2019-12
Working paper. Economics
ISSN 2340-5031

MAXIMUM LIKELIHOOD ESTIMATION OF SCORE-DRIVEN MODELS WITH DYNAMIC SHAPE PARAMETERS: AN APPLICATION TO MONTE CARLO VALUE-AT-RISK

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*

uc3m | Universidad Carlos III de Madrid

Maximum likelihood estimation of score-driven models with dynamic shape parameters: an application to Monte Carlo value-at-risk

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: Dynamic conditional score (DCS) models with time-varying shape parameters provide a flexible method for volatility measurement. The new models are estimated by using the maximum likelihood (ML) method, conditions of consistency and asymptotic normality of ML are presented, and Monte Carlo simulation experiments are used to study the precision of ML. Daily data from the Standard & Poor's 500 (S&P 500) for the period of 1950 to 2017 are used. The performances of DCS models with constant and dynamic shape parameters are compared. In-sample statistical performance metrics and out-of-sample value-at-risk backtesting support the use of DCS models with dynamic shape.

Keywords: Dynamic conditional score models; Score-driven shape parameters; Value-at-risk; Outliers.

JEL classification codes: C22; C52; C58

[*]Corresponding author. Postal address: School of Business, Universidad Francisco Marroquín, Calle Manuel F. Ayau, Zona 10, Guatemala City 01010, Guatemala. E-mail and telephone: sblazsek@ufm.edu, +502 2338-7783.

1. Introduction

When econometric models that include scale and shape parameters are applied to financial returns, then all those parameters appear in the volatility forecasting formulas. The scale parameter is dynamic in classical volatility models (e.g. Engle, 1982; Bollerslev, 1986, 1987; Nelson, 1991; Harvey et al., 1994; Harvey and Shephard, 1996; Kib et al., 1998; Barndorff-Nielsen and Shephard, 2002), but the shape of the probability distribution of returns is time-invariant in those models. In the present paper, new dynamic conditional score (DCS) models (Creal et al., 2011, 2013; Harvey, 2013) are introduced, within which several dynamic shape parameters are used in the measurement of volatility. In the new DCS models, news on asset value updates volatility, not only through scale but also through shape. The dynamics of all scale and shape parameters are estimated in one step. The new DCS models of this paper use 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) probability distributions with dynamic shape parameters. These EGB2-DCS, NIG-DCS and Skew-Gen-t-DCS models, respectively, are extensions of the DCS models with constant shape parameters from the body of literature (e.g. Harvey, 2013; Harvey and Sucarrat, 2014; Harvey and Lange, 2017). The results of this paper suggest that the DCS models with dynamic shape improve the performance of the DCS models with constant shape, because: (i) they have superior in-sample statistical performances, and (ii) they provide more accurate out-of-sample value-at-risk (VaR) measurements.

In this paper, the dynamic tail shape, skewness and peakedness of financial returns are estimated. For the dynamic tail shape of returns, different models appear in the literature. Quintos et al. (2001) construct tests of tail shape constancy that allow for an unknown breakpoint, and they apply those tests to stock price data for a period that covers the Asian financial crisis of 1997-1998. Galbraith and Zernov (2004) apply the same tests to the Dow Jones Industrial Average (DJIA) and Standard & Poor's 500 (S&P 500) indices. Bollerslev and Todorov (2011) suggest a nonparametric method of dynamic tail shape, which, in their study, is applied to high-frequency data from the S&P 500. Several works use options data for the estimation of dynamic tail shapes of the underlying assets (e.g. Bakshi et al., 2003; Bollerslev et al., 2009; Backus et al., 2011; Bollerslev and Todorov, 2014; Bollerslev et al., 2015). Kelly and Jiang (2014) identify a common variation in the tail shape of United States (US) stock returns, by using panel data models.

The DCS models of the present paper are estimated by using the maximum likelihood (ML) method, and the conditions of consistency and asymptotic normality of the ML estimates are presented. For each DCS model, the correct specification of the probability distributions of financial returns is verified up to the fourth moment. The precision of the ML estimator is also studied, by performing several Monte Carlo (MC) simulation experiments for known data generating processes. Daily log-return time series data are used from the S&P 500 index for the period of January 4, 1950 to December 30, 2017. The application of these data is relevant for practitioners for the effective estimation and prediction of volatility, VaR and expected shortfall on: (i) well-diversified US equity portfolios; (ii) S&P 500 futures and options contracts; (iii) exchange traded funds (ETFs) related to the S&P 500 index. According to the estimation results, the likelihood-based performance of the Skew-Gen-t-DCS model is superior to the likelihood-based performances of the EGB2-DCS and NIG-DCS models. The score-driven dynamics of the shape of financial returns are significant for all models. The ML estimation results show that the in-sample statistical performances of DCS models with dynamic shape parameters are superior to the in-sample statistical performances of DCS models with constant shape parameters.

In order to motivate the practical use of the new DCS models, out-of-sample VaR backtesting is performed for the S&P 500. Data for the period of September 2, 2008 to March 31, 2009 are used from the 2008 US financial crisis. The results for the S&P 500 show that predicted potential extreme losses for the trading day after each outlier are higher for the DCS models with dynamic shape parameters than for the DCS models with constant shape parameters. These suggest that DCS models with dynamic shape may be effective for the prediction of consecutive additive outliers. Motivated by these findings, a modified dataset is used for the S&P 500, in which additive outliers are duplicated. For the modified sample, the DCS models with dynamic shape parameters predict extreme losses for consecutive outliers, while the DCS models with constant shape parameters fail to predict extreme losses for the second outlier. These results may motivate the application of the new DCS specifications in VaR measurements of financial risk managers during periods of high market volatility.

The remainder of this paper is organized as follows: Section 2 introduces the DCS models with dynamic shape parameters. Section 3 presents the statistical inferences of those models, the conditions of the asymptotic properties of the ML estimates, and the MC experiments that study the precision of the ML estimator. Section 4 describes the data. Section 5 presents the empirical results. Section 6 presents the VaR backtesting results. Section 7 concludes.

2. Econometric methods

2.1. DCS models with dynamic shape parameters

The DCS models of the daily log-return on the S&P 500 index y_t are formulated as:

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

where μ_t and $\exp(\lambda_t)$ are the location and scale parameters, respectively. For ϵ_t, the EGB2, NIG and Skew-Gen-t distributions are used (Appendix A). For $\epsilon_t \sim \text{EGB2}[0, 1, \exp(\xi_t), \exp(\zeta_t)]$, both shape parameters are positive. For $\epsilon_t \sim \text{NIG}[0, 1, \exp(\nu_t), \exp(\nu_t)\tanh(\eta_t)]$, $\tanh(x)$ is the hyperbolic tangent function, and the absolute value of parameter $\exp(\nu_t)\tanh(\eta_t)$ is less than parameter $\exp(\nu_t)$. For $\epsilon_t \sim \text{Skew-Gen-}t[0, 1, \tanh(\tau_t), \exp(\nu_t) + 4, \exp(\eta_t)]$, shape parameter $\tanh(\tau_t)$ is in the interval $(-1, 1)$, degrees of freedom parameter $\exp(\nu_t) + 4$ is higher than four, and shape parameter $\exp(\eta_t)$ is positive. For these distributions, $E(y_t|y_1, \ldots, y_{t-1}) \equiv E(y_t|\mathcal{F}_{t-1}) \neq \mu_t$, since $E(\epsilon_t|\mathcal{F}_{t-1}) \neq 0$ (Appendix A). For ease of notation, $\rho_{k,t}$ is used as a general notation for ξ_t, ζ_t, ν_t, η_t and τ_t. The index k in $\rho_{k,t}$ refers to the k-th shape parameter of the distribution, which is determined by a transformation of $\rho_{k,t}$. For example, $\text{EGB2}[0, 1, \exp(\xi_t), \exp(\zeta_t)] = \text{EGB2}[0, 1, \exp(\rho_{1,t}), \exp(\rho_{2,t})]$. The density functions of the EGB2, NIG and Skew-Gen-t distributions are presented in Fig. 1, in which different values are used for the shape parameters and the density function of the standard normal distribution is also shown.

In the following, the score-driven filters for μ_t, λ_t and $\rho_{k,t}$ are presented. Firstly, μ_t is specified as:

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

where $|\phi| < 1$ and $u_{\mu,t}$ is the scaled score function of the log-likelihood (LL) with respect to μ_t (Appendix A). This location model can be related to the unobserved components models (UCMs) (Harvey, 1989), because a UCM is obtained by replacing $\theta u_{\mu,t-1}$ by a contemporaneous Gaussian i.i.d. error term. The location equation is jointly estimated with scale and shape, since in that way the model controls for possible dynamics in the expected return and the measurement of volatility dynamics is improved. The joint estimation is required for this model, due to the score-driven updating mechanism that is based on the conditional density of y_t (Appendix A). Secondly, λ_t is specified as:

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

where $|\beta| < 1$, $u_{\lambda,t}$ is the score function of the LL with respect to λ_t (Appendix A), and $\text{sgn}(x)$ is the signum function. 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. Thirdly, $\rho_{k,t}$ is specified as:

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

where $|\gamma_k| < 1$, and $u_{\rho,k,t}$ is the score function of the LL with respect to $\rho_{k,t}$ (Appendix A). For the EGB2 distribution, the two dynamic parameters that influence shape are denoted as $\rho_{1,t} = \xi_t$ and $\rho_{2,t} = \zeta_t$. For the NIG distribution, the two dynamic parameters that influence shape are denoted as $\rho_{1,t} = \nu_t$ and $\rho_{2,t} = \eta_t$. 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 $\rho_{3,t} = \eta_t$. For each distribution the constant shape parameter model is used as the benchmark, for which $\rho_{k,t} = \delta_k$.

Due to the score-driven updating, the information gain in the filters is optimal according to the Kullback–Leibler divergence measure (Blasques et al., 2015). For the updating of μ_t a scaled score function is used, for which the dynamic scaling parameters are defined by using the Fisher information matrix (Harvey, 2013). For λ_t, the score function is scaled by using a constant parameter (Harvey, 2013). Similarly, the score function is scaled by using a constant parameter for $\rho_{k,t}$. Several works suggest the use of the inverse of the Fisher information matrix (e.g. Creal et al., 2013; Harvey, 2013), which implies dynamic scaling for all score functions. The DCS models of the present paper may be improved in future works, by using more effective scaling mechanisms.

As an extension of this model, contemporaneous values and lags of exogenous explanatory variables may also be included on the right sides of Eqs. (2.2) to (2.4). Different ways of initialization are considered for each dynamic equation. For the results reported in this paper, μ_t is initialized 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, the results of this paper are robust to other ways of initialization. For example, the results for the case when parameters μ_0 and $\rho_{k,0}$ are used for initialization are similar to the results reported in this paper.

Several specifications in DCS models are robust to extreme observations, because the score functions in those models reduce the effects of outliers (e.g. Harvey, 2013). This property is studied here for the DCS models with dynamic shape parameters. By using the S&P 500 dataset, the outlier transformation of the score functions is presented in Fig. 2, which shows that extreme observations are never accentuated by the score functions of the new DCS models. Therefore, outliers appear within the unexpected return v_t rather than within the score functions that update the dynamic equations.

2.2. Model specification test

For the EGB2 and NIG distributions, the first four moments exist. For the Skew-Gen-t distribution, the degrees of freedom parameter specification ensures that the first four moments exist (i.e. the degrees of freedom parameter is > 4). The first four conditional moments of ϵ_t for the EGB2, NIG and Skew-Gen-t distributions are reported in Appendix A. Define the auxiliary error term as:

$$\epsilon_t^* = \frac{\epsilon_t - E(\epsilon_t|\mathcal{F}_{t-1};\Theta)}{\text{Var}^{1/2}(\epsilon_t|\mathcal{F}_{t-1};\Theta)} = \frac{\epsilon_t - E(\epsilon_t|\epsilon_1,\ldots,\epsilon_{t-1};\Theta)}{\text{Var}^{1/2}(\epsilon_t|\epsilon_1,\ldots,\epsilon_{t-1};\Theta)} \tag{2.5}$$

This transformation reduces the importance of those outliers that appear within ϵ_t. The robustness of model specification tests is increased when residuals are standardized according to Eq. (2.5) (see Li, 2004, Chapter 4). The model specification test of the present paper uses the following properties:

$$E(\epsilon_t^*|\mathcal{F}_{t-1};\Theta) = E(\epsilon_t^*|\epsilon_1^*,\ldots,\epsilon_{t-1}^*;\Theta) = 0 \tag{2.6}$$

$$E[(\epsilon_t^*)^2 - 1|\mathcal{F}_{t-1};\Theta] = E[(\epsilon_t^*)^2 - 1|\epsilon_1^*,\ldots,\epsilon_{t-1}^*;\Theta] = 0 \tag{2.7}$$

$$E[(\epsilon_t^*)^3 - \text{Skew}(\epsilon_t|\mathcal{F}_{t-1})|\mathcal{F}_{t-1};\Theta] = \tag{2.8}$$

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

$$E[(\epsilon_t^*)^4 - \text{Kurt}(\epsilon_t|\mathcal{F}_{t-1})|\mathcal{F}_{t-1};\Theta] = \tag{2.9}$$

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

Within the expectations of Eqs. (2.6) to (2.9), the variables are martingale difference sequences (MDSs). The MDS test with optimal lag-order selection (Escanciano and Lobato, 2009) is applied in the present paper, to verify the correct specification for each probability distribution.

3. Statistical inference

All models of this paper are estimated by using the ML method. Blasques et al. (2017, 2018) present the conditions for the asymptotic properties of ML for DCS models with one score-driven parameter. In the present paper, the same conditions are presented for DCS models with several score-driven parameters. The ML estimator is

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

where $\Theta = (\Theta_1, \ldots, \Theta_K)'$ is the vector of parameters and $\ln f(y_t|\mathcal{F}_{t-1}; \Theta)$ is presented in Appendix A. The following assumptions are used: (A1) $f(y_t|\mathcal{F}_{t-1}; \Theta) = p_0(y_t|\mathcal{F}_{t-1}; \Theta_0)$ for some Θ from the parameter set $\tilde{\Theta}$, where p_0 is the true conditional density and Θ_0 denotes the true values of the parameters. (A2) $\int_{\mathbb{R}} f(y_t|\mathcal{F}_{t-1}; \Theta)dy_t = 1$ for all y_t and Θ. (A3) $\tilde{\Theta} \in \mathbb{R}^K$ is compact. (A4) $\hat{\Theta}_{\text{ML}}$ is a unique solution to the problem of Eq. (3.1). (A5) $\text{LL}(\cdot; \Theta)$ is a Borel measurable function on $\mathbb{R}^T$. (A6) For each $(y_1, \ldots, y_T) \in \mathbb{R}^T$, $\text{LL}(y_1, \ldots, y_T; \cdot)$ is a continuous function on $\tilde{\Theta}$. (A7) $|\text{LL}(y_1, \ldots, y_T; \Theta)| < b(y_1, \ldots, y_T)$ for all Θ and $E[b(y_1, \ldots, y_T)] < \infty$. Under (A1) to (A7), the ML estimator is consistent: $\hat{\Theta}_{\text{ML}} \to_p \Theta_0$ as $T \to \infty$.

