Moments of a Family of Asymmetric Stochastic Volatility Models and the Stochastic News Impact Surface☆

Xiuping Maoa, Esther Ruizab,*, Helena Veigaab,c

aDepartment of Statistics, Universidad Carlos III de Madrid.
bInstituto Flores de Lemus, Universidad Carlos III de Madrid.
cFinancial Research Center/UNIDE, Avenida das Forças Armadas, 1600-083, Lisboa, Portugal.

Abstract

In this paper, we derive the statistical properties of a family of Stochastic Volatility (SV) models with leverage effect, which includes several popular models usually fitted to represent the dynamic evolution of asymmetric volatilities of financial returns. In spite of their popularity, the properties of some of these models were previously unknown in the literature. By comparing the properties of these models, we are able to establish the advantages and limitations of each of them and give some guidelines about which model to implement in practice. We also propose the Stochastic News Impact Surface (SNIS) to represent the asymmetric response of volatility to positive and negative shocks in the context of SV models. The SNIS is useful to show the added flexibility of SV models over GARCH models when representing conditionally heteroscedastic time series with leverage effect. Finally, we analyze the performance of a MCMC estimator of the parameters and volatilities of one model from this family and show that it has appropriate finite sample properties. Furthermore, estimating this model using this MCMC estimator, one can correctly identify the true nested specifications which are popularly implemented in empirical applications. All the results are illustrated by Monte Carlo experiments and by fitting the model to a series of daily S&P500 returns.

Keywords: EGARCH, Leverage effect, MCMC estimator, Threshold Stochastic Volatility, VaR, WinBUGS

☆Financial support from the Spanish Ministry of Education and Science, research projects ECO2009-08100 and ECO2012-32401, is acknowledged. The third author is also grateful for project MTM2010-17323. We thank Mike Wiper and the Bayesian study group of the Department of Statistics at UC3M for relevant suggestions about MCMC estimation.

September 14, 2014
1. Introduction

Asymmetric Stochastic Volatility (SV) models are often implemented to represent the asymmetric response of volatility to positive and negative past returns. Incorporating the leverage effect into SV models can have important implications from the point of view of financial models; see, for example, Hull and White (1987) in the context of the Black-Scholes formula, Nandi (1998) for pricing and hedging S&P500 index options and Lien (2005) for average optimal hedge ratios. A variety of alternative econometric specifications are available to choose among when dealing with SV models with leverage effect. In particular, Taylor (1994) and Harvey and Shephard (1996) propose incorporating the leverage effect through correlation between the level and log-volatility disturbances. Alternatively Demos (2002) and Asai and McAleer (2011) suggest adding a noise to the log-volatility equation specified as in the EGARCH model of Nelson (1991). Finally, Breidt (1996) and So et al. (2002) propose a Threshold SV model in which the parameters of the log-volatility equation change depending on whether past returns are positive or negative; see also Asai and McAleer (2006).

Although the asymmetric SV models mentioned above are often implemented to represent the dynamic dependence of volatilities, their statistical properties are either partially known or completely unknown. Consequently, it is not possible to establish their advantages and limitations when fitted to explain the empirical properties of financial returns. Furthermore, knowing the moments of returns implied by the different specifications can be important when estimating the parameters using estimators based on the Method of Moments (MM) as those proposed by, for example, Bollerslev and Zhou (2002) and Garcia et al. (2011). In this paper, we derive the statistical properties of a family of asymmetric SV models that we call generalized asymmetric SV (GASV). The GASV family is rather general including as particular cases some of the asymmetric SV models considered above. We focus on the ability of the GASV family to explain the empirical properties of interest when dealing with real financial time series, namely, excess kurtosis, positive and persistent autocorrelations of power-transformed absolute returns and negative cross-correlations between returns and
future power-transformed absolute returns. We propose one new specification named Threshold GASV (T-GASV) that belongs to the GASV family and nests some of the asymmetric SV models mentioned above. The analytical expressions of its statistical properties are obtained, therefore, as a marginal outcome of this analysis, we are able to point out the advantages and limitations of each of the restricted specifications.

A useful tool to describe how a particular model represents the asymmetric response of volatility to positive and negative past returns often observed in practice is the News Impact Curve (NIC) which was originally proposed by Engle and Ng (1993) in the context of GARCH models. Yu (2012) proposes an extension of the NIC to SV models based on measuring the effect of the level disturbance on the conditional variance. However, this is a rather difficult task due to the lack of observability of the volatility in SV models. In the spirit of Yu (2012), Takahashi et al. (2013) propose several methods to compute the news impact curve for SV models. In this paper, we suggest an alternative definition of the NIC in the context of SV models, which relates the volatility with the level and volatility disturbances. Therefore, we propose representing the response of volatility by a surface called Stochastic News Impact Surface (SNIS).\footnote{The SNIS proposed in this paper should not be confused with the News Impact Surface (NIS) defined in the context of multivariate models; see, for example, Asai and McAleer (2009), Savva (2009) and Caporin and McAleer (2011).} Analyzing the SNIS, we show that the asymmetric impact of the level disturbance on the volatility can be different depending on the volatility disturbance.

Although SV models are attractive for modeling volatility, their empirical implementation is limited by the difficulty involved in the estimation of their parameters which is complicated by the lack of a closed-form expression of the likelihood. Furthermore, the volatility itself is unobserved and cannot be directly estimated. Consequently, several simulation-based procedures have been proposed for the estimation of parameters and volatilities; see Broto and Ruiz (2004) for a survey. Among them, Monte Carlo Markov Chain (MCMC) based approaches have become popular given their good properties in estimating parameters and volatilities; see, for example, Omori et al. (2007), Omori and Watanabe (2008), Nakajima.
and Omori (2009), Abanto-Valle et al. (2010) and Tsiotas (2012) for MCMC estimators of SV models with leverage effect. In this paper, we consider a MCMC estimator implemented in the user-friendly and freely available WinBUGS software described by Meyer and Yu (2000). This estimator is based on a single-move Gibbs sampling algorithm and has been recently implemented in the context of asymmetric SV models by, for example, Yu (2012) and Wang et al. (2013). The MCMC estimator implemented by WinBUGS is appealing because it can handle non-Gaussian level disturbances without much programming effort. We carry out extensive Monte Carlo experiments and show that, it has adequate finite sample properties to estimate the parameters and volatilities of T-GASV model in situations similar to those encountered when analyzing time series of real financial returns. Furthermore, we show that its nested specifications can be adequately identified when the parameters are estimated using the WinBUGS software. Therefore, in empirical applications, researchers will be better off by fitting the general model proposed in this paper and letting the data choose the preferred specification of the volatility instead of choosing a particular ad hoc specification. Finally, the MCMC estimator is implemented to estimate the volatilities and Value at Risk (VaR) of daily S&P500 returns after fitting the new model proposed in this paper.

The rest of the paper is organized as follows. Section 2 defines the GASV family and derives its statistical properties. Section 3 proposes the SNIS to describe the asymmetric response of volatility. The shape of the SNIS is illustrated for several particular specifications of interest within the GASV family. The properties of the T-GASV and its nested specifications are analyzed and compared in Section 4. Section 5 conducts Monte Carlo experiments to analyze the finite sample properties of the MCMC estimator of the parameters and underlying volatilities of T-GASV model. Section 6 presents an empirical application to daily S&P500 returns. Finally, the main conclusions and some guidelines for future research are summarized in Section 7.
2. The statistical properties of the GASV family

In this section, we define the GASV family of asymmetric SV models and derive its statistical properties. In particular, we obtain the general conditions for stationarity and for the existence of integer moments of returns and absolute returns. Expressions of the marginal variance and kurtosis, the autocorrelations of power-transformed absolute returns and cross-correlations between returns and future power-transformed absolute returns are derived.

