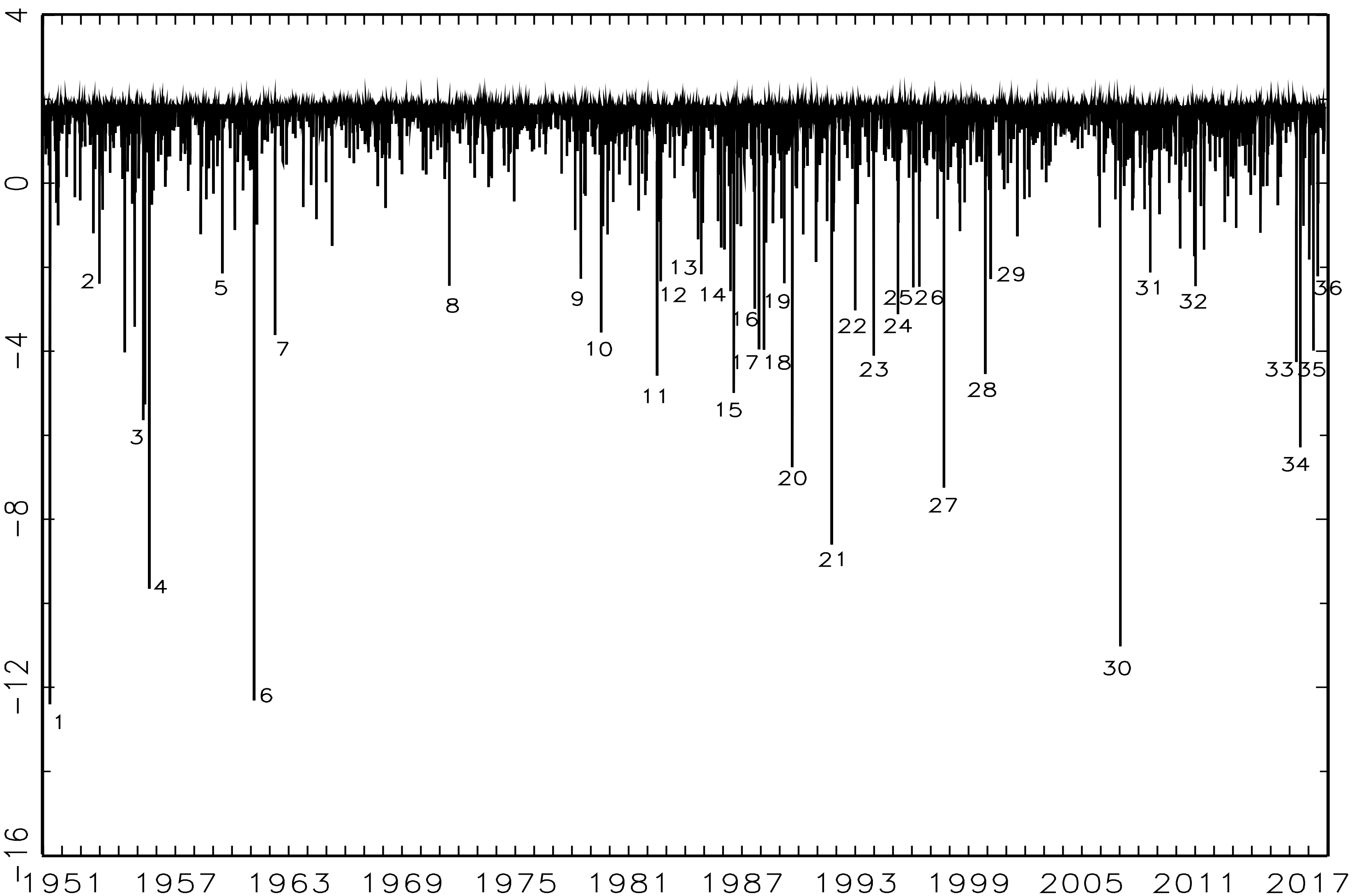

Fig. 5. Evolution of ν_t for the DCS-Skew-Gen-t model (τ_t, η_t constant; ν_t dynamic).

Note: We number the extreme events described in Appendix C.