The following results use some additional assumptions: (A8) Θ_0 is an interior point within $\tilde{\Theta} \in \mathbb{R}^K$. (A9) $\text{LL}(y_1, \ldots, y_T; \Theta)$ is twice continuously differentiable on all of the interior points of $\tilde{\Theta}$. (A10) $\partial[\int_{\mathbb{R}} f(y_t|\mathcal{F}_{t-1}; \Theta)dy_t]/\partial\Theta = \int_{\mathbb{R}}[\partial f(y_t|\mathcal{F}_{t-1}; \Theta)/\partial\Theta]dy_t$.

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

$$G_{ti}(\Theta) = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1}; \Theta)}{\partial \Theta_i} \tag{3.2}$$

for period $t = 1, \ldots, T$ and parameter $i = 1, \ldots, K$. 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 (A10), the ML estimator of Eq. (3.1) is equivalent to the following representation:

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

According to the mean-value expansion about Θ_0:

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

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

$$H_t(\Theta) = \frac{\partial^2 \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial\Theta\Theta'} \tag{3.5}$$

which is evaluated at K different mean values, indicated by $\bar{\Theta}$. Each $\bar{\Theta}$ is located between Θ_0 and $\hat{\Theta}_{\text{ML}}$ that is formally expressed as: $||\bar{\Theta} - \Theta_0|| \leq ||\hat{\Theta}_{\text{ML}} - \Theta_0||$, where $||\cdot||$ is the Euclidean norm.

The following results use some additional assumptions: (A11) $\partial[\int_{\mathbb{R}} G_t(\Theta)' f(y_t|\mathcal{F}_{t-1};\Theta)dy_t]/\partial\Theta = \int_{\mathbb{R}}[\partial G_t(\Theta)' f(y_t|\mathcal{F}_{t-1};\Theta)/\partial\Theta]dy_t$. (A12) The information matrix $\mathcal{I}(\Theta_0) \equiv -E[H_t(\Theta_0)]$ is positive definite. (A13) The elements of $\mathcal{I}(\Theta_0)$ are bounded in absolute value by function $b(y_1,\ldots,y_T)$ for all Θ and $E[b(y_1,\ldots,y_T)] < \infty$. The conditions of (A13) are studied in Section 3.1. Under (A1) to (A13), the $K \times K$ contribution to the information matrix for period t is given by:

$$I_t(\Theta_0) = -E[H_t(\Theta_0)|\mathcal{F}_{t-1}] = E[G_t(\Theta_0)' G_t(\Theta_0)|\mathcal{F}_{t-1}] \tag{3.6}$$

which is evaluated at the true values of parameters. From Eqs. (3.3) and (3.4):

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

$$\sqrt{T}(\hat{\Theta}_{\text{ML}} - \Theta_0) = \mathcal{I}^{-1}(\Theta_0)\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right] + o_p(1) \tag{3.8}$$

The following result uses the assumptions: (A14) A central limit theorem is satisfied for Eq. (3.8); (A14) is studied in Section 3.2. (A15) The DCS model is invertible (Blasques et al., 2017, 2018). Under (A1) to (A15), the asymptotic distribution of the ML estimates is given by:

$$\sqrt{T}(\hat{\Theta}_{\text{ML}} - \Theta_0) \to_d N_K\left[0_{K\times 1}, \mathcal{I}^{-1}(\Theta_0)\right] \quad \text{as} \quad T \to \infty \tag{3.9}$$

The asymptotic covariance matrix of $\hat{\Theta}_{\text{ML}}$ is estimated by using $[\sum_{t=1}^{T} G_t(\hat{\Theta}_{\text{ML}})' G_t(\hat{\Theta}_{\text{ML}})]^{-1}$.

3.1. Conditions of (A13)

For ease of notation, a DCS model with known constant shape parameters is considered first:

$$y_t = \mu_t + \exp(\lambda_t)\epsilon_t \tag{3.10}$$

$$\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1} \tag{3.11}$$

$$\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} \tag{3.12}$$

Variables μ_t and λ_t are re-parameterized, by using the unconditional means $E(\mu_t) = \tilde{c} = c/(1-\phi)$ and $E(\lambda_t) = \tilde{\omega} = \omega/(1-\beta)$, as follows:

$$\mu_t = \tilde{c}(1-\phi) + \phi\mu_{t-1} + \theta u_{\mu,t-1} \tag{3.13}$$

$$\lambda_t = \tilde{\omega}(1-\beta) + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} \tag{3.14}$$

for which $\Theta = (\tilde{c}, \phi, \theta, \tilde{\omega}, \beta, \alpha)$ and $K = 6$. This alternative form of the model is used, since the information matrix is simpler under this representation (Harvey, 2013, p. 34). The conditions for the covariance stationarity of y_t are $|\phi| < 1$ and $|\beta| < 1$. These conditions are named as Condition 1.

To study the finiteness of the elements of $\mathcal{I}(\Theta_0)$, matrix $\mathcal{I}(\Theta_0)$ is expressed as:

$$\mathcal{I}(\Theta_0) = E[I_t(\Theta_0)] = E\left\{E[G_t(\Theta_0)'G_t(\Theta_0)|\mathcal{F}_{t-1}]\right\} = E[G_t(\Theta_0)'G_t(\Theta_0)] \tag{3.15}$$

In the following, conditions under which all of the elements of $E[G_t(\Theta_0)'G_t(\Theta_0)] < \infty$ are presented. The elements of $G_t(\Theta_0)'$ are expressed, according to the chain rule, as follows:

$$G_t(\Theta_0)' = \begin{bmatrix} \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\theta} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\phi} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\tilde{c}} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\alpha} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\beta} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\tilde{\omega}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial\mu_t}{\partial\theta} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial\mu_t}{\partial\phi} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\mu_t} \times \frac{\partial\mu_t}{\partial\tilde{c}} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\lambda_t} \times \frac{\partial\lambda_t}{\partial\alpha} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\lambda_t} \times \frac{\partial\lambda_t}{\partial\beta} \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial\lambda_t} \times \frac{\partial\lambda_t}{\partial\tilde{\omega}} \end{bmatrix} \tag{3.16}$$

Four panels within the contribution to the information matrix are defined as follows:

$$I_t(\Theta_0) = E[G_t(\Theta_0)'G_t(\Theta_0)|\mathcal{F}_{t-1}] = E\left[\begin{pmatrix} A_{(3\times3)} & C_{(3\times3)} \\ C_{(3\times3)} & B_{(3\times3)} \end{pmatrix}|\mathcal{F}_{t-1}\right] \tag{3.17}$$

where in A only the derivatives of μ_t appear, in B only the derivatives of λ_t appear, and in C the derivatives of μ_t and λ_t appear. By using scalars from A, B and C, the following matrix is defined:

$$F_t = \left\{\begin{matrix} i_3\left[\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \mu_t}\right]^2 i_3' & i_3\left[\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \mu_t} \times \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \lambda_t}\right] i_3' \\ i_3\left[\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \mu_t} \times \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \lambda_t}\right] i_3' & i_3\left[\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \lambda_t}\right]^2 i_3' \end{matrix}\right\} \tag{3.18}$$

where i_3 is a (3×1) vector of ones, and the contribution to the information matrix can be written as:

$$I_t(\Theta_0) = E(F_t|\mathcal{F}_{t-1}) \circ D_t(\Theta_0) = E(F_t|\mathcal{F}_{t-1}) \circ \begin{bmatrix} \tilde{A}_{(3\times3)} & \tilde{C}_{(3\times3)} \\ \tilde{C}_{(3\times3)} & \tilde{B}_{(3\times3)} \end{bmatrix} \tag{3.19}$$

where $\circ$ denotes the Hadamard product, and $D_t(\Theta_0)$ is the outer product of:

$$\tilde{D}_t = [(\partial\mu_t/\partial\theta), (\partial\mu_t/\partial\phi), (\partial\mu_t/\partial\tilde{c}), (\partial\lambda_t/\partial\alpha), (\partial\lambda_t/\partial\beta), (\partial\lambda_t/\partial\tilde{\omega})] \tag{3.20}$$

with itself, i.e. $D_t(\Theta_0) = \tilde{D}_t'\tilde{D}_t$; in panel $\tilde{A}$ only the derivatives of μ_t appear, in panel $\tilde{B}$ only the derivatives of λ_t appear, and in panel $\tilde{C}$ the derivatives of both μ_t and λ_t appear. $D_t(\Theta_0)$ is not in the conditional expectation in Eq. (3.19), since it is determined by $\mathcal{F}_{t-1}$. $\mathcal{I}(\Theta_0)$ can be written as:

$$\mathcal{I}(\Theta_0) = E[I_t(\Theta_0)] = E(F_t) \circ E[D_t(\Theta_0)] + M_t = E(F_t) \circ E\begin{bmatrix} \tilde{A}_{(3\times3)} & \tilde{C}_{(3\times3)} \\ \tilde{C}_{(3\times3)} & \tilde{B}_{(3\times3)} \end{bmatrix} + M_t \tag{3.21}$$

where M_t $(K \times K)$ includes the covariances between the elements of F_t and $D_t(\Theta_0)$.

In the remainder of this section, the conditions of the finiteness of $E(F_t)$, $E[D_t(\Theta_0)]$ and M_t are presented. With respect to the finiteness of $E(F_t)$, matrix F_t can be written as:

$$F_t = \begin{bmatrix} i_3(u_{\mu,t}^2/k_t^2)i_3' & i_3(u_{\mu,t} \times u_{\lambda,t}/k_t)i_3' \\ i_3(u_{\mu,t} \times u_{\lambda,t}/k_t)i_3' & i_3(u_{\lambda,t}^2)i_3' \end{bmatrix} \tag{3.22}$$

where k_t is the time-varying scaling parameter. The form of k_t is different for different DCS models. For example, from Eq. (A.8) of Appendix A, for EGB2-DCS the form of k_t is:

$$\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta_0)}{\partial \mu_t} = u_{\mu,t} \times \{\Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)]\} \exp(2\lambda_t) = \frac{u_{\mu,t}}{k_t} \tag{3.23}$$

For k_t of NIG-DCS and Skew-Gen-t-DCS, see Eqs. (A.22) and (A.34) in Appendix A, respectively. Based on Eq. (3.22), it is necessary that the unconditional means of $(u_{\mu,t}^2/k_t^2)$, $u_{\lambda,t}^2$ and $(u_{\mu,t} \times u_{\lambda,t}/k_t)$ to be finite. This condition is named as Condition 2.

With respect to the finiteness of $E[D_t(\Theta_0)]$, the following equations are used (e.g. Harvey, 2013):

$$\tilde{D}_t' = \begin{bmatrix} \frac{\partial \mu_t}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \phi} \\ \frac{\partial \mu_t}{\partial \tilde{c}} \\ \frac{\partial \lambda_t}{\partial \alpha} \\ \frac{\partial \lambda_t}{\partial \beta} \\ \frac{\partial \lambda_t}{\partial \tilde{\omega}} \end{bmatrix} = \begin{bmatrix} X_{\mu,t-1} \times \frac{\partial \mu_{t-1}}{\partial \theta} + u_{\mu,t-1} \\ X_{\mu,t-1} \times \frac{\partial \mu_{t-1}}{\partial \phi} + \mu_{t-1} - \tilde{c} \\ X_{\mu,t-1} \times \frac{\partial \mu_{t-1}}{\partial \tilde{c}} + 1 - \phi \\ X_{\lambda,t-1} \times \frac{\partial \lambda_{t-1}}{\partial \alpha} + u_{\lambda,t-1} \\ X_{\lambda,t-1} \times \frac{\partial \lambda_{t-1}}{\partial \beta} + \lambda_{t-1} - \tilde{\omega} \\ X_{\lambda,t-1} \times \frac{\partial \lambda_{t-1}}{\partial \tilde{\omega}} + 1 - \beta \end{bmatrix} \tag{3.24}$$

where $X_{\mu,t} \equiv \phi + \theta(\partial u_{\mu,t}/\partial \mu_t)$ and $X_{\lambda,t} \equiv \beta + \alpha(\partial u_{\lambda,t}/\partial \lambda_t)$. Eq. (3.24) provides the following conditions for the finiteness of $E[D_t(\Theta_0)]$: 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 $|E(X_{\mu,t}X_{\lambda,t})| < 1$ (see the proof in Appendix B). In addition, for the DCS model with score-driven μ_t and λ_t, it is also necessary that: (i) 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 that is formed by all possible pairs of those variables is also finite (see the proof in Appendix B); (ii) the unconditional second moment of each element of the outer product of the vector $(X_{\mu,t}, X_{\lambda,t}, u_{\mu,t}, u_{\lambda,t})$ with itself is finite (see the proof in Appendix B). With respect to the latter point, it is noteworthy that the first four moments of ϵ_t are finite for all DCS models of this paper. These conditions are named as Condition 3.

With respect to the finiteness of the covariances within M_t, which represent the covariances between the elements of F_t and $D_t(\Theta_0)$, it is required that the second moments of all variables in the covariances are finite. The variables in the covariances can be seen in Eqs. (3.22) and (3.24). The covariances between the elements of F_t and $D_t(\Theta_0)$ are finite under Condition 3.