2.1. Model description

Let $y_t$ be the return at time $t$, $\sigma^2_t$ its volatility, $h_t \equiv \log \sigma^2_t$ and $\epsilon_t$ be an independent and identically distributed (IID) sequence with mean zero and variance one. The GASV family is given by

$$y_t = \exp(h_t/2) \epsilon_t, \quad t = 1, \ldots, T$$

(1)

$$h_t - \mu = \phi(h_{t-1} - \mu) + f(\epsilon_{t-1}) + \eta_{t-1},$$

(2)

where $f(\epsilon_{t-1})$ is any function of $\epsilon_{t-1}$ for which no restrictions are imposed further than being a function of $\epsilon_{t-1}$ but not of the other disturbance in the model, $\eta_{t-1}$. Therefore, given $\epsilon_t$, $f(\epsilon_t)$ is observable. The volatility noise, $\eta_t$, is a Gaussian white noise with variance $\sigma^2_\eta$. It is assumed to be independent of $\epsilon_t$ for all leads and lags. The scale parameter, $\mu$, is related with the marginal variance of returns, while $\phi$ is related with the rate of decay of the autocorrelations of power-transformed absolute returns towards zero and, consequently, with the persistence of the volatility shocks. Note that, in equations (1) and (2), the return at time $t-1$ is correlated with the volatility at time $t$. Furthermore, if $f(\cdot)$ is not an even function, then positive and negative past returns with the same magnitude have different effects on volatility.

It is important to note that although the specification of log-volatility in (2) is rather

---

$^2$The normality of $\eta_t$ when $f(\epsilon_{t-1}) = 0$ has been justified by, for example, Andersen et al. (2001a) and Andersen et al. (2001b, 2003).
general, it rules out models in which the persistence, \( \phi \), and/or the variance of the volatility noise, \( \sigma^2_\eta \), are time-varying. Finally, note that the only assumption made about the distribution of the level disturbance, \( \epsilon_t \), is that it is an IID sequence with mean zero and variance one. As a consequence, \( \epsilon_t \) is strictly stationary. In the related literature, different assumptions about this distribution have been considered. Originally, Jacquier et al. (1994) and Harvey and Shephard (1996) assume that \( \epsilon_t \) is a Gaussian process. Although this is the most popular assumption, there has been other proposals that consider heavy-tailed distributions such as the Student-t distribution or the Generalized Error Distribution (GED)\(^3\); see, for example, Chen et al. (2008), Choy et al. (2008) and Wang et al. (2011, 2013). Several authors also include skewness in the distribution of \( \epsilon_t \) by assuming an asymmetric GED distribution as in Cappuccio et al. (2004) and Tsiotas (2012) or a skew-Normal and a skew-Student-t distributions as in Nakajima and Omori (2012) and Tsiotas (2012).

2.2. Moments of returns

We now derive the statistical properties of the GASV family in equations (1) and (2). Theorem 2.1 establishes sufficient conditions for the stationarity of \( y_t \) and derives the expression of \( E(|y_t|^c) \) and \( E(y_t^c) \) for any positive integer \( c \).

**Theorem 2.1.** Define \( y_t \) by the GASV family in equations (1) and (2). The process \( \{y_t\} \) is strictly stationary if \(|\phi| < 1\). Further, if \( \epsilon_t \) follows a distribution such that both \( E(\exp(0.5c f(\epsilon_t))) \) and \( E(|\epsilon_t|^c) \) exist and are finite for some positive integer \( c \), then \( \{|y_t|\} \) and \( \{y_t\} \) have finite, time-invariant moments of arbitrary order \( c \) which are given by

\[
E(|y_t|^c) = \exp\left(\frac{c\mu}{2}\right) E(|\epsilon_t|^c) \exp\left(\frac{c^2 \sigma^2_\eta}{8(1 - \phi^2)}\right) P(0.5c\phi^{i-1})
\]  

---

\(^3\)The GED distribution with parameter \( \nu \) is described by Harvey (1990) and has the attractiveness of including distributions with different tail thickness as, for example, the Normal when \( \nu = 2 \), the Double Exponential when \( \nu = 1 \) and the Uniform when \( \nu = \infty \). The GED distribution has heavy tails if \( \nu < 2 \).
and

\[ E(y_t^c) = \exp\left(\frac{c\mu}{2}\right) E(\epsilon_t^c) \exp\left(-\frac{c^2 \sigma^2}{8(1 - \phi^2)}\right) P(0.5c\phi^{i-1}), \quad (4) \]

where \( P(b_i) \equiv \prod_{i=1}^{\infty} E(\exp(b_i f(\epsilon_{t-i}))). \)

**Proof.** See Appendix A.1. \[\Box\]

Theorem 2.1 establishes the strict stationarity of \( y_t \) if \(|\phi| < 1\) and the existence of the expectation of \( y_t^2 \) if further \( E(\exp(f(\epsilon_t))) < \infty \). Consequently, under these two conditions, \( y_t \) is also weakly stationary.

Note that according to expression (4), if \( \epsilon_t \) has a symmetric distribution, then all odd moments of \( y_t \) are zero. Furthermore, from expression (3), it is straightforward to obtain expressions of the marginal variance and kurtosis of \( y_t \) as the following corollaries show.

**Corollary 2.1.** Under the conditions of Theorem 2.1 with \( c = 2 \) and taking into account that \( E(y_t) = 0 \), the marginal variance of \( y_t \) is directly obtained from (3) as follows

\[ \sigma_y^2 = \exp\left(\mu + \frac{\sigma^2}{2(1 - \phi^2)}\right) P(\phi^{i-1}). \quad (5) \]

**Corollary 2.2.** Under the conditions of Theorem 2.1 with \( c = 4 \), the kurtosis of \( y_t \) can be obtained as \( E(y_t^4)/(E(y_t^2))^2 \) using expression (3) with \( c = 4 \) and \( c = 2 \) as follows

\[ \kappa_y = \kappa_\epsilon \exp\left(\frac{\sigma^2}{1 - \phi^2}\right) \frac{P(2\phi^{i-1})}{[P(\phi^{i-1})]^2}, \quad (6) \]

where \( \kappa_\epsilon \) is the kurtosis of \( \epsilon_t \).

The kurtosis of the basic symmetric Autoregressive SV (ARSV) model considered by Harvey et al. (1994) is given by \( \kappa_\epsilon \exp\left(\frac{\sigma^2}{1 - \phi^2}\right). \) Therefore, this kurtosis is multiplied by the factor \( r = \frac{P(2\phi^{i-1})}{[P(\phi^{i-1})]^2} \) in the GASV family.

Note that, the expression of \( E(|y_t|^c) \) in (3) depends on \( f(\cdot) \) and on the distribution of
Therefore, in order to obtain closed-form expressions of the variance and kurtosis of returns, one needs to assume a particular distribution of $\epsilon_t$ and a specification of $f(\epsilon_t)$. We will particularize these expressions for some popular distributions and specifications in Section 4. Also, it is important to note that even for those cases in which the function $f(\cdot)$ and/or the distribution of $\epsilon_t$ are such that they do not allow to obtain closed-form expressions of the moments, expression (3) can always be used to simulate them as far as they are finite.