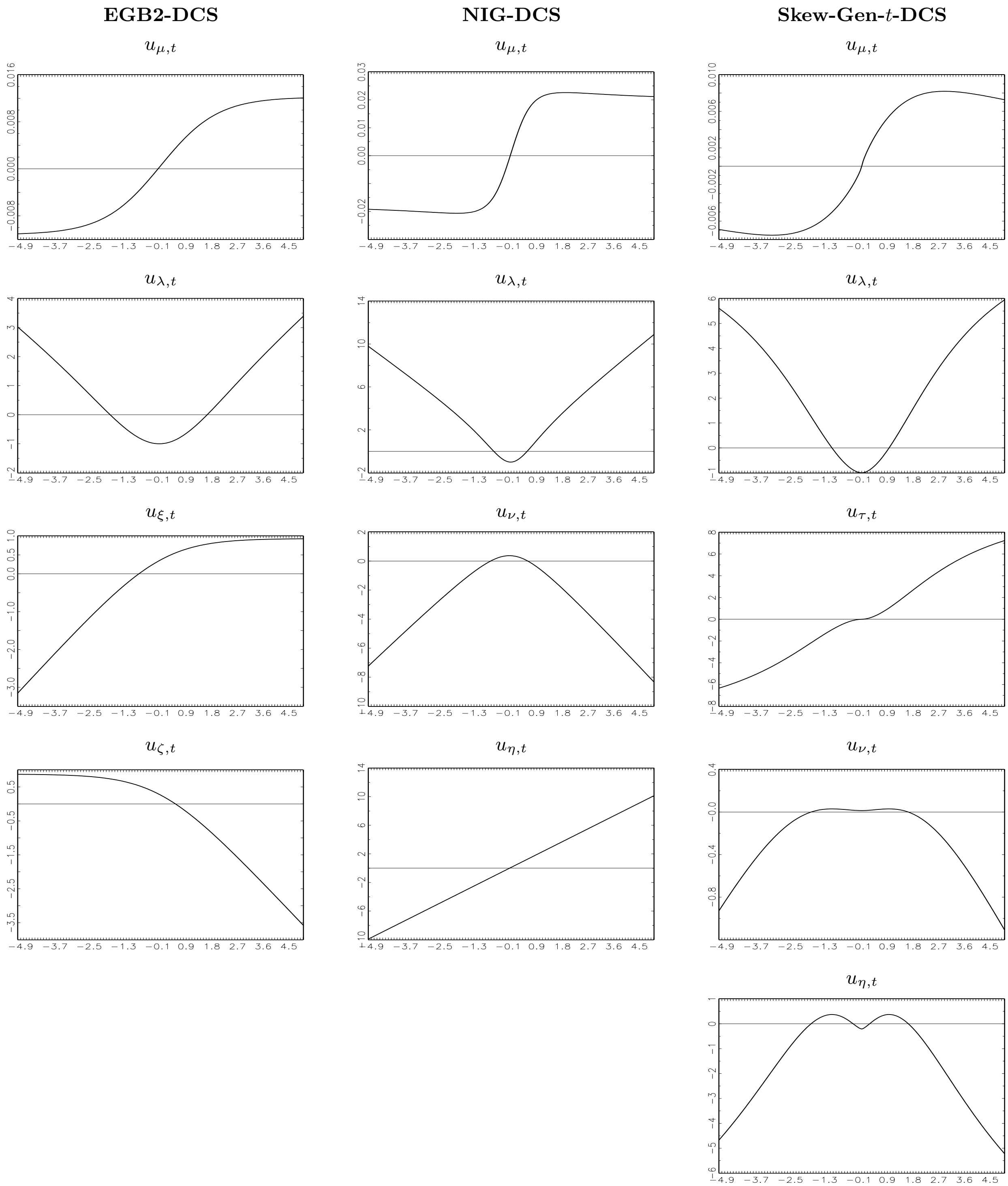

Fig. 6. Score functions, as a function of ϵ_t.

Notes: We present the estimates for those DCS specifications for which all shape parameters are time-varying. For all of the cases, we use the unconditional mean of the score-driven parameters in the computation.