3.2. Conditions of (A14)

By using Eq. (3.8), the following result is equivalent to Eq. (3.9):

$$T^{-1/2}\sum_{t=1}^{T} G_t(\Theta_0)' \to_d N_K\left[0_{K\times 1}, \mathcal{I}(\Theta_0)\right] \tag{3.25}$$

By using the Cramér-Wold Device (e.g. White, 1984), Eq. (3.25) is true if for all $a \in \mathbb{R}^K$:

$$\frac{a'T^{-1/2}\sum_{t=1}^{T} G_t(\Theta_0)'}{[a'\mathcal{I}(\Theta_0)a]^{1/2}} \to_d N(0,1) \tag{3.26}$$

where a is a $(K \times 1)$ vector. To show the asymptotic result in Eq. (3.26), for ease of notation $Z_t \equiv a'G_t(\Theta_0)'$ is introduced and Definition 5.15 and Theorem 5.16 from the work of White (1984) are used. According to those two theorems, if (E1) Z_t is stationary and ergodic, (E2) $E(Z_t^2) < \infty$, and (E3) Z_t is an adapted mixingale (with respect to this last point, see Hamilton, 1994, p. 190):

$$\left\{E[E(Z_t|\mathcal{F}_{t-m})^2]\right\}^{1/2} = E|E(Z_t|\mathcal{F}_{t-m})| \leq c_t\gamma_m \text{ as } m \to \infty \text{ and } \gamma_m \to 0 \tag{3.27}$$

where $\mathcal{F}_{t-m}$ has started in the infinite past and is available until period $t-m$ (including period $t-m$), c_t is a finite non-negative sequence and $\gamma_m = O[1/(m^{1+\varepsilon})]$ for $\varepsilon > 0$, then Eq. (3.26) holds.

The proof of Eq. (3.26) uses the following condition: Condition 4 is that ϵ_t is stationary and ergodic. (E1) can be shown as follows: Theorem 3.35 of White (1984) shows that a measurable function transforms stationary and ergodic variables into stationary and ergodic variables. The transformations of ϵ_t in $Z_t = a'G_t(\Theta_0)'$ satisfy the conditions of that theorem under (A1) to (A13) and Conditions 1 to 4. The results of Brandt (1986) and Diaconis and Freedman (1999) on stochastic recurrence equations (SREs) support this conclusion. In relation to those results, Eq. (3.24) can be written as:

$$\begin{bmatrix} \frac{\partial \mu_t}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \phi} \\ \frac{\partial \mu_t}{\partial \tilde{c}} \\ \frac{\partial \lambda_t}{\partial \alpha} \\ \frac{\partial \lambda_t}{\partial \beta} \\ \frac{\partial \lambda_t}{\partial \tilde{\omega}} \end{bmatrix} = \begin{bmatrix} X_{\mu,t-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & X_{\mu,t-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & X_{\mu,t-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & X_{\lambda,t-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & X_{\lambda,t-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & X_{\lambda,t-1} \end{bmatrix} \begin{bmatrix} \frac{\partial \mu_{t-1}}{\partial \theta} \\ \frac{\partial \mu_{t-1}}{\partial \phi} \\ \frac{\partial \mu_{t-1}}{\partial \tilde{c}} \\ \frac{\partial \lambda_{t-1}}{\partial \alpha} \\ \frac{\partial \lambda_{t-1}}{\partial \beta} \\ \frac{\partial \lambda_{t-1}}{\partial \tilde{\omega}} \end{bmatrix} + \begin{bmatrix} u_{\mu,t-1} \\ \mu_{t-1} - \tilde{c} \\ 1 - \phi \\ u_{\lambda,t-1} \\ \lambda_{t-1} - \tilde{\omega} \\ 1 - \beta \end{bmatrix} \tag{3.28}$$

and the following compact notation is used for the previous equation:

$$\tilde{D}'_t = A^*_{t-1}\tilde{D}'_{t-1} + B^*_{t-1} \tag{3.29}$$

which is a SRE. Under (A1) to (A13), Conditions 1 to 4 and by using Theorem 3.35 of White (1984), A^*_{t-1} and B^*_{t-1} are stationary and ergodic. Moreover, from $E(X^2_{\mu,t}) < 1$ and $E(X^2_{\lambda,t}) < 1$ of Condition 3 and by using the Cauchy–Schwarz inequality: $|E(X_{\mu,t})| < 1$ and $|E(X_{\lambda,t})| < 1$ in Eq. (3.28). By using the results of Brandt (1986) and Diaconis and Freedman (1999), $\tilde{D}'_t$ is stationary and ergodic. Under (A1) to (A13), Conditions 1 to 4 and Theorem 3.35 of White (1984), $Z_t = a'G_t(\Theta_0)'$ is also stationary and ergodic. (E2) can be shown as follows:

$$E(Z_t^2) = E\{[a'G_t(\Theta_0)']^2\} = \text{Var}[a'G_t(\Theta_0)'] + E^2[a'G_t(\Theta_0)'] = \tag{3.30}$$

$$a'\text{Var}[G_t(\Theta_0)']a + \{a'E[G_t(\Theta_0)']\}^2 = a'E[G_t(\Theta_0)'G_t(\Theta_0)]a = a'\mathcal{I}(\Theta_0)a < \infty$$

where $E[G_t(\Theta_0)'] = 0$ and the finiteness of $\mathcal{I}(\Theta_0)$ are used under (A1) to (A13) and Conditions 1 to 4.

According to (E3), the m-step ahead forecast $E(Z_t|\mathcal{F}_{t-m})$ converges in absolute expected value to the unconditional mean of zero as $m \to \infty$ (Hamilton, 1994). This can be shown as follows: The unconditional mean of Z_t is zero, because $E(Z_t) = E[a'G_t(\Theta_0)'] = a'E[G_t(\Theta_0)'] = a'0_{K\times 1} = 0$ under (A1) to (A13) and Conditions 1 to 4. Moreover, Eq. (3.16) shows that $G_t(\Theta_0)'$ is the Hadamard product of a $K \times 1$ vector of the score functions and $\tilde{D}'_t$. Under (A1) to (A13) and Conditions 1 to 4, the conditional and unconditional means of the score functions are zero. Therefore, the m-step ahead forecasts of the score functions converge to zero as $m \to \infty$. Under the same conditions, Lemma 6 (Harvey, 2013, p. 36) is applied to $\tilde{D}'_t$, which provides the following forecasting results:

$$\lim_{m\to\infty} E\left(\frac{\partial\mu_t}{\partial\theta}|\mathcal{F}_{t-m}\right) = \lim_{m\to\infty} E\left(\frac{\partial\mu_t}{\partial\phi}|\mathcal{F}_{t-m}\right) = 0 \tag{3.31}$$

$$\lim_{m\to\infty} E\left(\frac{\partial\lambda_t}{\partial\alpha}|\mathcal{F}_{t-m}\right) = \lim_{m\to\infty} E\left(\frac{\partial\lambda_t}{\partial\beta}|\mathcal{F}_{t-m}\right) = 0 \tag{3.32}$$

$$\lim_{m\to\infty} E\left(\frac{\partial\mu_t}{\partial\tilde{c}}|\mathcal{F}_{t-m}\right) = \frac{1-\phi}{1-E(X_{\mu,t})} \tag{3.33}$$

$$\lim_{m\to\infty} E\left(\frac{\partial \lambda_t}{\partial \tilde{\omega}}|\mathcal{F}_{t-m}\right) = \frac{1-\beta}{1-E(X_{\lambda,t})} \quad (3.34)$$

The forecasts of Eqs. (3.31) and (3.32) converge to zero and the forecasts of Eqs. (3.33) and (3.34) converge to a non-zero constant. As a consequence $\lim_{m\to\infty} E[G_t(\Theta_0)'|\mathcal{F}_{t-m}] = 0$,

$$\lim_{m\to\infty} E[Z_t|\mathcal{F}_{t-m}] = \lim_{m\to\infty} E[a'G_t(\Theta_0)'|\mathcal{F}_{t-m}] = \lim_{m\to\infty} a'E[G_t(\Theta_0)'|\mathcal{F}_{t-m}] = 0 \quad (3.35)$$

which implies that $E|E(Z_t|\mathcal{F}_{t-m})| \leq c_t\gamma_m$ as $m \to \infty$ and $\gamma_m \to 0$.

3.3. The asymptotic properties of ML for two models

Proposition 1: For Eqs. (3.10) to (3.12), $\sqrt{T}(\hat{\Theta}_{\mathrm{ML}} - \Theta_0) \to_d N[0_{K\times 1}, \mathcal{I}^{-1}(\Theta_0)]$ as $T \to \infty$, under (A1) to (A15) and the following conditions: Condition 1 is that $|\phi| < 1$ and $|\beta| < 1$. Condition 2 is that $E(u^2_{\mu,t}/k^2_t)$, $E(u^2_{\lambda,t})$ and $E(u_{\mu,t} \times u_{\lambda,t}/k_t)$ are finite. Condition 3 is that $E(X^2_{\mu,t}) < 1$, $E(X^2_{\lambda,t}) < 1$ and $|E(X_{\mu,t}X_{\lambda,t})| < 1$, where $X_{\mu,t} = \phi + \theta(\partial u_{\mu,t}/\partial \mu_t)$ and $X_{\lambda,t} = \beta + \alpha(\partial u_{\lambda,t}/\partial \lambda_t)$. Condition 3 also requires that: (i) 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 that is formed by all possible pairs of those variables is also finite; (ii) the unconditional second moment of each element of the outer product of the vector $(X_{\mu,t}, X_{\lambda,t}, u_{\mu,t}, u_{\lambda,t})$ with itself is finite. Condition 4 is that ϵ_t is strictly stationary and ergodic.

Conditions 1 to 4 can be extended to DCS models with score-driven shape parameters. For example, the following EGB2-DCS model with $\epsilon_t \sim \mathrm{EGB2}[0, 1, \exp(\xi_t), \exp(\zeta_t)]$ is considered:

$$y_t = \mu_t + \exp(\lambda_t)\epsilon_t \quad (3.36)$$

$$\mu_t = c + \phi\mu_{t-1} + \theta u_{\mu,t-1} \quad (3.37)$$

$$\lambda_t = \omega + \beta\lambda_{t-1} + \alpha u_{\lambda,t-1} \quad (3.38)$$

$$\xi_t = \delta_1 + \gamma_1\xi_{t-1} + \kappa_1 u_{\xi,t-1} \quad (3.39)$$

$$\zeta_t = \delta_2 + \gamma_2\zeta_{t-1} + \kappa_2 u_{\zeta,t-1} \quad (3.40)$$

Proposition 2: For Eqs. (3.36) to (3.40), $\sqrt{T}(\hat{\Theta}_{\mathrm{ML}} - \Theta_0) \to_d N[0_{K\times 1}, \mathcal{I}^{-1}(\Theta_0)]$ as $T \to \infty$, under (A1) to (A15) and the following conditions: Condition 1 is that $|\phi| < 1$, $|\beta| < 1$, $|\gamma_1| < 1$ and $|\gamma_2| < 1$. Condition 2 is that $E(u^2_{\mu,t}/k^2_t)$, $E(u^2_{\lambda,t})$, $E(u^2_{\xi,t})$, $E(u^2_{\zeta,t})$, $E(u_{\mu,t} \times u_{\lambda,t}/k_t)$, $E(u_{\mu,t} \times u_{\xi,t}/k_t)$,

$E(u_{\mu,t} \times u_{\zeta,t}/k_t)$, $E(u_{\lambda,t} \times u_{\xi,t})$, $E(u_{\lambda,t} \times u_{\zeta,t})$, and $E(u_{\xi,t} \times u_{\zeta,t})$ are finite. Condition 3 is that $E(X^2_{\mu,t}) < 1$, $E(X^2_{\lambda,t}) < 1$, $E(X^2_{\xi,t}) < 1$, $E(X^2_{\zeta,t}) < 1$, $|E(X_{\mu,t}X_{\lambda,t})| < 1$, $|E(X_{\mu,t}X_{\xi,t})| < 1$, $|E(X_{\mu,t}X_{\zeta,t})| < 1$, $|E(X_{\lambda,t}X_{\xi,t})| < 1$, $|E(X_{\lambda,t}X_{\zeta,t})| < 1$ 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 $X_{\zeta,t} = \gamma_2 + \kappa_2(\partial u_{\zeta,t}/\partial \zeta_t)$. Condition 3 also requires that: (i) 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 that the unconditional mean of each product that is formed by all possible pairs of those variables is also finite; (ii) the unconditional second moment of each element of the outer product of the vector $(X_{\mu,t}, X_{\lambda,t}, X_{\xi,t}, X_{\zeta,t}, u_{\mu,t}, u_{\lambda,t}, u_{\xi,t}, u_{\zeta,t})$ with itself is finite. Condition 4 is that ϵ_t is strictly stationary and ergodic.

3.4. MC simulation experiments

For all MC experiments, zero mean $\mu_t = 0$, unit scale $\exp(\lambda_t) = 1$, and score-driven shape parameters are used for $t = 1, \ldots, T$. Two sets of true parameter values are used: The first set assumes high persistence for the shape parameters (i.e. γ_k=0.95 for all k). The second set assumes low persistence for the shape parameters (i.e. γ_k=0.15 for all k). The true values of all parameters are presented in Table 1. By using those true values, $1,000$ trajectories are simulated and each trajectory includes $T = 10,000$ periods. With respect to the sample size T, it is noteworthy that DCS models with dynamic shape need a large sample size for the reliable estimation of tail shape dynamics.

By using the ML method, the parameters of the DCS models with dynamic shape parameters are estimated for each trajectory. In Table 1, the 5%, 50% and 95% quantiles of the 1,000 parameter estimates are reported. For the high-persistence case, the medians give a good approximation of the true values and the 90% confidence intervals of the quantiles include all true values. For the low-persistence case, the medians give a good approximation of most of the true values; the only exceptions are some of the constant parameters, for which the true value is not within the 90% confidence interval (e.g. δ_2 for EGB2-DCS and NIG-DCS). The 90% confidence intervals also show that the ML estimation is more precise for the high persistence case than it is for the low persistence case. The medians indicate that dynamic parameter γ_k and parameters of the score functions κ_k are precisely estimated for all cases.

4. Data

Daily data are used from the closing prices of the S&P 500 index p_t for the period of January 4, 1950 to December 30, 2017 (source: Bloomberg). In Table 2, descriptive statistics of daily log-returns on the

S&P 500 y_t are presented, where $y_t = \ln(p_t/p_{t-1})$ for $t = 1, \ldots, T$ and pre-sample data are used for p_0. The following results can be highlighted: 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. The evolution of y_t is presented in Fig. 3, where extreme observations are indicated by using the $\overline{\mu} \pm 5\overline{\sigma}$ interval; $\overline{\mu}$ and $\overline{\sigma}$ are the estimates of mean and standard deviation, respectively. In the same figure, the concentration of outliers during the period of the 2008 US financial crisis can be observed.