2.3. Dynamic dependence

Looking at the dynamic dependence of returns when they are defined as in (1) and (2), it is easy to see that they are a martingale difference. However, they are not serially independent as the conditional heteroscedasticity generates non-zero autocorrelations of power-transformed absolute returns. The following theorem derives the autocorrelation function (acf) of power transformed absolute returns.

**Theorem 2.2.** Consider a stationary process $y_t$ defined by equations (1) and (2) with $|\phi| < 1$. If $\epsilon_t$ follows a distribution such that $E(\exp(0.5 cf(\epsilon_t))) < \infty$ and $E(|\epsilon_t|^c) < \infty$ for some positive integer $c$, then the $\tau$-th order autocorrelation of $|y_t|^c$ is finite and given by

$$
\rho_c(\tau) = \frac{\left\{ E(|\epsilon_t|^c) E(|\epsilon_t|^c \exp(0.5c\phi^{-1}f(\epsilon_t))) \exp \left( \frac{\phi^c \sigma^2 \gamma^2}{4(1 - \phi^2)} \right) P(0.5c(1 + \phi^c)\phi^{i-1}) \right\} T(\tau, 0.5c\phi^{i-1})}{\left\{ E(|\epsilon_t|^{2c}) \exp \left( \frac{c^2 \sigma^2 \gamma^2}{4(1 - \phi^2)} \right) P(c\phi^{i-1}) - [E(|\epsilon_t|^c) P(0.5c\phi^{i-1})]^2 \right\}},
$$

where $T(n, b_i) \equiv \prod_{i=1}^{n-1} E(\exp(b_i f(\epsilon_t-i)))$ if $n > 1$ while $T(1, b_i) \equiv 1$.

**Proof.** See Appendix A.2.

Notice that, in practice, most authors dealing with real time series of financial returns
focus on the autocorrelations of squared and absolute returns, $\rho_2(\tau)$ and $\rho_1(\tau)$, which can be obtained from (7) when $c = 2$ and $c = 1$, respectively.

The leverage effect is reflected in the cross-correlations between power-transformed absolute returns and lagged returns. The following theorem gives general expressions of these cross-correlations.

**Theorem 2.3.** Consider a stationary process $y_t$ defined by equations (1) and (2) with $|\phi| < 1$. If $\epsilon_t$ follows a distribution such that $E(\exp(0.5cf(\epsilon_t))) < \infty$ and $E(|\epsilon_t|^{2c}) < \infty$ for some positive integer $c$, then the $\tau$-th order cross-correlation between $y_t$ and $|y_{t+\tau}|^c$ for $\tau > 0$ is finite and given by

$$
\rho_{c1}(\tau) = \frac{E(|\epsilon_t|^c) \exp\left(\frac{2c\phi^\tau - 1}{8(1 - \phi^2)} \sigma_\eta^2\right) E(\epsilon_t \exp(0.5c\phi^{\tau - 1} f(\epsilon_t))) P(0.5 + c\phi^\tau) \phi^{i-1})}{T(\tau, 0.5c\phi^{i-1}) \sqrt{P(c\phi^{i-1})}} \sqrt{E(|\epsilon_t|^{2c}) \exp\left(\frac{c^2\sigma_\eta^2}{4(1 - \phi^2)}\right) P(c\phi^{i-1}) - [E(|\epsilon_t|^c) P(0.5c\phi^{i-1})]^2}.
$$

(8)

*Proof.* See Appendix A.3.

3. The Stochastic News Impact Surface

Besides the cross-correlations between returns and future power-transformed absolute returns, another useful tool to describe the asymmetric response of volatility is the News Impact Curve (NIC) originally proposed by Engle and Ng (1993) in the context of GARCH models. The NIC is defined as the function relating past return shocks to current volatility with all lagged conditional variances evaluated at the unconditional variance of returns. The NIC has been widely implemented when dealing with GARCH-type models; see, for example, Maheu and McCurdy (2004). Extending the NIC to SV models is not straightforward due to the presence of the volatility disturbance in the latter models. As far as we know, there are two attempts in the literature to propose a NIC function for SV models. The first is attributed to Yu (2012) who proposes a function that relates the conditional variance to the
lagged return innovation, $\epsilon_{t-1}$, holding all other lagged returns equal to zero. Given that, in SV models the conditional variance is not directly specified, this definition of the NIC requires solving high-dimensional integrals using numerical methods making its computation a difficult task. Furthermore, the NIC proposed by Yu (2012) is based on integrating over the latent volatilities and, therefore, useful information about the differences between the effects of $\epsilon_t$ on $\sigma_{t+1}$ for different values of $\eta_t$ can be lost. The second attempt is due to Takahashi et al. (2013) that specifies the news impact function for SV models in the spirit of Yu (2012) as the volatility at time $t + 1$ conditional on returns at time $t$. However, in order to obtain an U-shaped NIC, Takahashi et al. (2013) proposes to incorporate the dependence between returns and volatility by considering their joint distribution. This idea is implemented by using a Bayesian MCMC scheme or a simple rejection sampling.

It is important to note that, in the context of GARCH models, because there is just one disturbance, the volatility at time $t$, $\sigma_t^2$, coincides with the conditional variance, $\text{Var}(y_t|y_1, \ldots, y_{t-1})$. Consequently, when Engle and Ng (1993) propose relating past returns to current volatility, this amounts to relating past returns with conditional variances. However, in SV models, the volatility and the conditional variance are different objects. Therefore, in this paper, we propose measuring the effect of past shocks, $\epsilon_{t-1}$ and $\eta_{t-1}$, on the volatility instead of on the conditional variance as proposed by Yu (2012). Taking into account the information provided by the two disturbances involved in the model, we define the Stochastic News Impact Surface (SNIS) as the surface that relates $\sigma_t^2$ with $\epsilon_{t-1}$ and $\eta_{t-1}$. As in Engle and Ng (1993), we evaluate the lagged volatilities at the marginal variance, so that, we consider that at time $t - 1$, the volatility is equal to an “average” volatility and analyze the effect of level shocks, $\epsilon_{t-1}$, and volatility shocks, $\eta_{t-1}$, on the volatility at time $t$. Therefore, the SNIS is given by

$$\text{SNIS}_t = \exp((1 - \phi)\mu)\sigma_y^{2\phi} \exp(f(\epsilon_{t-1}) + \eta_{t-1}).$$

Note that the shape of SNIS does not depend on the distribution of $\epsilon_t$ as it is a function of $f(\epsilon_{t-1})$ and $\eta_{t-1}$.
For illustrating the SNIS, we consider the following specification of $f(\cdot)$

$$f(\epsilon_t) = \alpha I(\epsilon_t < 0) + \gamma_1 \epsilon_t + \gamma_2 |\epsilon_t|,$$  

(10)

where $I(\cdot)$ is an indicator function that takes value one when the argument is true and zero otherwise. We denote the model defined by equations (1), (2) and (10) as Threshold GASV (T-GASV). This specification is interesting because it nests several popular models previously proposed in the literature to represent asymmetric volatilities in the context of SV models. For example, when $\alpha = \gamma_2 = 0$ and $\epsilon_t$ follows a Gaussian distribution, we obtain the asymmetric autoregressive SV (A-ARSV) model of Harvey and Shephard (1996). On the other hand, when $\alpha = 0$ the model reduces to the EGARCH plus error model of Demos (2002) and Asai and McAleer (2011), denoted as E-SV. Finally, when only $\alpha \neq 0$, equation (10) resumes to a threshold model where only the constant changes depending on the sign of past returns. By changing the threshold in the indicator variable, we allow the leverage effect to be different depending on the size of $\epsilon_t$.