5. Empirical results

In this section, the ML estimation results for the S&P 500 are presented for the EGB2-DCS (Table 3), NIG-DCS (Table 4) and Skew-Gen-t-DCS (Tables 5(a) and (b)) models. LL-based performances of those models are compared in Table 6. The evolutions of $\rho_{k,t}$ and λ_t are presented for all models in Figs. 4, 5, 6(a) and 6(b). The dates of some extreme events are highlighted in Fig. 7.

Tables 3 to 5 show the following results: For most of the specifications, ϕ parameter which measures the persistence 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, highly significant ω, α, α^* and β parameters are found for the scale. For most of the specifications, the dynamic parameters of shape (i.e. γ_1, γ_2 and γ_3) are significant and positive. For all of the specifications, the scaling parameter of the score function with respect to shape (i.e. κ_1, κ_2 and κ_3) is significantly different from zero (i.e. the DCS models are identified; Harvey, 2013).

All estimates of ϕ, β, γ_1, γ_2 and γ_3 are less than one in absolute value (Tables 3 to 5). Thus, Condition 1 is supported. In Tables 3 to 5, the estimates of $C_\mu = E(X_{\mu,t}^2)$, $C_\lambda = E(X_{\lambda,t}^2)$, $C_{\rho,1} = E(X_{\rho,1,t}^2)$, $C_{\rho,2} = E(X_{\rho,2,t}^2)$, $C_{\rho,3} = E(X_{\rho,3,t}^2)$, $C_{\mu,\lambda} = |E(X_{\mu,t}X_{\lambda,t})|$, $C_{\mu,\rho,1} = |E(X_{\mu,t}X_{\rho,1,t})|$, $C_{\mu,\rho,2} = |E(X_{\mu,t}X_{\rho,2,t})|$, $C_{\mu,\rho,3} = |E(X_{\mu,t}X_{\rho,3,t})|$, $C_{\lambda,\rho,1} = |E(X_{\lambda,t}X_{\rho,1,t})|$, $C_{\lambda,\rho,2} = |E(X_{\lambda,t}X_{\rho,2,t})|$, $C_{\lambda,\rho,3} = |E(X_{\lambda,t}X_{\rho,3,t})|$, $C_{\rho,1,\rho,2} = |E(X_{\rho,1,t}X_{\rho,2,t})|$, $C_{\rho,1,\rho,3} = |E(X_{\rho,1,t}X_{\rho,3,t})|$ and $C_{\rho,2,\rho,3} = |E(X_{\rho,2,t}X_{\rho,3,t})|$ are also reported. 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, the augmented Dickey–Fuller (1979) (hereinafter, ADF) unit root test with constant is performed, which supports those conditions. Condition 4 is a maintained assumption in this paper.

To study assumption (A1), results of the model specification test of Section 2.2 are reported in

Tables 3 to 5. For EGB2-DCS, the MDS null hypothesis of the Escanciano–Lobato test is rejected for all specifications, with respect to skewness and kurtosis. For NIG-DCS, the MDS null hypothesis of the Escanciano–Lobato test is never rejected. For Skew-Gen-t-DCS, eight different specifications with respect to dynamic versus constant shape are considered, and for four out of eight specifications the MDS null hypothesis of the Escanciano–Lobato test is not rejected up to the fourth moment.

In-sample model performances are compared by using the following metrics: LL, Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan-Quinn criterion (HQC), and the likelihood-ratio (LR) test (Tables 3 to 6). For those metrics, the statistical performance of at least one of the DCS specifications with dynamic shape parameters is superior to the statistical performance of the DCS model with constant shape parameters. For the LR test, at least one of the DCS specifications with dynamic shape parameters is significantly superior to the nested DCS specification with constant shape parameters. The results also suggest 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.

The evolutions of $\rho_{k,t}$ and λ_t are presented in Figs. 4 to 6. Those figures indicate that: (i) the shape parameters are time-varying for all models; (ii) for the DCS specifications with dynamic shape, the shape parameters identify the dates of outliers. As an example, in Fig. 7, the identification of outliers is studied for DCS-Skew-Gen-t with constant τ_t, dynamic ν_t and constant η_t. In that figure, some numbers indicate those trading days when ν_t is relatively low, which suggests that the probability of there being an outlier is relatively high. Important events are reported in Appendix C for those trading days that are indicated in Fig. 7. Those events may have significantly impacted the US stock market.

6. VaR backtesting

In this section VaR backtesting applications are presented, in which the VaR measurements of the DCS models with constant and dynamic shape parameters are compared. The results provide the following insight for practitioners about the quality of VaR measurement for the new DCS specifications: The DCS models with dynamic shape effectively predict consecutive additive outliers, while the DCS models with constant shape fail to predict the second outlier.

As aforementioned, extreme observations in the S&P 500 log-returns are concentrated during the period of the 2008 US financial crisis (Fig. 3). This motivates the consideration of the period of September 2, 2008 to March 31, 2009 (i.e. 146 trading days) in the VaR backtesting applications. During that period there were 24 outliers (Fig. 3). The design of the VaR backtesting procedure of this paper is according to the approach of the Basel Committee (1996), in which a 1-day VaR is estimated out-of-sample at the 99% confidence level for each of the most recent 250 trading days. In the present paper, a VaR(1 day, 99%) is estimated for each of the 146 days of the backtesting period. A rolling-window estimation approach is used for all DCS models, and 10,000 observations are included into each rolling window. As alternatives, the use of 2,500 and 5,000 observations is also considered for the rolling windows, but more robust ML estimates are obtained for $T = 10,000$. VaR is approximated after the parameter estimation by using MC simulation, for which 100,000 possible log-returns are simulated for the trading day after the last observation of each rolling window. VaR(1 day, 99%) is defined by the 1% quantile of the log-return simulations. The performance of VaR is compared for the following models: (i) EGB2-DCS with constant and dynamic shape parameters; (ii) NIG-DCS with constant and dynamic shape parameters; (iii) Skew-Gen-t-DCS with constant and dynamic shape parameters. All shape parameters are time-varying for the DCS specifications with dynamic shape parameters.

One of the backtests that are suggested by the Basel Committee (1996) is the 'traffic light approach', which is based on the number of those trading days during the period defined by the last 250 trading days, for which the realized return is lower than the VaR estimate. That number represents the 'VaR failures' during the backtesting period. In the present paper, the VaR failures for the 146-day backtesting period are counted in a similar way. Furthermore, the test of Kupiec (1995) is also applied to evaluate whether the proportion of VaR(1 day, 99%) failures is significantly higher than 1% during the backtesting period. The null hypothesis of the Kupiec test is that the VaR model is appropriate.

The number of VaR failures and the Kupiec test results for the 146-day backtesting period are presented in Table 7(a). The number of VaR failures does not differ for the DCS specifications with constant and dynamic shape. For EGB2-DCS, 4 VaR failures are found for both the constant- and dynamic-shape models. The results also show that VaR for EGB2-DCS is inappropriate according to the Kupiec test, at the 10% level of significance. For NIG-DCS and Skew-Gen-t-DCS, 2 VaR failures are found for both the constant- and dynamic-shape models, and the Kupiec test results support the quality of VaR measurements. Conclusions from Table 7(a): (i) the DCS models with dynamic shape parameters do not provide better VaR measurements than the DCS models with constant shape parameters for the backtesting period of the S&P 500; (ii) the VaR measurements for the NIG-DCS and Skew-Gen-t-DCS models are superior to the VaR measurement for the EGB2-DCS model.

To provide a further analysis of VaR, the evolution of S&P 500 log-returns and the evolution of the VaR estimates are presented for all models in Fig. 8. That figure provides the following interesting insight. It shows that, after each extreme observation, the predicted potential extreme loss of the next day is higher for the DCS specification with dynamic shape parameter than for the DCS specification with constant shape parameter. This result is robust, because it is obtained for the EGB2, NIG and Skew-Gen-t distributions, and it provides the following intuition: If outliers occur on consecutive trading days, then: (i) the VaR measurements for DCS models with constant shape parameters may fail for the second outlier; (ii) the VaR measurements for DCS models with dynamic shape parameters will detect the second outlier. The correct detection of consecutive additive outliers is important in the VaR backtesting literature, and it motivated the work of Christofferssen (1998). The null hypothesis of the Christofferssen test is that the arrival times of VaR failures are independent. If that null hypothesis is rejected, for example, due to consecutive VaR failures within the backtesting period, then the econometric model is not updated correctly after extreme observations.

To study the issue of consecutive outliers, a modified S&P 500 dataset is used for the backtesting period. For each of those days when the VaR fails (see the dates in Table 7(a)), the outlier is duplicated for the next day. The results for the 146-day backtesting period are presented in Table 7(b). For EGB2-DCS, 3 VaR failures are found for both the constant- and dynamic-shape models. Although the number of VaR failures is identical for both models, two of those VaR failures are on consecutive trading days for the constant-shape model, while the three VaR failures are on non-consecutive trading days for the dynamic-shape model. For NIG-DCS and Skew-Gen-t-DCS, 3 VaR failures are found for the constant-shape models and 2 VaR failures are found for the dynamic-shape models. It is important to highlight that consecutive VaR failures are observed for all DCS models with constant shape parameters, while the VaR failures are never on consecutive days for the DCS models with dynamic shape parameters. The VAR backtesting application indicates that, during periods of high market volatility, the VaR measurements of DCS models with dynamic shape parameters are protected against consecutive additive outliers, but the VaR measurements of DCS models with constant shape parameters fail on the trading day of the second additive outlier.

7. Conclusions

In this paper, new score-driven models with dynamic shape parameters have been suggested, which improve the performance of the score-driven models with constant shape parameters, since: (i) they have superior in-sample statistical performances, and (ii) they provide more accurate out-of-sample VaR measurements. Log-return time series data have been used from the S&P 500 index for the period of 1950 to 2017. All DCS models have been estimated by using the ML method, the conditions of the asymptotic properties of the ML estimator have been provided, and the use of the ML estimator has been supported by performing MC simulation experiments. The in-sample statistical performances of DCS models with dynamic shape parameters have been superior to the in-sample statistical performances of DCS models with constant shape parameters. VaR backtesting for a period during the 2008 US financial crisis has indicated that the DCS models with dynamic shape parameters effectively predict extreme losses for consecutive additive outliers, while the DCS models with constant shape parameters are incorrectly updated after the first outlier and fail to predict extreme losses for the second outlier. The in-sample statistical performance and out-of-sample VAR backtesting results of this paper may motivate the consideration of DCS models with dynamic shape parameters in the VaR measurement practices of financial risk managers for periods of high market volatility.

Acknowledgments

Previous versions of this paper were presented at the GESG Research Seminar (Universidad Francisco Marroquín, March 7, 2019), Cambridge-INET Conference on Score-Driven and Nonlinear Time Series Models (University of Cambridge, March 27-29, 2019), and IX-th Workshop in Time Series Econometrics (University of Zaragoza, April 4-5, 2019). The authors are thankful to Matthew Copley, Juan Carlos Escanciano, Andrew Harvey, Jason Jones and to all seminar and conference participants for the helpful comments and suggestions. Astrid Ayala and Szabolcs Blazsek acknowledge funding from the School of Business of Universidad Francisco Marroquín. Alvaro Escribano acknowledges funding from the Spanish Ministry of Economy, Industry and Competitiveness (ECO2015-68715-R, ECO2016-00105-001), Consolidation Grant (#2006/04046/002), and Maria de Maeztu Grant (MDM 2014-0431).

References

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.

Basel Committee, 1996. Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements. Available at www.bis.org.

Blasques, F., Koopman, S.J., Lucas, A., 2015. Information-theoretic optimality of observation-driven time series models for continuous responses. Biometrika 102 (2): 325–343. doi: 10.1093/biomet/asu076.

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., 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.

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 $Y_{n+1} = A_n Y_n + B_n$ 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.

Christofferssen, P., 1998. Evaluating interval forecasts. International Economic Review 39(4): 841–862. doi: 10.2307/2527341.

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.

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.

Kupiec, P.H., 1995. Techniques for verifying the accuracy of risk measurement models. The Journal of Derivatives 3(2): 73–84. doi: 10.3905/jod.1995.407942.

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.

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.

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

Appendix A

In this appendix, for each error specification, 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 $u_{\rho,k,t}$ are presented.

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

$$E(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \Psi^{(0)}[\exp(\xi_t)] - \Psi^{(0)}[\exp(\zeta_t)] \quad (A.1)$$

$$\text{Var}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)] \quad (A.2)$$

$$\text{Skew}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \Psi^{(2)}[\exp(\xi_t)] - \Psi^{(2)}[\exp(\zeta_t)] \quad (A.3)$$

$$\text{Kurt}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \Psi^{(3)}[\exp(\xi_t)] + \Psi^{(3)}[\exp(\zeta_t)] \quad (A.4)$$

respectively; Θ is the vector of parameters and $\Psi^{(i)}(x)$ is the polygamma function of order i. For the EGB2-DCS model, $y_t|\mathcal{F}_{t-1} \sim \text{EGB2}[\mu_t, \exp(-\lambda_t), \exp(\xi_t), \exp(\zeta_t)]$. The conditional mean and the conditional volatility of y_t are

$$E(y_t|\mathcal{F}_{t-1};\Theta) = \mu_t + \exp(\lambda_t)\left\{\Psi^{(0)}[\exp(\xi_t)] - \Psi^{(0)}[\exp(\zeta_t)]\right\} \quad (A.5)$$

$$SD(y_t|\mathcal{F}_{t-1};\Theta) = \exp(\lambda_t)\{\Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)]\}^{1/2} \tag{A.6}$$

respectively. The log of the conditional density of y_t is

$$\ln f(y_t|\mathcal{F}_{t-1};\Theta) = \exp(\xi_t)\epsilon_t - \lambda_t - \ln\Gamma[\exp(\xi_t)] - \ln\Gamma[\exp(\zeta_t)] \tag{A.7}$$

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

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

$$\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t} = u_{\mu,t} \times \{\Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)]\}\exp(2\lambda_t) = \frac{u_{\mu,t}}{k_t} \tag{A.8}$$

where

$$u_{\mu,t} = \{\Psi^{(1)}[\exp(\xi_t)] + \Psi^{(1)}[\exp(\zeta_t)]\}\exp(\lambda_t)\left\{[\exp(\xi_t)+\exp(\zeta_t)]\frac{\exp(\epsilon_t)}{\exp(\epsilon_t)+1} - \exp(\xi_t)\right\} \tag{A.9}$$

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

$$u_{\lambda,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \lambda_t} = [\exp(\xi_t)+\exp(\zeta_t)]\frac{\epsilon_t\exp(\epsilon_t)}{\exp(\epsilon_t)+1} - \exp(\xi_t)\epsilon_t - 1 \tag{A.10}$$

Thirdly, the score function with respect to ξ_t is

$$u_{\xi,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \xi_t} = \exp(\xi_t)\epsilon_t - \exp(\xi_t)\Psi^{(0)}[\exp(\xi_t)] \tag{A.11}$$

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

Fourthly, the score function with respect to ζ_t is

$$u_{\zeta,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \zeta_t} = -\exp(\zeta_t)\Psi^{(0)}[\exp(\zeta_t)] \tag{A.12}$$

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

(2) For the NIG-DCS model, $\epsilon_t \sim \text{NIG}[0, 1, \exp(\nu_t), \exp(\nu_t)\tanh(\eta_t)]$, 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. The conditional mean, conditional variance, conditional skewness and conditional kurtosis of ϵ_t are given by:

$$E(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \frac{\tanh(\eta_t)}{[1-\tanh^2(\eta_t)]^{1/2}} \tag{A.13}$$

$$\text{Var}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \frac{\exp(-\nu_t)}{[1-\tanh^2(\eta_t)]^{3/2}} \tag{A.14}$$

$$\text{Skew}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \frac{3\tanh(\eta_t)}{\exp(\nu_t/2)\left[1-\tanh^2(\eta_t)\right]^{1/4}} \tag{A.15}$$

$$\text{Kurt}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = 3 + \frac{3\left[1+4\tanh^2(\eta_t)\right]}{\exp(\nu_t)\left[1-\tanh^2(\eta_t)\right]^{1/2}} \tag{A.16}$$

respectively. For the NIG-DCS model, the conditional distribution of y_t is

$$y_t|\mathcal{F}_{t-1} \sim \text{NIG}[\mu_t, \exp(\lambda_t), \exp(\nu_t-\lambda_t), \exp(\nu_t-\lambda_t)\tanh(\eta_t)] \tag{A.17}$$

The conditional mean and the conditional volatility of y_t are

$$E(y_t|\mathcal{F}_{t-1};\Theta) = \mu_t + \frac{\exp(\lambda_t)\tanh(\eta_t)}{[1-\tanh^2(\eta_t)]^{1/2}} \tag{A.18}$$

$$SD(y_t|\mathcal{F}_{t-1};\Theta) = \left\{\frac{\exp(2\lambda_t - \nu_t)}{[1-\tanh^2(\eta_t)]^{3/2}}\right\}^{1/2} \tag{A.19}$$

respectively. The log of the conditional density of y_t is

$$\ln f(y_t|\mathcal{F}_{t-1};\Theta) = \nu_t - \lambda_t - \ln(\pi) + \exp(\nu_t)[1-\tanh^2(\eta_t)]^{1/2} \tag{A.20}$$

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

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. Firstly, the score function with respect to μ_t is

$$\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t} = -\exp(\nu_t - \lambda_t)\tanh(\eta_t) + \frac{\epsilon_t}{\exp(\lambda_t)(1+\epsilon_t^2)} \tag{A.21}$$

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

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

$$u_{\mu,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t} \times \exp(2\lambda_t) = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t} \times k_t \tag{A.22}$$

Secondly, the score function with respect to λ_t is

$$u_{\lambda,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \lambda_t} = -1 - \exp(\nu_t)\tanh(\eta_t)\epsilon_t + \frac{\epsilon_t^2}{1+\epsilon_t^2} \tag{A.23}$$

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

Thirdly, the score function with respect to ν_t is

$$u_{\nu,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \nu_t} = 1 + \exp(\nu_t)[1-\tanh^2(\eta_t)]^{1/2} + \exp(\nu_t)\tanh(\eta_t)\epsilon_t \tag{A.24}$$

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

Fourthly, the score function with respect to η_t is

$$u_{\eta,t} = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \eta_t} = \exp(\nu_t)\mathrm{sech}^2(\eta_t)\epsilon_t - \exp(\nu_t)\tanh(\eta_t)\mathrm{sech}(\eta_t) \tag{A.25}$$

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

(3) 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. The conditional mean, conditional variance, conditional skewness and conditional kurtosis of ϵ_t, respectively, are:

$$E(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \frac{2\tanh(\tau_t)[\exp(\nu_t)+4]^{\exp(-\eta_t)}B\left\{\frac{2}{\exp(\eta_t)}, \frac{\exp(\nu_t)+3}{exp(\eta_t)}\right\}}{B\left\{\frac{1}{\exp(\eta_t)}, \frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right\}} \tag{A.26}$$

$$\mathrm{Var}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = [\exp(\nu_t)+4]^{2\exp(-\eta_t)} \times \tag{A.27}$$

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

$$\text{Skew}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \frac{2\tanh(\tau_t)[\exp(\nu_t)+4]^{3\exp(-\eta_t)}}{B^3\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]}\times \tag{A.28}$$

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

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

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

$$\text{Kurt}(\epsilon_t|\mathcal{F}_{t-1};\Theta) = \frac{[\exp(\nu_t)+4]^{4\exp(-\eta_t)}}{B^4\left[\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right]}\times \tag{A.29}$$

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

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

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

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

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

respectively; $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the beta function and $\Gamma(x)$ is the gamma function. For the Skew-Gen-t-DCS model, the conditional distribution of y_t is

$$y_t|\mathcal{F}_{t-1} \sim \text{Skew-Gen-}t[\mu_t, \exp(\lambda_t), \tanh(\tau_t), \exp(\nu_t)+4, \exp(\eta_t)] \tag{A.30}$$

The conditional mean of y_t is

$$E(y_t|\mathcal{F}_{t-1};\Theta) = \mu_t + 2\exp(\lambda_t)\tanh(\tau_t)[\exp(\nu_t)+4]^{\exp(-\eta_t)} \times \frac{B\left\{\frac{2}{\exp(\eta_t)},\frac{\exp(\nu_t)+3}{\exp(\eta_t)}\right\}}{B\left\{\frac{1}{\exp(\eta_t)},\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right\}} \tag{A.31}$$

The conditional volatility of y_t is

$$SD(y_t|\mathcal{F}_{t-1};\Theta) = \exp(\lambda_t)[\exp(\nu_t)+4]^{\exp(-\eta_t)}\times \tag{A.32}$$

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

The log of the conditional density of y_t is

$$\ln f(y_t|\mathcal{F}_{t-1};\Theta) = \eta_t - \lambda_t - \ln(2) - \frac{\ln[\exp(\nu_t)+4]}{\exp(\eta_t)} - \ln\Gamma\left[\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right] \tag{A.33}$$

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

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

Firstly, the score function with respect to μ_t is

$$\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t}= \tag{A.34}$$

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

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

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

$$u_{\lambda,t}=\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \lambda_t}=\frac{|\epsilon_t|^{\exp(\eta_t)}[\exp(\nu_t)+5]}{|\epsilon_t|^{\exp(\eta_t)}+[1+\tanh(\tau_t)\mathrm{sgn}(\epsilon_t)]^{\exp(\eta_t)}[\exp(\nu_t)+4]}-1 \tag{A.35}$$

Thirdly, the score function with respect to τ_t is

$$u_{\tau,t}=\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \tau_t}=\frac{[\exp(\nu_t)+5]|\epsilon_t|^{\exp(\eta_t)}\mathrm{sgn}(\epsilon_t)\mathrm{sech}(\tau_t)}{[\mathrm{sgn}(\epsilon_t)\mathrm{sinh}(\tau_t)+\cosh(\tau_t)]}\times \tag{A.36}$$

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

Fourthly, the score function with respect to ν_t is

$$u_{\nu,t}=\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \nu_t}=-\frac{\exp(\nu_t-\eta_t)}{\exp(\nu_t)+4}-\exp(\nu_t-\eta_t)\Psi^{(0)}\left[\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right] \tag{A.37}$$

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

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

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

Fifthly, the score function with respect to η_t is

$$u_{\eta,t}=\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \eta_t}=1+\frac{\ln[\exp(\nu_t)+4]}{\exp(\eta_t)}+\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\Psi^{(0)}\left[\frac{\exp(\nu_t)+4}{\exp(\eta_t)}\right] \tag{A.38}$$

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

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

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

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

Appendix B

In this appendix, the conditions under which the expected value of each of the nine elements of $\tilde{C}$ from Eq. (3.21) is finite are studied. $\tilde{C}$ is 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.

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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (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 Eq. (B.1), 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 Eq. (B.1) is constant under Condition 3. By using the law of iterated expectations in Eq. (B.1), covariances appear on the right side of the equation. Those covariances are finite 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.2), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the third term on the right side of Eq. (B.2), 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 Eq. (B.2) is zero, since $E(\lambda_t - \tilde{\omega}) = 0$. By using the law of iterated expectations in Eq. (B.2), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. For the second term on the right side of Eq. (B.2), write the expectation:

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.3) is the first lag of the second term on the right side of Eq. (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 side 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 side is constant under Condition 3. By using the law of iterated expectations in Eq. (B.3), covariances appear on the right side of the equation. Those covariances are finite 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.4), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the second term on the right side of Eq. (B.4), 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 Eq. (B.4), 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 Eq. (B.4), the law of iterated expectations gives zero, because $E(u_{\mu,t}) = 0$. By using the law of iterated expectations in Eq. (B.4), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.5), the absolute value of the autoregressive parameter is < 1 under Condition 3. For the second term on the right side of Eq. (B.5), 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 Eq. (B.5), the law of iterated expectations gives zero, because $E(u_{\lambda,t})=0$. By using the law of iterated expectations in Eq. (B.5), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. For the third term on the right side of Eq. (B.5), write the expectation:

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.6) is the first lag of the third term on the right side of Eq. (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 Eq. (B.6) is zero, since $E(\mu_t-\tilde{c})=0$. The third term on the right side of Eq. (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 side of Eq. (B.6) is constant under Condition 3. By using the law of iterated expectations in Eq. (B.6), covariances appear on the right side of the equation. Those covariances are finite 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.7), the absolute value of the autoregressive parameter is <1 under Condition 3. By using the law of iterated expectations in Eq. (B.7), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. In the following, the covariance stationarity of the (i) second, (ii) third and (iii) fourth terms of Eq. (B.7), respectively, is analyzed:

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

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.8) is the first lag of the second term on the right side of Eq. (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 Eq. (B.8) is the first lag of the fourth term on the right side of Eq. (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 Eq. (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 Eq. (B.8) is zero, because $E(\mu_t-\tilde{c})=0$. By using the law of iterated expectations in Eq. (B.8), covariances appear on the right side of the equation. Those covariances are finite under Condition 3.

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

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.9) is the first lag of the third term on the right side of Eq. (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 Eq. (B.9) is the first lag of the fourth term on the right side of Eq. (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 Eq. (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 Eq. (B.9) is zero, because $E(\lambda_t - \tilde{\omega}) = 0$. By using the law of iterated expectations in Eq. (B.9), covariances appear on the right side of the equation. Those covariances are finite under Condition 3.

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

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.10) is the first lag of the fourth term on the right side of Eq. (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 Eq. (B.10) are zero, because $E(\mu_t - \tilde{c}) = E(\lambda_t - \tilde{\omega}) = 0$. The fourth term on the right side of Eq. (B.10) is constant under Condition 3. By using the law of iterated expectations in Eq. (B.10), covariances appear on the right side of the equation. Those covariances are finite 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.11), the absolute value of the autoregressive parameter is < 1 under Condition 3. The second term on the right side of Eq. (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 Eq. (B.11) is zero, because $E(\mu_t - \tilde{c}) = 0$. By using the law of iterated expectations in Eq. (B.11), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. For the third term on the right side of Eq. (B.11), write the expectation:

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.12) is the first lag of the third term on the right side of Eq. (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 Eq. (B.12) is zero, because $E(\mu_t - \tilde{c}) = 0$. The third term on the right side of Eq. (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 Eq. (B.12) is zero, because $E(u_{\mu,t}) = 0$. By using the law of iterated expectations in Eq. (B.12), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.13), the absolute value of the autoregressive parameter is < 1 under Condition 3. The second term on the right side of Eq. (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 Eq. (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 Eq. (B.13) is zero, because $E(u_{\lambda,t}) = 0$. By using the law of iterated expectations in Eq. (B.13), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.14), the absolute value of the autoregressive parameter is < 1 under Condition 3. The third term on the right side of Eq. (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 Eq. (B.14) is zero, because $E(\lambda_t - \tilde{\omega}) = 0$. By using the law of iterated expectations in Eq. (B.14), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. For the second term on the right side of Eq. (B.14), write the expectation:

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

The law of iterated expectations is used for the previous equation. The first term on the right side of Eq. (B.15) is the first lag of the second term on the right side of Eq. (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 Eq. (B.15) is zero, because $E(\lambda_{t-2} - \tilde{\omega}) = 0$. The third term on the right side of Eq. (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 Eq. (B.15) is zero, because $E(u_{\lambda,t}) = 0$. By using the law of iterated expectations in Eq. (B.15), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. 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:

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

The law of iterated expectations is used for the previous equation. For the first term on the right side of Eq. (B.16), the absolute value of the autoregressive parameter is < 1 under Condition 3. The second term on the right side of Eq. (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 Eq. (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 Eq. (B.16) is constant. By using the law of iterated expectations in Eq. (B.16), covariances appear on the right side of the equation. Those covariances are finite under Condition 3. Thus, $E[(\partial\mu_t/\partial\tilde{c}) \times (\partial\lambda_t/\partial\tilde{\omega})]$ is finite.

Appendix C

In this appendix, the circumstances of the extreme events that are numbered in Figure 7 are described.

(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 that was 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 the neighbourhood of Love Canal in Niagara Falls, New York, 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 the Pinochet regime in Chile. July 7, 1986: the Supreme Court struck down the 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 withdraw its troops from Afghanistan. April 7, 1988: the US performed a nuclear test at Nevada Test Site. April 9, 1988: the 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 a 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 USD 6.75 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 explosion 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 allow 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.

Table 1. Monte Carlo simulation experiments for DCS models with dynamic shape parameters.