The top panel of Figure 1 plots the SNIS of the T-GASV model with \{\phi, \sigma^2_\eta, \alpha, \gamma_1, \gamma_2\} = \{0.98, 0.05, 0.07, -0.08, 0.1\} and $\exp((1 - \phi)\mu)\sigma^2_\phi = 1$. We can observe that the SNIS shows a discontinuity due to the presence of the indicator function in (10). The leverage effect is very clear when the volatility shock is positive. The most important feature of the SNIS plotted in Figure 1 is that it shows that the leverage effect of SV models is different depending on the values of the volatility shock. In practice, when $\eta_{t-1}$ is negative, the leverage effect is weaker. When $\eta_{t-1} = 0$, we obtain the NIC of the corresponding GARCH-type model which is also plotted in Figure 1. It is important to observe that by introducing $\eta_t$ in the T-GASV model, more flexibility is added to represent the leverage effect.

Next, we illustrate the shape of SNIS of the E-SV model. For this purpose, we consider the same parameters as above with $\alpha = 0$ and plot the corresponding SNIS in the second panel of Figure 1. In this case, we can observe that there is not any discontinuity but the
effect of $\epsilon_{t-1}$ on $\sigma_t$ still depends on $\eta_{t-1}$. As before, we also plot the NIC of the EGARCH model of Nelson (1991) by considering $\eta_t = 0$.

Finally, in the bottom panel of Figure 1, we illustrate the shape of the SNIS of the A-ARSV model of Harvey and Shephard (1996) by considering further that $\gamma_2 = 0$. In this case, we can observe that the leverage effect is very weak for negative log-volatility shocks.

Summarizing, Figure 1 shows that, for the T-GASV model and the particular parameter values considered, given a value of the lagged volatility shock, $\eta_{t-1}$, the response of volatility is stronger when $\epsilon_{t-1}$ is negative than when it is positive with the same magnitude. Furthermore, this asymmetric response depends on the log-volatility noise, $\eta_{t-1}$. The leverage effect is clearly stronger when $\eta_{t-1}$ is positive and large than when it is negative.

4. Alternative Asymmetric SV models

As mentioned above, appropriate choices of the function $f(\cdot)$ and of the distribution of $\epsilon_t$ allow obtaining closed-form expressions of the moments of returns. In this section, we derive these expressions for the T-GASV model when $\epsilon_t$ follows a GED distribution. Furthermore, we also consider some of the most popular specifications nested in the T-GASV model and compare their properties to see which are closer to the empirical properties observed in real time series of financial returns.

4.1. T-GASV model

Consider the T-GASV model defined in equations (1), (2) and (10) with $\epsilon_t \sim GED(\nu)$. If $\nu > 1$, then the conditions in Theorem 2.1 are satisfied and a closed-form expression of $E(|y_t|^c)$ can be derived; see Appendix B.1 for the corresponding expectations. In particular, the marginal variance of $y_t$ is given by equation (5) with

$$P(b_i) = \prod_{i=1}^{\infty} \left\{ \sum_{k=0}^{\infty} \left( \frac{\Gamma(1/\nu)}{\Gamma(3/\nu)} \right) \frac{\Gamma((k+1)/\nu)}{2\Gamma(1/\nu)k!} b_i^k \left[ (\gamma_1 + \gamma_2)^k + \exp(\alpha b_i)(\gamma_2 - \gamma_1)^k \right] \right\}, \quad (11)$$
where $\Gamma(\cdot)$ is the Gamma function. Note that in order to compute $P(\cdot)$, one needs to truncate the corresponding infinite product and summation. Our experience is that truncating the product at $i = 500$ and the summation at $k = 1000$ gives very stable results. Similarly, the kurtosis can be obtained as given in expression (6) with $P(\phi^{i-1})$ and $P(2\phi^{i-1})$ given as in expression (11).

Given that the Gaussian distribution is a special case of the GED distribution when $\nu = 2$, closed-form expressions of $E(|y_t|^c)$ can also be obtained in this case; see Appendix B.2 for the corresponding expectations. In particular, the marginal variance is given by expression (5) while the kurtosis is given by expression (6) with

$$P(b_i) = \prod_{i=1}^{\infty} \left\{ \exp \left( \alpha b_i + \frac{b_i^2(\gamma_1 - \gamma_2)^2}{2} \right) \Phi(b_i(\gamma_2 - \gamma_1)) + \exp \left( \frac{b_i^2(\gamma_1 + \gamma_2)^2}{2} \right) \Phi(b_i(\gamma_2 + \gamma_1)) \right\},$$

where $\Phi(\cdot)$ is the Normal cumulative distribution function.

When $\nu < 1$, we cannot obtain analytical expressions of $E(|y_t|^c)$. However, in Appendix B.1, we show that $E(|y_t|^c)$ in equation (3) is finite if $\gamma_2 + \gamma_1 \leq 0$ and $\gamma_2 - \gamma_1 \leq 0$. Finally, if $\nu = 1$, the conditions for the existence of $E(|y_t|^c)$ in equation (3) are $\gamma_2 + \gamma_1 < 2\sqrt{2}/c$ and $\gamma_2 - \gamma_1 < 2\sqrt{2}/c$.

As mentioned in the previous section, the kurtosis of the T-GASV model is equal to the kurtosis of the basic symmetric ARSV model multiplied by the factor, $r = \frac{P(2\phi^{i-1})}{P(\phi^{i-1})^2}$. We illustrate its shape in Figure 2 which plots it as a function of the leverage parameters $\alpha$ and $\gamma_1$ when $\gamma_2 = 0.1$ and 0 for three different persistence parameters, namely, $\phi = 0.5, 0.9$ and 0.98 assuming Gaussian errors. First of all, we can observe that the factor is always larger than 1. Therefore, the T-GASV generates returns with higher kurtosis than the corresponding basic symmetric ARSV model. Second, the effects of the parameters $\alpha$, $\gamma_1$

\footnote{The same conditions should be satisfied for the finiteness of $E(|y_t|^c)$ when $\epsilon_t$ follows a Student-t distribution with $d > 2$ degrees of freedom.}
and \( \gamma_2 \) on the kurtosis of returns are very different depending on the persistence. The kurtosis increases with \( \alpha \), \( |\gamma_1| \) and \( \gamma_2 \). However, their effects are only appreciable when \( \phi \) is close to 1.

The expectations needed to obtain closed-form expressions of the autocorrelations in expression (7) and cross-correlations in (8) have been derived in Appendix B.1 for the T-GASV model with parameter \( \nu > 1 \) and in Appendix B.2 for the particular case of the Normal distribution, i.e. \( \nu = 2 \). As above, when \( \nu \leq 1 \), we can only obtain conditions for the existence of the autocorrelations and cross-correlations. As these autocorrelations are highly non-linear functions of the parameters, it is not straightforward to analyze the role of each parameter on their shape. Furthermore, by comparing the autocorrelations in (7) for absolute and squared returns, it is not easy to conclude whether the T-GASV model is able to generate the Taylor effect defined by the autocorrelations of absolute returns being larger than those of squares; see Ruiz and Pérez (2012) for an analysis of the Taylor effect in the context of symmetric SV models. Consequently, in order to illustrate how these moments depend on each of the parameters, we focus on the model with parameters \( \phi = 0.98 \), \( \sigma^2 = 0.05 \) and Gaussian errors.