EGB2-DCS	True values #1	5% quantile	Median	95% quantile	True values #2	5% quantile	Median	95% quantile
δ_1	−0.0500	−0.0721	−0.0526	−0.0375	−0.0500	−0.1045	−0.0753	−0.0413
γ_1	0.9500	0.9307	0.9490	0.9636	0.1500	0.0004	0.1509	0.5167
κ_1	0.0500	0.0399	0.0498	0.0606	0.0500	0.0331	0.0504	0.0686
δ_2	−0.0500	−0.0596	−0.0501	−0.0415	−0.0500	−0.0468	−0.0251	−0.0065
γ_2	0.9500	0.9407	0.9500	0.9583	0.1500	0.0002	0.1529	0.5239
κ_2	−0.0500	−0.0582	−0.0500	−0.0424	−0.0500	−0.0678	−0.0494	−0.0304
NIG-DCS	True values #1	5% quantile	Median	95% quantile	True values #2	5% quantile	Median	95% quantile
δ_1	0.0500	0.0290	0.0616	0.1287	0.0500	0.0146	0.0642	0.1051
γ_1	0.9500	0.8917	0.9479	0.9754	0.1500	0.0007	0.1495	0.7644
κ_1	0.0500	0.0263	0.0445	0.0673	0.0500	0.0001	0.0481	0.1193
δ_2	−0.0500	−0.0837	−0.0674	−0.0541	−0.0500	−0.1750	−0.1406	−0.0918
γ_2	0.9500	0.9284	0.9420	0.9534	0.1500	0.0002	0.1445	0.4403
κ_2	0.0500	0.0399	0.0452	0.0509	0.0500	0.0328	0.0487	0.0647
Skew-Gen-*t*-DCS	True values #1	5% quantile	Median	95% quantile	True values #2	5% quantile	Median	95% quantile
δ_1	−0.0200	−0.0352	−0.0294	−0.0133	−0.04	−0.1057	−0.0859	−0.0669
γ_1	0.9500	0.9394	0.9489	0.9587	0.15	0.0002	0.1511	0.3201
κ_1	0.0500	0.0445	0.0485	0.0542	0.05	0.0405	0.0494	0.0580
δ_2	0.0800	0.0367	0.0729	0.2177	1.3	0.5043	1.0005	1.4774
γ_2	0.9500	0.9466	0.9495	0.9854	0.15	0.1061	0.1382	0.1524
κ_2	0.0500	0.0479	0.0501	0.2987	0.05	0.0404	0.0497	0.0591
δ_3	0.0300	0.0180	0.0317	0.0709	0.05	0.0175	0.0828	0.1621
γ_3	0.9500	0.9082	0.9492	0.9765	0.15	0.0005	0.1504	0.6737
κ_3	0.0500	0.0319	0.0501	0.0782	0.05	0.0040	0.0512	0.0981

Notes: Number of simulated trajectories: 1,000; sample size: $T = 10,000$. For all models $\mu_t = 0$ and $\exp(\lambda_t) = 1$ for $t = 1, \dots, T$. True values #1 imply high persistence for the shape parameters. True values #2 imply low persistence for the shape parameters. The data generating processes are given by: For the EGB2-DCS model, $y_t \sim$ EGB2$[0, 1, \exp(\xi_t), \exp(\zeta_t)]$, where $\xi_t = \delta_1 + \gamma_1 \xi_{t-1} + \kappa_1 u_{\xi,t-1}$ and $\zeta_t = \delta_2 + \gamma_2 \zeta_{t-1} + \kappa_2 u_{\zeta,t-1}$. For the NIG-DCS model, $y_t \sim$ NIG$[0, 1, \exp(\nu_t), \exp(\nu_t)\tanh(\eta_t)]$, where $\nu_t = \delta_1 + \gamma_1 \nu_{t-1} + \kappa_1 u_{\nu,t-1}$ and $\eta_t = \delta_2 + \gamma_2 \eta_{t-1} + \kappa_2 u_{\eta,t-1}$. 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 $\tau_t = \delta_1 + \gamma_1 \tau_{t-1} + \kappa_1 u_{\tau,t-1}$, $\nu_t = \delta_1 + \gamma_1 \nu_{t-1} + \kappa_1 u_{\nu,t-1}$ and $\eta_t = \delta_2 + \gamma_2 \eta_{t-1} + \kappa_2 u_{\eta,t-1}$. For the initial condition, the unconditional mean of each dynamic shape parameter is used. For example, $\xi_1 = \delta_1/(1 - \gamma_1)$.

Table 2. Descriptive statistics for daily log-returns on the S&P 500 index, $y_t = \ln(p_t/p_{t-1})$.

Start date	January 4, 1950
End date	December 30, 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
$\text{Corr}(y_t, y_{t-1})$	0.0269
$\text{Corr}(y_t^2, y_{t-1})$	-0.0877

Source of data: Bloomberg

Table 3. Parameter estimates and model specification tests, 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 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 4. Parameter estimates and model specification tests, 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 5(a). Parameter estimates and model specification tests, 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 5(b). Parameter estimates and model specification tests, 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 6. AIC-, BIC- and HQC-based comparison of statistical performance.

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

Notes: ✓indicates that all MDS model specification tests support the model (see Tables 3 to 5).

Table 7(a). VaR backtesting for log-returns on the S&P 500.

Start date of backtesting period	September 2, 2008					
End date of backtesting period	March 31, 2009					
Number of trading days	146					
EGB2-DCS	Constant shape			Dynamic shape		
Number of VaR failures	4			4		
Kupiec test statistic	3.0278			3.0278		
Kupiec test p-value	0.0819			0.0819		
	Date	Log-return	MC-VaR(1 day, 99%)	Date	Log-return	MC-VaR(1 day, 99%)
First failure	September 15, 2008	−0.0486	−0.0396	September 9, 2018	−0.0356	−0.0349
Second failure	September 17, 2008	−0.0490	−0.0485	September 15, 2018	−0.0486	−0.0396
Third failure	September 29, 2008	−0.0926	−0.0573	September 29, 2008	−0.0926	−0.0566
Fourth failure	October 15, 2008	−0.0948	−0.0947	October 15, 2008	−0.0948	−0.0934
NIG-DCS	Constant shape			Dynamic shape		
Number of VaR failures	2			2		
Kupiec test statistic	0.1809			0.1809		
Kupiec test p-value	0.6706			0.6706		
	Date	Log-return	MC-VaR(1 day, 99%)	Date	Log-return	MC-VaR(1 day, 99%)
First failure	September 15, 2018	−0.0486	−0.0413	September 15, 2018	−0.0486	−0.0418
Second failure	September 29, 2008	−0.0926	−0.0597	September 29, 2008	−0.0926	−0.0604
Skew-Gen-t-DCS	Constant shape			Dynamic shape		
Number of VaR failures	2			2		
Kupiec test statistic	0.1809			0.1809		
Kupiec test p-value	0.6706			0.6706		
	Date	Log-return	MC-VaR(1 day, 99%)	Date	Log-return	MC-VaR(1 day, 99%)
First failure	September 15, 2018	−0.0486	−0.0407	September 15, 2018	−0.0486	−0.0430
Second failure	September 29, 2008	−0.0926	−0.0590	September 29, 2008	−0.0926	−0.0593

Table 7(b). VaR backtesting for log-returns on the S&P 500 with consecutive additive outliers.

Start date of backtesting period	September 2, 2008					
End date of backtesting period	March 31, 2009					
Number of trading days	146					
EGB2-DCS	Constant shape			Dynamic shape		
Number of VaR failures	3			3		
Kupiec test statistic	1.2575			1.2575		
Kupiec test p-value	0.2621			0.2621		
	Date	Log-return	MC-VaR(1 day, 99%)	Date	Log-return	MC-VaR(1 day, 99%)
First failure	September 15, 2008	−0.0486	−0.0397	September 9, 2008	−0.0356	−0.0354
Second failure	September 29, 2008	−0.0926	−0.0696	September 15, 2008	−0.0486	−0.0439
Third failure	September 30, 2008	−0.0926	−0.0903	September 29, 2008	−0.0926	−0.0608
NIG-DCS	Constant shape			Dynamic shape		
Number of VaR failures	3			2		
Kupiec test statistic	1.2575			0.1809		
Kupiec test p-value	0.2621			0.6706		
	Date	Log-return	MC-VaR(1 day, 99%)	Date	Log-return	MC-VaR(1 day, 99%)
First failure	September 15, 2008	−0.0486	−0.0413	September 15, 2008	−0.0486	−0.0420
Second failure	September 29, 2008	−0.0926	−0.0655	September 29, 2008	−0.0926	−0.0641
Third failure	September 30, 2008	−0.0926	−0.0877			
Skew-Gen-t-DCS	Constant shape			Dynamic shape		
Number of VaR failures	3			2		
Kupiec test statistic	1.2575			0.1809		
Kupiec test p-value	0.2621			0.6706		
	Date	Log-return	MC-VaR(1 day, 99%)	Date	Log-return	MC-VaR(1 day, 99%)
First failure	September 15, 2008	−0.0486	−0.0407	September 15, 2008	−0.0486	−0.0438
Second failure	September 29, 2008	−0.0926	−0.0653	September 29, 2008	−0.0926	−0.0618
Third failure	September 30, 2008	−0.0926	−0.0874			

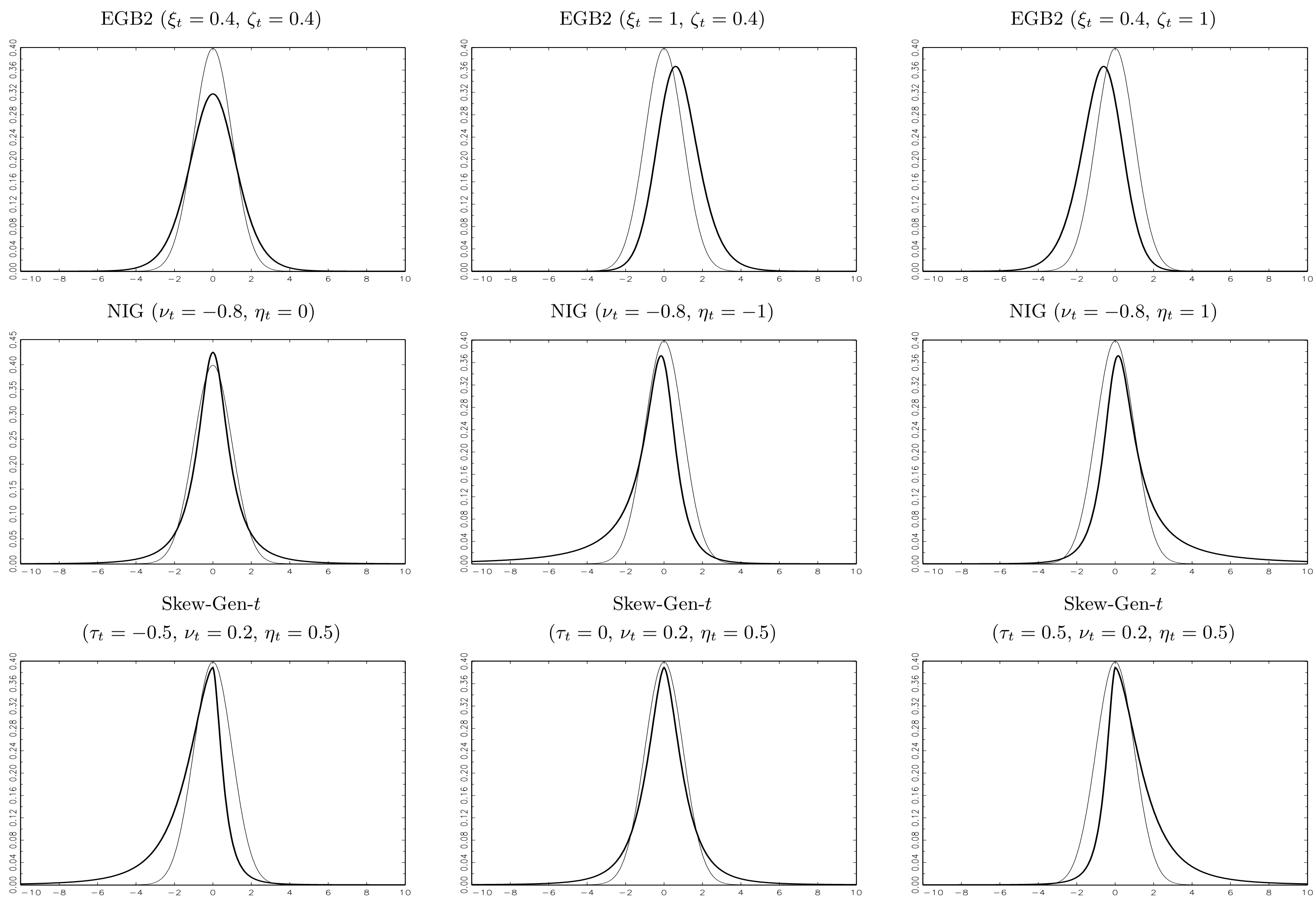

Fig. 1. Density functions for the EGB2, NIG, Skew-Gen-t distributions (thick lines) and the N(0,1) distribution (thin lines).

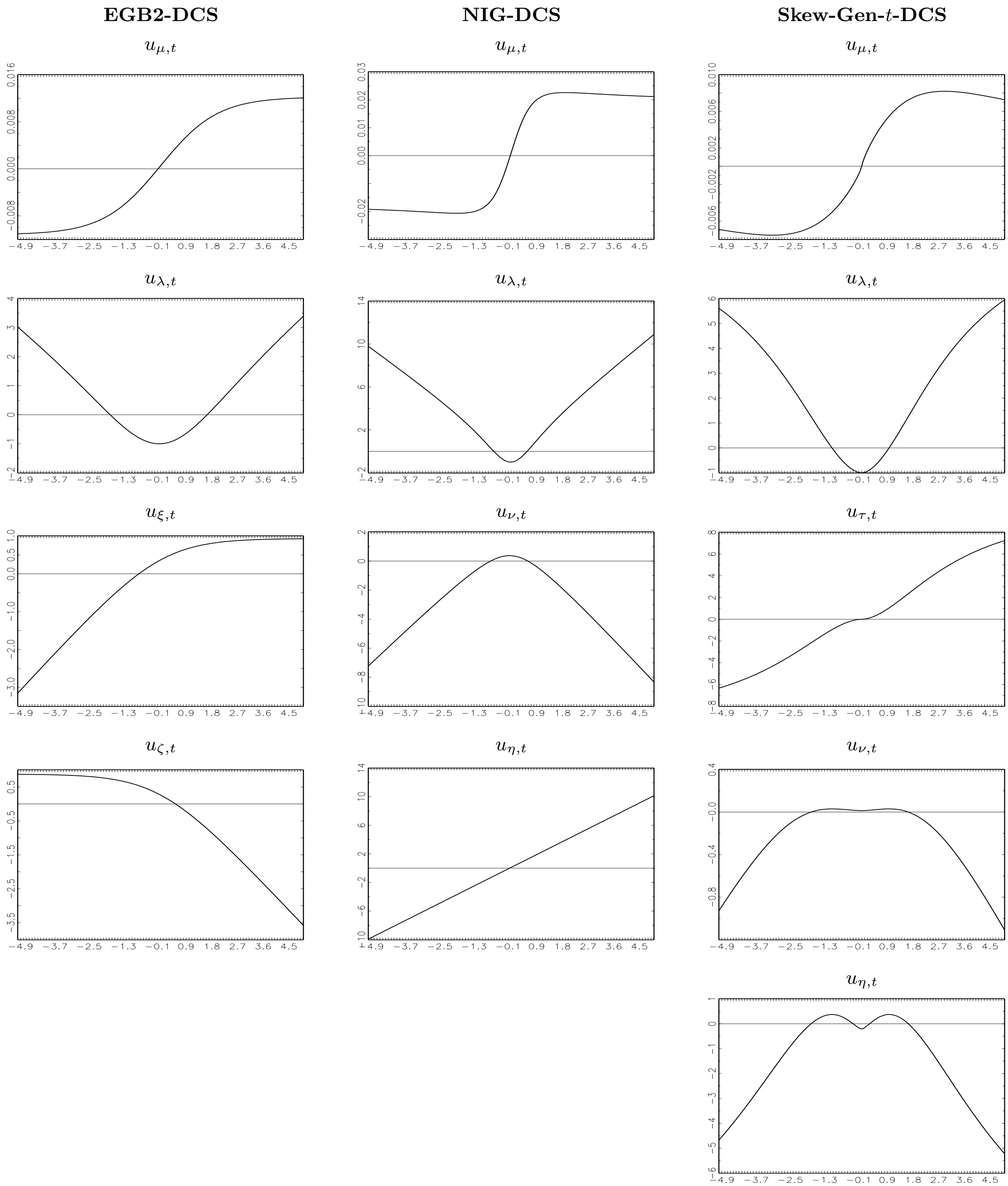

Fig. 2. Score functions of DCS models with dynamic shape parameters.

Notes: In the DCS models of this figure, all of the shape parameters are dynamic. The score functions are presented as a function of ϵ_t. In the calculations of this figure, the score function formulas of Appendix A are used, in which the dynamic parameters are replaced by the estimates of their unconditional means (e.g. λ_t is replaced by $\hat{E}(\lambda_t) = \hat{\omega}/(1-\hat{\beta})$).

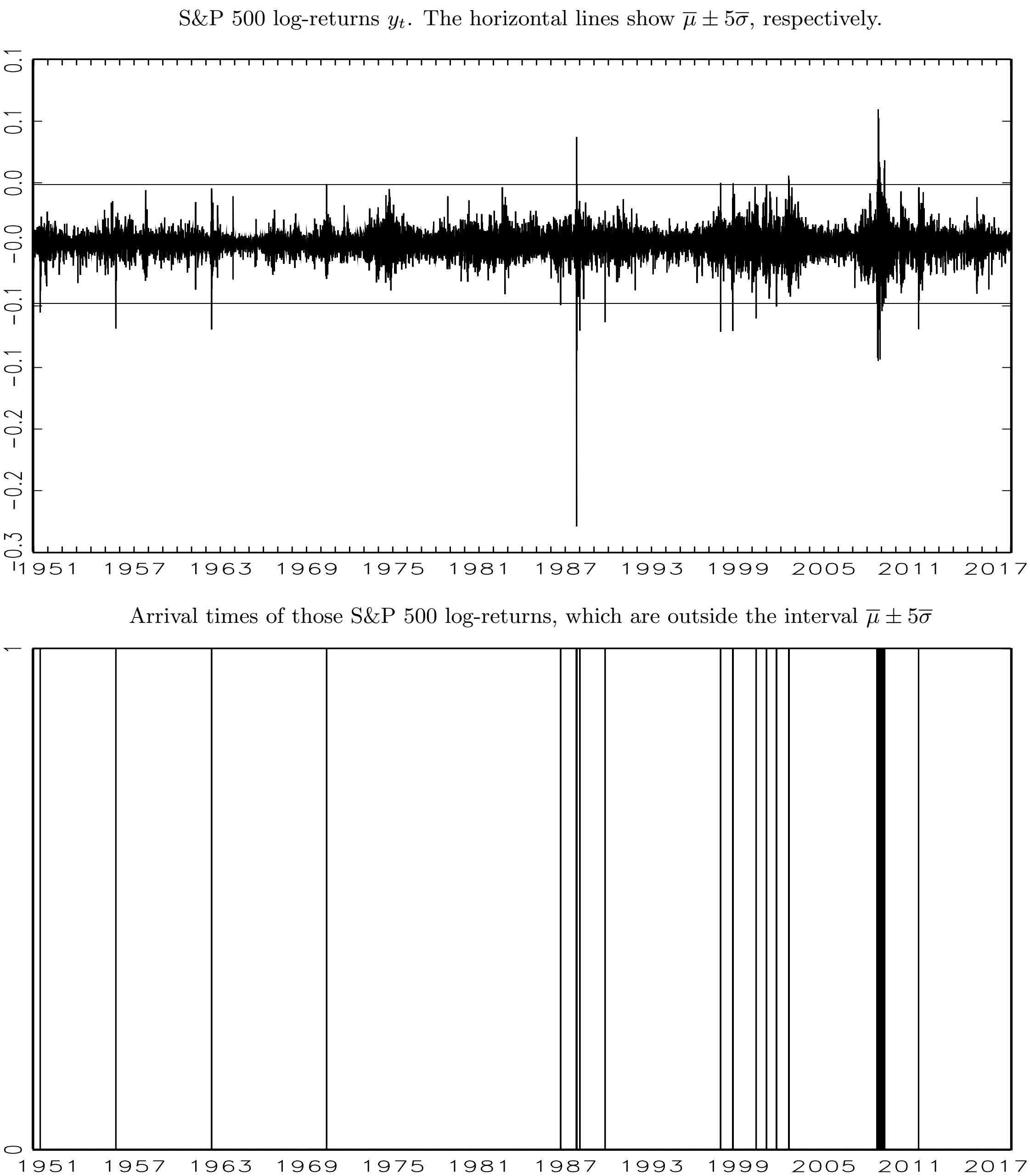

Fig. 3. S&P 500 log-returns y_t and outliers for the period of January 4, 1950 to December 30, 2017.
Notes: $\overline{\mu}$ and $\overline{\sigma}$ are the estimates of mean and standard deviation, respectively, of y_t. During the period of September 2, 2008 to March 31, 2009 (i.e. 146 trading days), there were 24 outliers

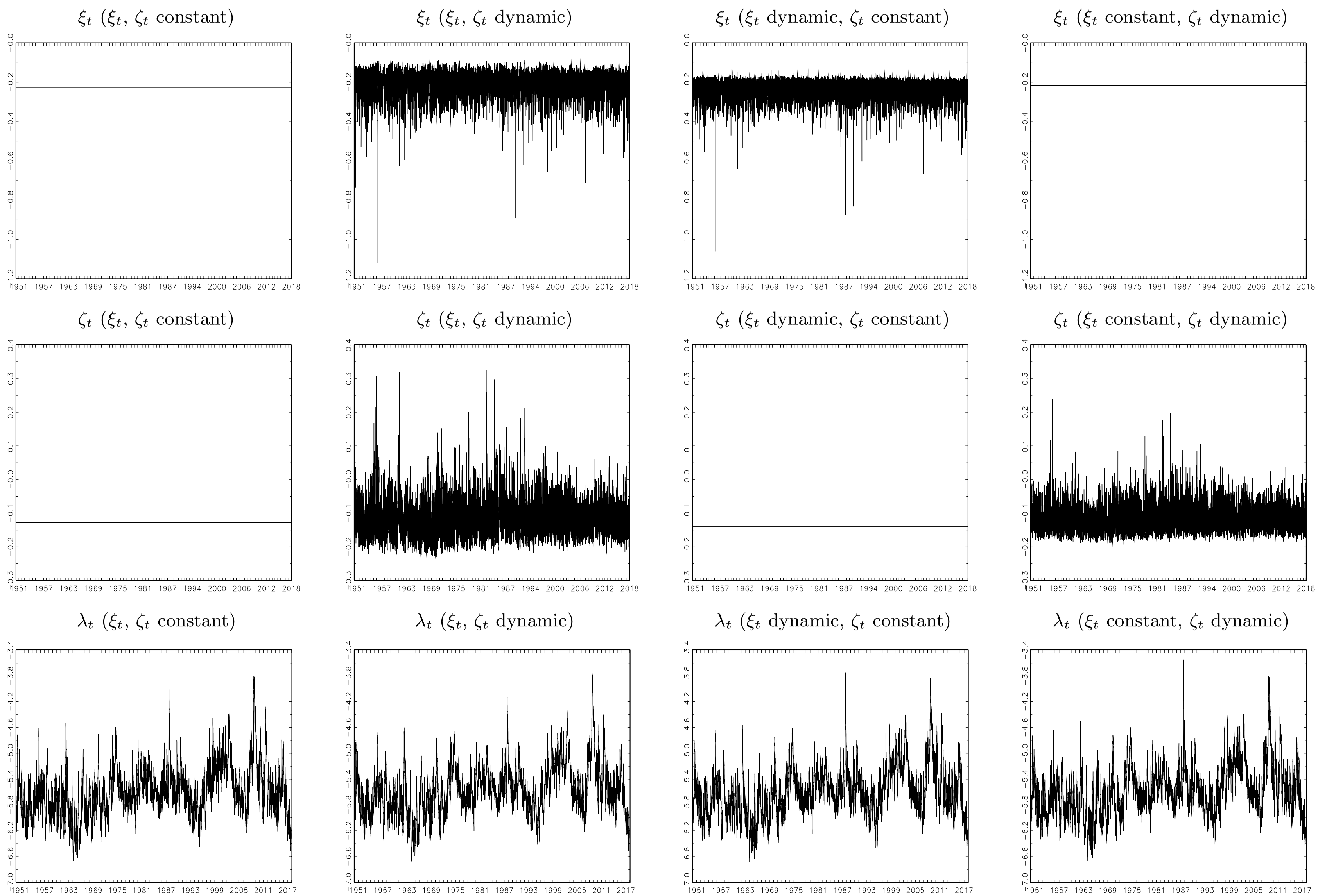

Fig. 4. EGB2-DCS model. *Note*: Each column shows a different shape parameter specification.

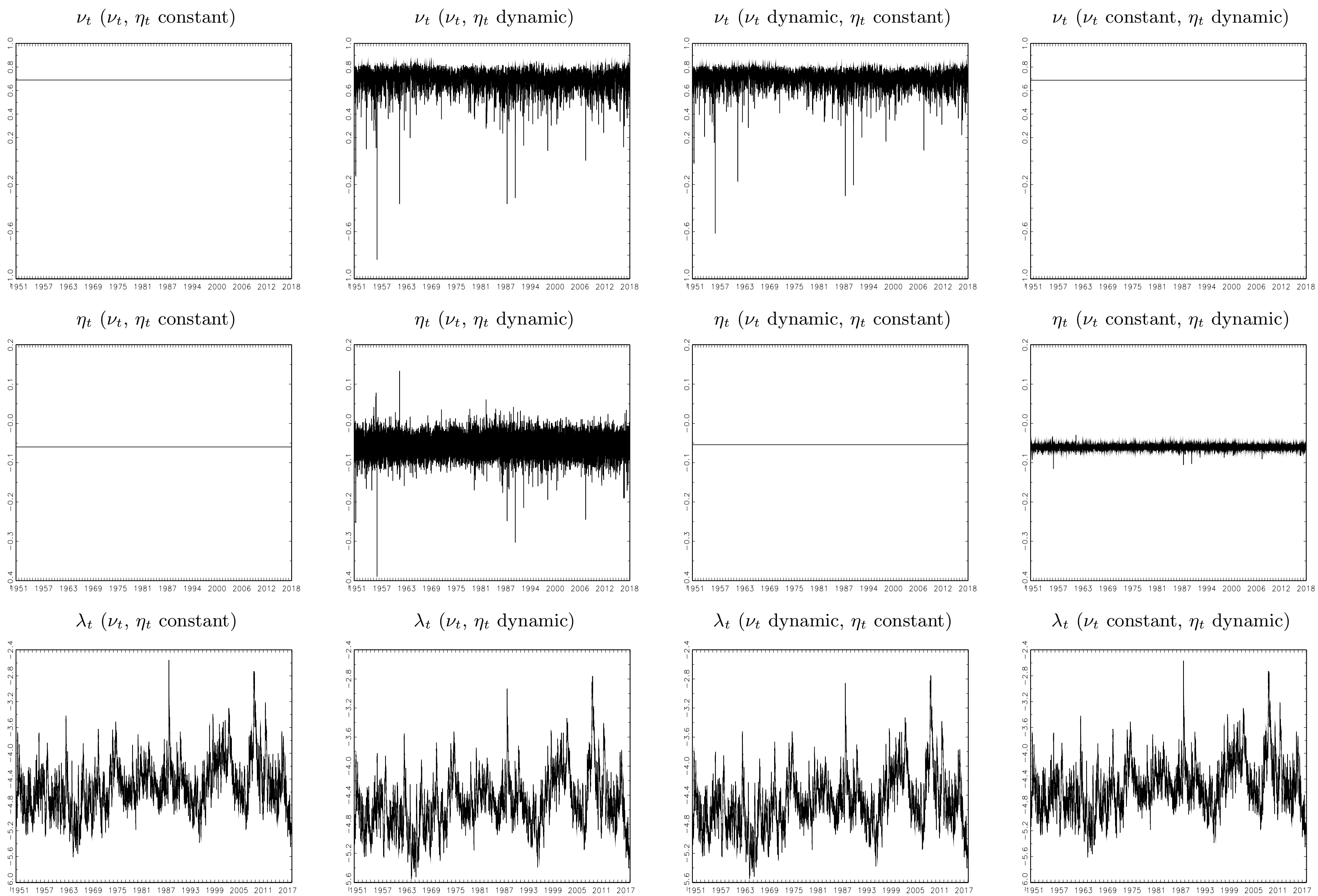

Fig. 5. NIG-DCS model. *Note*: Each column shows a different shape parameter specification.