The first order autocorrelations of squared and absolute returns, namely, \( \rho_2(1) \) and \( \rho_1(1) \), are plotted in the first row of Figure 3 as functions of the leverage parameters, \( \gamma_1 \) and \( \alpha \). In the top left panel of Figure 3, which corresponds to the autocorrelations of squares, we can observe that they are larger, the larger is \( \gamma_2 \). However, both surfaces are rather flat and, consequently, the leverage parameters do not have large effects on the first order autocorrelations of squares. The corresponding first order autocorrelations of absolute returns are plotted in the top right panel of Figure 3. The autocorrelations of absolute returns are also larger the larger is the parameter \( \gamma_2 \). However, we can observe that the autocorrelations of absolute returns increase with the threshold parameter \( \alpha \). The effect of \( \gamma_1 \) on the autocorrelations of absolute returns is much milder. Finally, comparing \( \rho_1(1) \) with \( \rho_2(1) \), we can conclude that, the Taylor effect is stronger the larger is the leverage effect, regardless of whether this is due to \( \alpha \) or \( \gamma_1 \).
In the second row of Figure 3, we illustrate the effect of the parameters on the cross-correlations between $y_t$ and $y_{t+1}^2$ and $|y_{t+1}|$, $\rho_{21}(1)$ and $\rho_{11}(1)$, respectively. First of all, observe that the first order cross-correlations between returns and future absolute and squared returns are indistinguishable for the two values of $\gamma_2$ considered in Figure 3. Second, for a given value of $\gamma_2$, it is obvious that increasing the leverage parameters $\alpha$ and $|\gamma_1|$ increases the absolute cross-correlations. Note that $|\gamma_1|$ drags $\rho_{21}(1)$ in an approximately linear way while the effect of $\alpha$ is non-linear. On the other hand, the absolute cross-correlations between returns and future absolute returns have an approximately linear relationship with $\gamma_1$ and $\alpha$ and are clearly larger than those between returns and future squared returns. Therefore, it seems that when identifying conditional heteroscedasticity and leverage effect in practice, it is preferable to work with absolute returns instead of squared returns.

Figure 3 focuses on the first order autocorrelations and cross-correlations, but gives no information on the shape of the acf and the cross-correlation function (ccf) for different lags. To illustrate these shapes and the role of the distribution of $\epsilon_t$ on the acf and ccf, the first row of Figure 4 plots the acf of squared and absolute returns and the ccf between returns and future squared and absolute returns for the same specification of $f(\cdot)$ considered above with parameters $\alpha = 0.07$, $\gamma_2 = 0.1$, $\gamma_1 = -0.08$ and four different values of the GED parameter, $\nu = 1.5, 1.7, 2$ and $2.5$. As expected, the acfs of $|y_t|$ and $y_t^2$ in the first two panels have an exponential decay. Furthermore, fatter tails of $\epsilon_t$ imply smaller autocorrelations of both absolute and squared returns; see Carnero et al. (2004) for similar conclusions in the context of symmetric SV models. The ccf plotted in the last two panels show that the parameter $\nu$ of the GED distribution has a very mild influence on the cross-correlations, especially for $\rho_{11}(\tau)$.

To put it briefly, both $\nu$ and $\gamma_2$ increase the flexibility of the model to represent the volatility clustering while have little influence on the leverage effect. On the other hand, $\gamma_1$ affects the leverage effect and this effect is reinforced by the inclusion of $\alpha$, which could influence slightly the autocorrelations of absolute returns.
4.2. Popular specifications of the leverage effect

One of the most popular specifications of SV models with leverage effect is the Gaussian A-ARSV model of Taylor (1994) and Harvey and Shephard (1996) which is obtained from T-GASV model when \( \epsilon_t \) is Gaussian and \( \alpha = \gamma_2 = 0 \); see Bartolucci and De Luca (2003), Yu et al. (2006) and Tsiotas (2012)\(^5\) among many others for empirical applications.

The moments of the Gaussian A-ARSV model have been already derived in the literature by Taylor (1994, 2007), Demos (2002), Ruiz and Veiga (2008) and Pérez et al. (2009). By using the expressions in the previous subsection, we can also obtain closed-form expressions of the moments when the errors are GED. As an illustration, Figure 4 plots the acfs and ccfs of the A-ARSV model for the same parameter values of the T-GASV model represented in the first row of Figure 4 except that \( \alpha = \gamma_2 = 0 \). We can observe that the autocorrelations of squared and absolute returns and the absolute cross-correlations are slightly smaller than those of the corresponding T-GASV models. Therefore, including \( \gamma_2 \) and \( \alpha \) in the T-GASV model allows for stronger volatility clustering and leverage effect. Smaller autocorrelations are observed when the tails of the distribution of the return disturbance, \( \epsilon_t \), are fatter. Once more, the thickness of the tails has very mild influence on the cross-correlations and, therefore, on the leverage effect.

The E-SV model of Demos (2002) and Asai and McAleer (2011) can also be obtained by assuming Normality of \( \epsilon_t \) and \( f(\epsilon_t) \) being defined as in (10) with \( \alpha = 0 \).\(^6\) The parameter \( \gamma_2 \) measures the dependence of \( h_t \) on past absolute return disturbances in the same form as in the EGARCH model.\(^7\) It nests the A-ARSV model when \( \gamma_2 = 0 \). Demos (2002) derives the acf of \( y_t \) and the ccf between \( y_t \) and \( y_t^2 \); see also Asai and McAleer (2011).\(^8\)

\(^5\)Tsiotas (2012) allows the return disturbance to follow several asymmetric and fat-tailed distributions.
\(^6\)Asai and McAleer (2011) also consider E-SV model with Student-t return errors.
\(^7\)In independent work, Asai et al. (2012) mention a specification of the volatility similar to the T-GASV model with \( f(\epsilon_t) \) defined as in (10) with long-memory. However, they do not develop further the statistical properties of the model.
\(^8\)It is important to point out that the E-SV model has also been implemented by specifying the log-volatility using \( y_{t-1} \) instead of \( \epsilon_{t-1} \) in the volatility equation; see Danielsson (1998) and Asai and McAleer (2005). In this case, although the estimation of the parameters is usually easier, the derivation of the properties is harder.
Using the results in this paper, we can obtain the properties when the errors have a GED distribution. The third row of Figure 4 plots the autocorrelations and cross-correlations for an E-SV model with the same parameter values of the T-GASV model considered above except that \( \alpha = 0 \). Comparing the plots of the A-ARSV and E-SV models in Figure 4, we can observe that adding \( |\epsilon_{t-1}| \) into the A-ARSV model generates larger autocorrelations of squares and absolute returns but not larger Taylor effect. However, as expected, the cross-correlations are almost identical. Therefore, the E-SV model is more flexible than the A-ARSV to represent wider patterns of volatility clustering but not of volatility leverage.

Figure 4 also illustrates that the E-SV model is not identified by the autocorrelations of squared and absolute returns and the cross-correlations between returns and future squared and absolute returns, when the parameter of the GED distribution of \( \epsilon_t, \nu \), is not fixed. Observe that, given a particular E-SV model, we may find an A-ARSV model with almost the same autocorrelations and cross-correlations. Compare, for example, the autocorrelations of the E-SV model with \( \nu = 2 \) and those of the A-ARSV model with \( \nu = 2.5 \). Further, the cross-correlations are indistinguishable in any case. However, these two models generate returns with different kurtoses. Therefore, if the parameter \( \nu \) is a free parameter, we cannot identify the parameters \( \gamma_2 \) and \( \sigma_0^2 \) using the information of the autocorrelations and cross-correlations. However, the distribution of returns implied by both models is different and therefore, this information should be used to estimate the parameters.