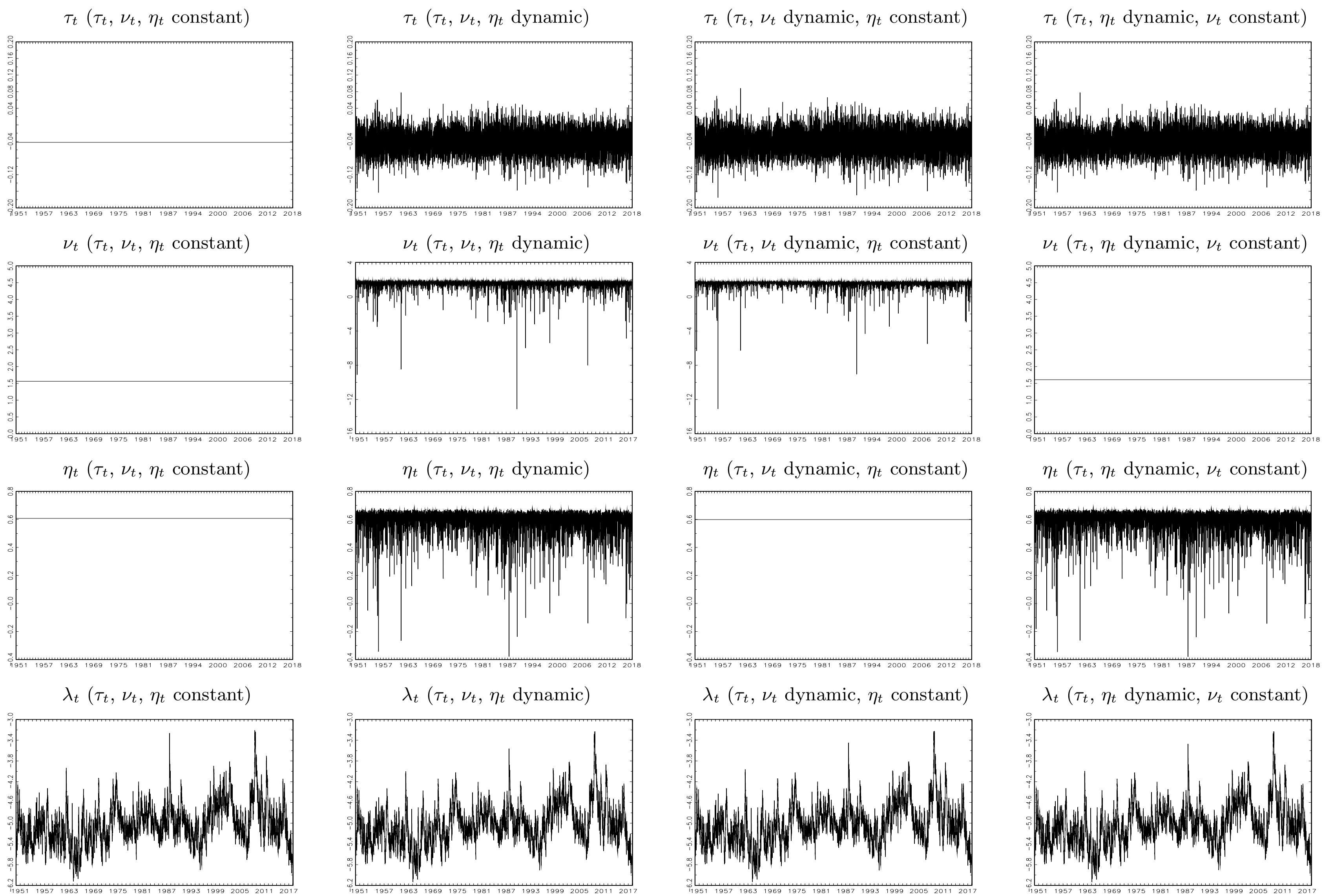

Fig. 6(a). Skew-Gen-t-DCS model. *Note*: Each column shows a different shape parameter specification.

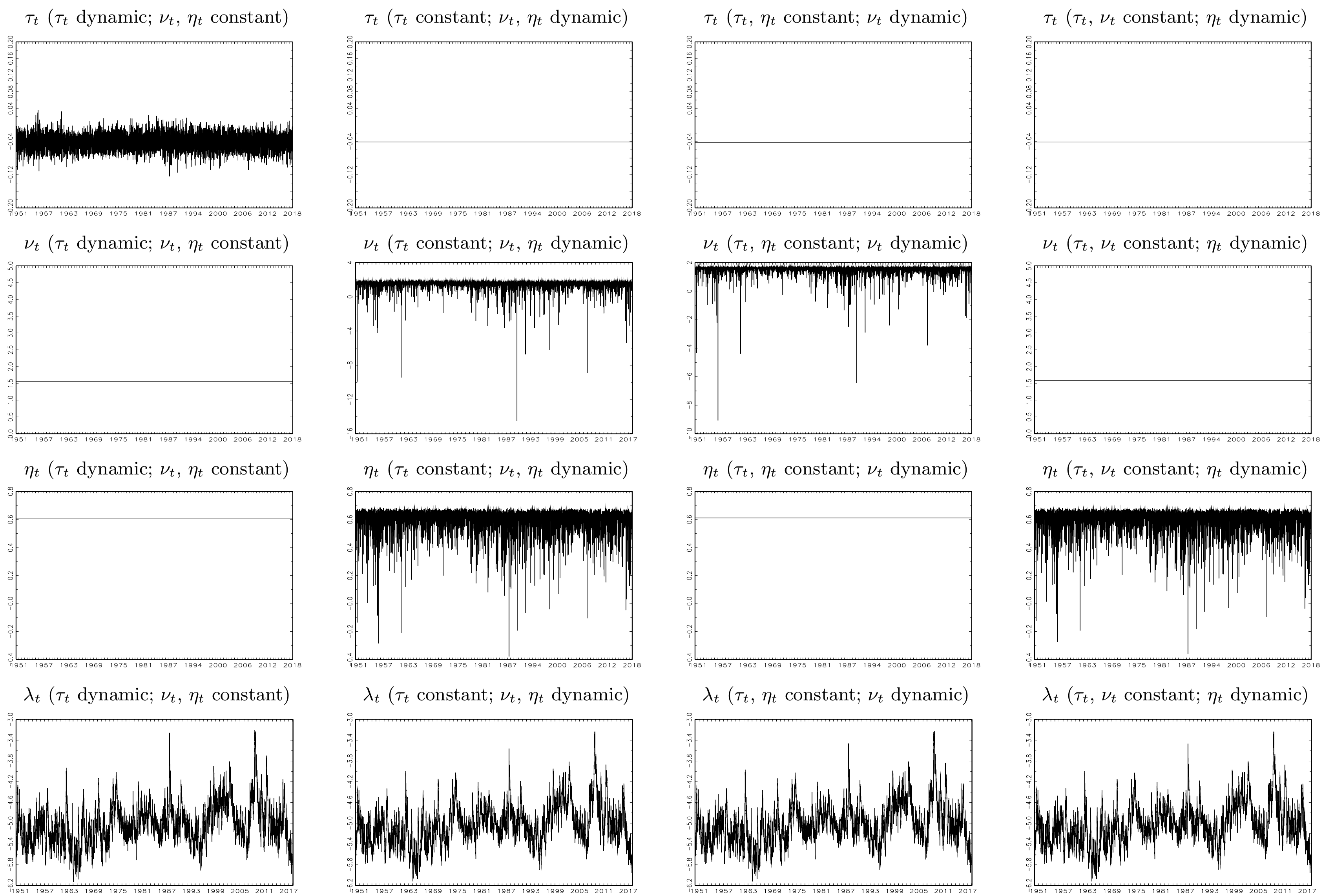

Fig. 6(b). Skew-Gen-t-DCS model. *Note*: Each column shows a different shape parameter specification.

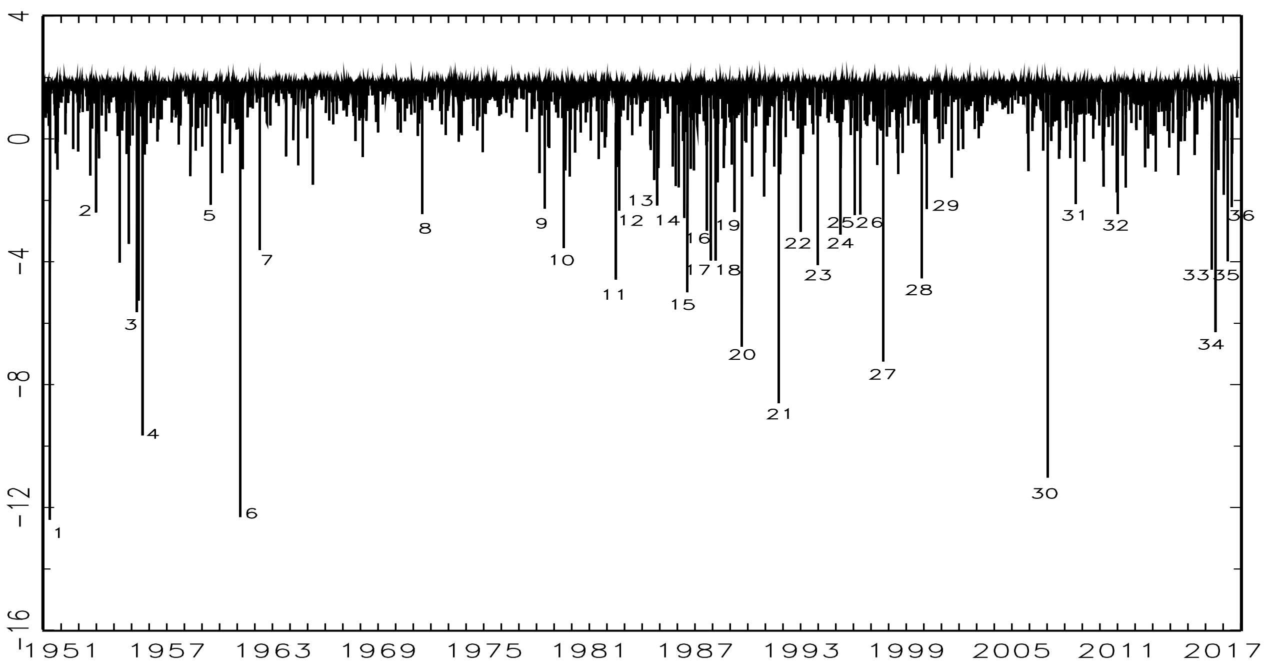

Fig. 7. Evolution of ν_t for the Skew-Gen-t-DCS model with constant τ_t, dynamic ν_t and constant η_t.
Note: The numbers indicate the extreme events from Appendix C.

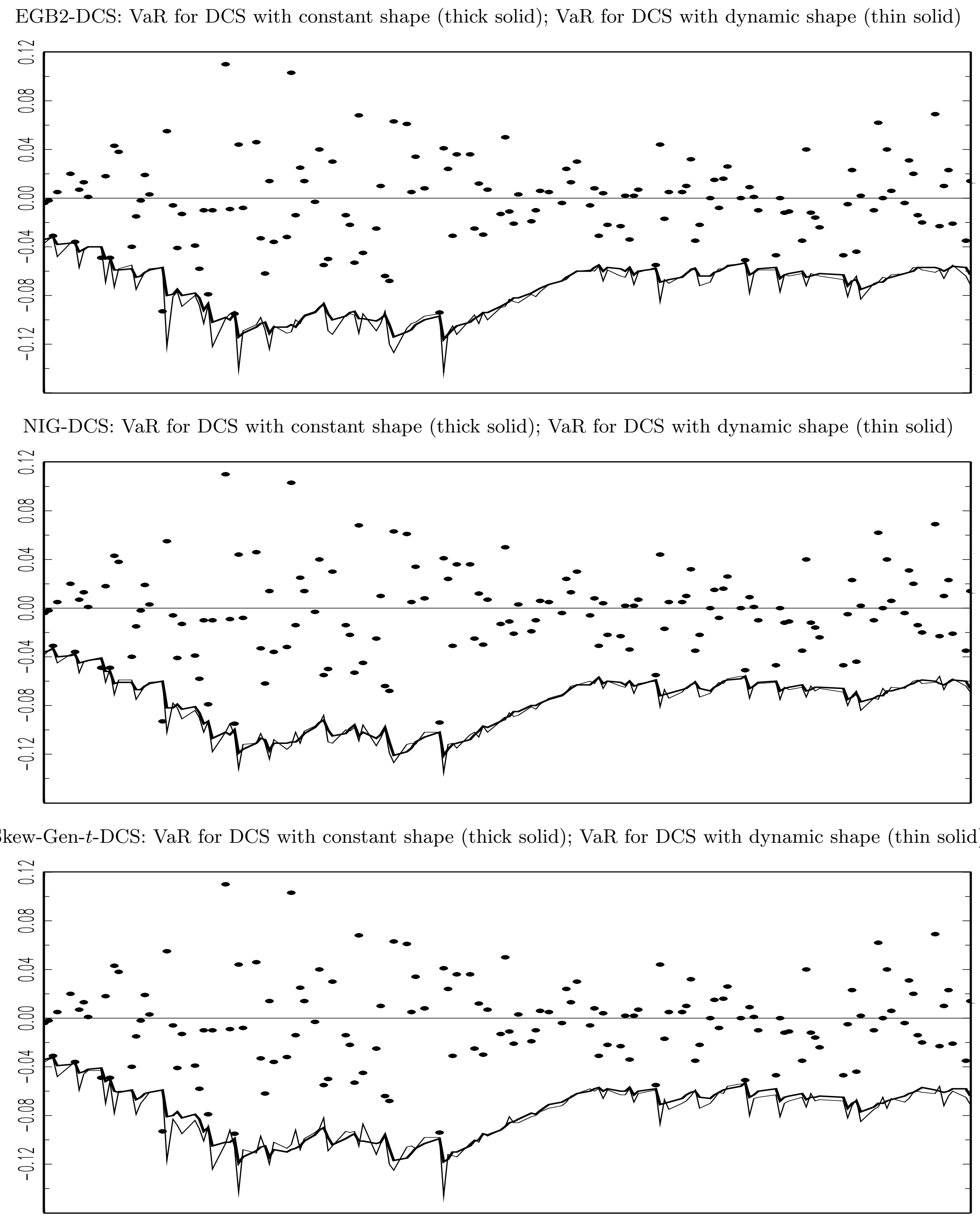

Fig. 8. Log-returns on the S&P 500 and VaR(1 day, 99%) for the period of September 2, 2008 to March 31, 2009. *Notes*: Solid circles indicate the log-returns on the S&P 500 for the backtesting period. For the dynamic-shape specifications of this figure, all of the shape parameters are dynamic.