By comparing the T-GASV and E-SV models, we can observe that the autocorrelations are almost identical. Only the autocorrelations of absolute returns of the T-GASV are slightly larger; see also Figure 3. Including \( \alpha \) only has a paltry effect on the volatility clustering that the model can represent. However, the cross-correlations of the T-GASV model are stronger than those of the E-SV model. Therefore, \( \alpha \) allows for a more flexible pattern of the leverage effect.

Finally, the last specification of \( f(\cdot) \) considered in this paper specifies the log-volatility with different constant levels depending on the sign of past returns as follows \( f(\epsilon_t) = \alpha I(\epsilon_t < \)}
This specification has been previously considered by Asai and McAleer (2006) and it is a restricted version of the Threshold SV model proposed by Breidt (1996) and So et al. (2002); see Mao et al. (2013) for the relevance of this specification when compared with the more general Threshold SV model. We denote this model as restricted Threshold SV (RT-SV). Once more, the statistical properties of the RT-SV model can be obtained using the results in the Appendices. The last row of Figure 4 illustrates the shape of the autocorrelations of squared and absolute returns and the cross-correlations between returns and future squared and absolute returns, for a RT-SV model with the same values of the parameters $\phi$, $\sigma_\eta^2$ and $\nu$ as those considered above. Comparing the autocorrelations of squares and absolute returns of the T-GASV model represented in the top panel and the RT-SV model, we can observe that the latter are slightly smaller than the former. However, the cross-correlations are clearly smaller in the RT-SV model. Actually, these cross-correlations are the smallest among those of all the models considered. It seems that the presence of $\alpha$ in the T-GASV model is reinforcing the role of the leverage parameter $\gamma_1$.

5. Finite sample performance of a MCMC estimator of the parameters

Stochastic volatility models are attractive because of their flexibility to represent a high range of the dynamic properties of time series of financial returns often observed when dealing with real data. This flexibility can be attributed to the presence of a further disturbance associated with the volatility process. However, as a consequence of the volatility being unobservable, it is not possible to obtain an analytical expression of the likelihood function. Furthermore, one needs to implement filters to obtain estimates of the latent unobserved volatilities. Thus, the main limitation of SV models is the difficulty involved in the estimation of the parameters and volatilities; see Broto and Ruiz (2004) for a survey on alternative procedures to estimate SV models. In this context, simulation based MCMC procedures are becoming very popular because of their good properties and flexibility to
deal with different specifications and distributions of the errors.\textsuperscript{9} The first Bayesian MCMC approach to estimate SV models with leverage effect was developed by Jacquier et al. (2004). After that, there have been several proposals that try to improve the properties of the MCMC estimators. For example, Omori et al. (2007), Omori and Watanabe (2008) and Nakajima and Omori (2009) implement the efficient sampler of Kim et al. (1998) to SV models with Student-t errors and leverage effect based on $\log y_t^2$. Based on the work of Shephard and Pitt (1997) and Watanabe and Omori (2004), Abanto-Valle et al. (2010) estimate an asymmetric SV model assuming scale mixtures of Normal return distributions while SV models with skew-Student-t and skew-Normal return errors are estimated by Tsiotas (2012) using MCMC. Among the alternative MCMC estimators available in the literature, in this paper, we consider the estimator described by Meyer and Yu (2000) who propose to estimate the A-ARSV model using the user-friendly and freely available WinBUGS software. The estimator uses the single-move Gibbs sampling algorithm; see Yu (2012) and Wang et al. (2013) for empirical implementations. This estimator is attractive because it reduces the coding effort allowing its empirical implementation to real time series of financial returns.

In this section and in the empirical application, we consider the T-GASV model with $\gamma_2 = 0$ and $\epsilon_t \sim GED(\nu)$, denoted as RT-GASV.\textsuperscript{10} Next, we describe briefly the algorithm. Let $p(\theta)$ be the joint prior distribution of the unknown parameters $\theta = \{\mu, \phi, \alpha, \gamma_1, \sigma_\eta^2, \nu\}$. Following Meyer and Yu (2000), the prior densities of $\phi$ and $\sigma_\eta^2$ are $\phi = 2\phi^* - 1$ with $\phi^* \sim Beta(20, 1.5)$ and $\sigma_\eta^2 = 1/\tau^2$ with $\tau \sim IG(2.5, 0.025)$, respectively, where $IG(\cdot, \cdot)$ is the inverse Gaussian distribution.\textsuperscript{11} The remaining prior densities are chosen to be

\textsuperscript{9}There are several alternative procedures proposed in the literature to estimate SV models with leverage effect. For example, Bartolucci and De Luca (2003) propose a likelihood estimator based on the quadrature methods of Fridman and Harris (1998). Alternatively, Harvey and Shephard (1996) propose a Quasi Maximum Likelihood procedure while Sandmann and Koopman (1998) implement a Simulated Maximum Likelihood procedure. Finally Liesenfeld and Richard (2003) propose a Maximum Likelihood approach based upon an efficient importance sampling.

\textsuperscript{10}As mentioned in the Section 4, we might have the problem of parameter identification. Therefore, we exclude $\gamma_2$ from the model.

\textsuperscript{11}Although the prior of $\phi^*$ is very informative, when it is changed to $Beta(1, 1)$, the results are very similar.
uninformative, that is, \( \mu \sim N(0, 10), \alpha \sim N(0.05, 10), \gamma_1 \sim N(-0.05, 10) \) and \( \nu \sim U(0, 4) \). These priors are assumed to be independent. The joint prior density of \( \theta \) and \( h \) is given by

\[
p(\theta, h) = p(\theta)p(h_0) \prod_{t=1}^{T+1} p(h_t|h_{t-1}, \theta).
\] (13)

The likelihood function is then given by

\[
p(y|\theta, h) = \prod_{t=1}^{T} p(y_t|h_t, \theta).
\] (14)

Note that the conditional distribution of \( y_t \) given \( h_t \) and \( \theta \) is \( y_t|h_t, \theta \sim GED(\nu) \). We make use of the scale mixtures of Uniform representation of the GED distribution proposed by Walker and Gutiérrez-Peña (1999) for obtaining the conditional distribution of \( y_t \) given \( \nu \) and \( h_t \), which is given by

\[
y_t|u, h_t \sim U\left(-\frac{\exp(h_t/2)}{\sqrt{2\Gamma(3/\nu)/\Gamma(1/\nu)}} u^{1/\nu}, \frac{\exp(h_t/2)}{\sqrt{2\Gamma(3/\nu)/\Gamma(1/\nu)}} u^{1/\nu}\right),
\] (15)

where \( u|\nu \sim Gamma(1+1/\nu, 2^{-\nu/2}) \). Given the initial values \( (\theta^{(0)}, h^{(0)}) \), the Gibbs sampler generates a Markov Chain for each parameter and volatility in the model through the following steps:

\[
\theta_1^{(1)} \sim p(\theta_1|\theta_2^{(0)}, \ldots, \theta_K^{(0)}, h^{(0)}, y);
\]
\[
\vdots
\]
\[
\theta_K^{(1)} \sim p(\theta_1|\theta_2^{(1)}, \ldots, \theta_{K-1}^{(1)}, h^{(0)}, y);
\]
\[
h_1^{(1)} \sim p(h_1|\theta_1^{(1)}, h_2^{(0)}, \ldots, h_{T+1}^{(0)}, y);
\]
\[
\vdots
\]
\[
h_{T+1}^{(1)} \sim p(h_{T+1}|\theta_1^{(1)}, h_1^{(1)}, \ldots, h_T^{(1)}, y).
\]

The estimates of the parameters and volatilities are the means of the Markov Chain. The
posterior joint distribution of the parameters and volatilities is given by

\[
p(\theta, h|y) \propto p(\theta)p(h_0) \prod_{t=1}^{T+1} p(h_t|h_{t-1}, y, \theta) \prod_{t=1}^{T} p(y_t|h_t, \theta).
\]

(16)

In this section, we carry out extensive Monte Carlo experiments to analyze the finite sample performance of the MCMC estimator when estimating both the parameters and the underlying volatilities of RT-GASV model. We consider two designs for the Monte Carlo experiments. First, \(R\) replicates are generated by the RT-GASV model with parameters \((\mu, \phi, \alpha, \gamma_1, \sigma^2, \nu) = (0, 0.98, 0.07, -0.08, 0.05, 1.5)\). All the parameters are then estimated using the MCMC estimator. The total number of iterations in the MCMC procedure is 20,000 after a burn-in of 10,000. The results are based on \(R = 500\) replicates of series with sample sizes \(T = 500, 1000\) and 2000. Table 1 reports the average and standard deviation of the posterior means together with the average of the posterior standard deviations of each parameter through the Monte Carlo replicates for the first design. We observe that the Monte Carlo averages of the posterior means are rather close to the true parameter values, indicating almost no finite sample biases for series of sizes \(T = 1000\) and 2000. Also, it is important to point out that the average of the posterior standard deviations is rather close to the Monte Carlo standard deviation of the posterior means. Consequently, inference based on the posterior distributions seems to be adequate when the sample size is as large as 1000. When \(T = 500\), the estimation could suffer from small parameter bias.

Second, we also want to check whether by fitting the T-GASV model we are able to identify the true restricted specifications when the distribution of \(\epsilon_t\) is Gaussian. With this purpose, we generate \(R = 500\) replicates of sizes \(T = 500\) and 1000 from each of the restricted models, A-ARSV and RT-SV, with Gaussian return errors and fit the T-GASV model while treating \(\nu\) as being known, \(\nu = 2\). The results, reported in Table 2, provide evidence that when \(\nu\) is known it is possible to identify the true data generating process (DGP) by fitting the more general T-GASV model even when the sample size is as small as \(T = 500\).  

21
Summarizing the Monte Carlo results on the MCMC estimator considered in this paper, we can conclude that: i) If the sample size is moderately large, the posterior distribution gives an adequate representation of the finite sample distribution with the posterior mean being an unbiased estimator of the true parameter value. ii) The true restricted specifications are correctly identified after fitting the proposed T-GASV model when $\nu$ is known.

When dealing with conditional heteroscedastic models, practitioners are interested not only in the parameter estimates but also, and more importantly, in the volatility estimates. Consequently, in the Monte Carlo experiments above, at each time period $t$ and for each replicate $i$, we also compute the relative prediction error of volatility, $e_t^{(i)} = (\sigma_t^{(i)} - \hat{\sigma}_t^{(i)})/\sigma_t^{(i)}$, where $\sigma_t^{(i)}$ is the simulated true volatility at time $t$ in the $i$-th replicate and $\hat{\sigma}_t^{(i)}$ is its MCMC estimate. Table 3 reports the average and standard deviation through time of $m_t = \sum_{i=1}^R e_t^{(i)}/R$ together with the average through time of the standard deviations given by $s_t = \sqrt{\sum_{i=1}^R (e_t^{(i)} - m_t)^2/(R - 1)}$ when $T = 500$ and 1000. These quantities have been computed for the Monte Carlo experiments conducted above. Consider first the results when the RT-GASV model is the true DGP. We observe that the estimates of the volatility are unbiased. Further, when the restricted models are the DGPs but the general T-GASV model is fitted, the errors are also insignificant and with similar standard deviations. In all cases the relative errors are negative. Therefore, the MCMC estimated volatilities are insignificantly larger than the true underlying volatilities.

6. Empirical application

6.1. Data description and estimation results

In this section, the RT-GASV model is fitted to represent the dynamic dependence of daily S&P500 returns observed from September 1, 1998 to July 25, 2014 with $T = 4000$ observations. The returns, computed as usual as $y_t = 100 \times \triangle \log P_t$, where $P_t$ is the adjusted close price from yahoo.finance on day $t$, have kurtosis 10.402 and skewness -0.154 which is not statistically significant. Therefore, it seems that it is not necessary to consider a skewed
distribution of the return errors. The raw prices together with their corresponding returns are plotted in Figure 5 which suggests the presence of volatility clustering with episodes of large volatilities associated with periods of negative movements in prices. Furthermore, this association between large volatilities and negative returns can also be observed in the negative sample cross-correlations between returns and future squared and absolute returns plotted in Figure 6. It is clear that the volatility clustering and leverage effect are present in the daily S&P500 returns. Consequently, the RT-GASV model is fitted first estimating $\nu$ as a free parameter and second assuming that the errors are Gaussian. Our objective is to observe empirically whether the estimated volatilities and the corresponding Value at Risk (VaR) are affected by the distribution of $\epsilon_t$. For completeness, we also fit the other two restricted models. All the parameters and volatilities have been estimated implementing the MCMC estimator of WinBUGS.

Table 4 reports the posterior mean and the 95% credible interval of the MCMC estimator of each parameter. The left panel reports the results of those models with GED errors while the right panel for the models with Normal errors. First, check the results of the Gaussian models. We can observe that when the RT-GASV model is fitted, the credible intervals for the threshold parameter $\alpha$ contain the zero. The Monte Carlo experiments in the previous section suggest that fitting the general RT-GASV model proposed in this paper, one could identify the true restricted specification of the log-volatilities if the distribution of $\epsilon_t$ is Gaussian. Consequently, it seems that the threshold parameter is not needed to represent the conditional heteroscedasticity of the S&P500 returns when $\epsilon_t \sim N(0, 1)$. Moreover, the DIC of the RT-SV model is larger than those of the other two models; see Berg et al. (2004), Wang et al. (2013) and Tsiotas (2012) for the use of the DIC to compare models in the context of SV models. Second, when $\epsilon_t \sim GED(\nu)$, the DIC of the RT-SV is similar to that of RT-GASV model and larger than that of the A-ARSV model. The same conclusion can be drawn under the log-likelihood criterion which is also reported in Table 4. Furthermore, under both criteria, models with GED errors outperform the counterparts with Gaussian errors. We conclude that the RT-GASV fits the S&P500 returns as well as RT-SV when
\( \epsilon_t \sim GED(\nu) \) and better than A-ARSV model. Figure 6 plots the plug-in moments implied by the estimated asymmetric SV models together with the corresponding sample moments. The plug-in moments given by the models with GED errors are always closer to the sample moments comparing with those of the corresponding models with Normal errors.

Given the apparent similarity in-sample between the RT-GASV and RT-SV specifications with GED errors, next we check whether they can generate significant differences when predicting the VaRs out-of-sample.

6.2. Forecasting VaR

In this subsection, we perform an out-of-sample comparison of the ability of the alternative asymmetric SV models considered in this paper, with \( \epsilon_t \) following either a GED or a Normal distribution, when evaluating the one-step-ahead VaR of the daily S&P500 returns. Given the extremely heavy computations involved in the estimation of the one-step-ahead VaR based on the MCMC estimator, we compute it using data from January 4, 2010 to July 25, 2014. The parameters are estimated using a rolling-window scheme fixing \( T = 1006 \) observations.\(^\text{12}\) Moreover, one-step-ahead VaRs are obtained starting on January 2, 2014 until July 25, 2014 as

\[
VaR_{t+1|t}(m) = q\hat{\sigma}_{t+1|t},
\]

with \( q \) being the 5\% quantile of the distribution with parameter \( \nu \) estimated in model \( m \) or the 5\% quantile of the Normal distribution when \( \nu = 2 \) and \( \hat{\sigma}_{t+1|t} \) is the estimated one-step-ahead volatility. Finally, we obtain 142 one-step-ahead VaRs.

In order to evaluate the adequacy of the interval forecasts provided by the VaRs computed as in equation (17) for each of the models, Table 5 reports the failure rates. We can observe that the failure rate of the RT-GASV model with GED error is the smallest and the closest to

\(^{12}\text{Checking the estimates obtained, we observe that all the estimates are very stable over the year considered in the rolling window estimation.}\)
the level 0.05. Therefore, our RT-GASV model with GED error provides the best prediction of volatilities for this S&P500 return series.

7. Conclusions

In this paper, we derive the statistical properties of a general family of asymmetric SV models named as GASV. Some of the most popular asymmetric SV models usually implemented when modeling heteroscedastic series with leverage effect can be included within the GASV family. In particular, the A-ARSV model which incorporates the leverage effect through the correlation between the disturbances in the level and log-volatility equations, the E-SV model which adds a noise to the log-volatility equation specified as an EGARCH model and a restricted T-SV model, in which the constant of the volatility equation is different depending on whether one-lagged returns are positive or negative, are included within the GASV family. We propose a new model named T-GASV which belongs to the GASV family and nests these particular specifications. Closed-form expressions of the statistical properties of these models are obtained, some of which were previously unknown in the literature. Therefore, we are able to analyze the role of each parameter in these models. Particularly, closed-form expressions of the variance, kurtosis, autocorrelations of power-transformed absolute returns and cross-correlations between returns and future power-transformed absolute returns are obtained when the disturbance of the log-volatility equation is Gaussian and the disturbance of the level equation follows a GED distribution with parameter strictly larger than one. We show that some care should be put in the specification of the log-volatility equation as, depending on the specification, some moments of returns can be indistinguishable for different combinations of parameters and distributions.

Another contribution of this paper is the proposal of the SNIS to describe the asymmetric response of volatility to positive and negative past returns in the context of SV models. One attractive feature of the SNIS is that it allows to observe how the asymmetric response of the volatility is different depending on the size and sign of the volatility shock.

Moreover, we analyze the finite sample properties of a MCMC estimator of the parameters
and volatilities of the T-GASV model using the WinBUGS software. We show that estimating the proposed restricted T-GASV model allows to correctly identify the true data generating process when the distribution of the level disturbance is known. Finally, the restricted T-GASV model is fitted to estimate the volatilities of S&P500 daily returns. For this particular data set, when estimating the VaRs our model with GED errors outperforms the benchmarks considered in this paper.

Several possible extensions of this paper could be of interest for future research. First, our focus is on univariate models. However, multivariate asymmetric models are attracting a great deal of interest in the literature; see, for example, Harvey et al. (1994), Asai and McAleer (2006), Chan et al. (2006), Chib et al. (2006), Jungbacker and Koopman (2006) and Yu and Meyer (2006). Deriving the statistical properties of multivariate GASV models is in our research agenda. Second, Rodríguez and Ruiz (2012) compare the properties of alternative asymmetric GARCH models to see which is closer to the empirical properties often observed when dealing with financial returns. In this paper, we compare alternative SV models. Comparing the properties of these alternative SV models with those of the best candidates within the GARCH family is also left for further research.
Figure 1: SNIS of different GASV models with $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ and $exp((1 - \phi)\mu)\sigma_y^{2\phi} = 1$. Top panel $f(\epsilon_t) = \alpha I(\epsilon_t < 0) + \gamma_1\epsilon_t + \gamma_2|\epsilon_t|$, middle panel $f(\epsilon_t) = \gamma_1\epsilon_t + \gamma_2|\epsilon_t|$ and bottom panel $f(\epsilon_t) = \gamma_1\epsilon_t$. The parameter values are $\{\alpha, \gamma_1, \gamma_2\} = \{0.07, -0.08, 0.1\}$. 
Figure 2: Ratio between the kurtoses of the T-GASV model and the symmetric ARSV model with Gaussian errors when $\gamma_2 = 0.1$ (left column) and 0 (right column) for three different values of the persistence parameter, $\phi = 0.5$ (first row), $\phi = 0.9$ (middle row) and $\phi = 0.98$ (bottom row).
Figure 3: First order autocorrelations of squares (top left panel), first order autocorrelations of absolute returns (top right panel), first order cross-correlations between returns and future squared returns (bottom left panel) and first order cross-correlations between returns and future absolute returns (bottom right panel) of different T-GASV models with parameters $\phi = 0.98$ and $\sigma^2_\eta = 0.05$. 

[Diagram of autocorrelations and cross-correlations]
Figure 4: Autocorrelations of squares (first column), autocorrelations of absolute returns (second column), cross-correlations between returns and future squared returns (third column) and cross-correlations between returns and future absolute returns (fourth column) for different specifications of asymmetric SV models. The first row corresponds to a T-GASV model with $\alpha = 0.07, \phi = 0.98, \sigma^2 = 0.05, \gamma_1 = -0.08, \gamma_2 = 0.1$ and $\nu = 1.5$ (solid lines), $\nu = 1.7$ (dashed lines), $\nu = 2$ (dotted lines) and $\nu = 2.5$ (dashdot lines). The second row corresponds to the A-ARSV with $\alpha = \gamma_2 = 0$. The third row matches along with the E-SV model with $\alpha = 0$. Finally, the last row corresponds to the RT-SV model with $\gamma_1 = \gamma_2 = 0$. 
Figure 5: S&P500 daily prices (bottom line) and returns (top line) observed from September 1, 1998 up to July 25, 2014.
Figure 6: Sample autocorrelations of squares (first column), autocorrelations of absolute returns (second column), cross-correlations of returns and future squared returns (third column) and cross-correlations between absolute returns and lagged returns (fourth column) together with the corresponding plug-in moments obtained after fitting the RT-GASV (first row), A-ARSV (second row) or RT-SV (third row) models to the daily S&P500 returns. The continuous lines correspond to the moments implied by the models estimated with a Gaussian distribution while the dotted lines correspond to the models estimated when the distribution is GED.
### Table 1: Monte Carlo results of the MCMC estimator of the parameters of the RT-GASV model. Reported are the values of the Monte Carlo average and standard deviation (in parenthesis) of the posterior means together with the Monte Carlo average of the posterior standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\phi$</th>
<th>$\alpha$</th>
<th>$\gamma_1$</th>
<th>$\sigma^2_\eta$</th>
<th>$\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>0</td>
<td>0.98</td>
<td>0.07</td>
<td>-0.08</td>
<td>0.05</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.268</td>
<td>0.952</td>
<td>0.108</td>
<td>-0.074</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.445)</td>
<td>(0.063)</td>
<td>(0.126)</td>
<td>(0.077)</td>
<td>(0.068)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.803</td>
<td>0.018</td>
<td>0.121</td>
<td>0.066</td>
<td>0.034</td>
</tr>
<tr>
<td>T=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.046</td>
<td>0.974</td>
<td>0.077</td>
<td>-0.082</td>
<td>0.055</td>
<td>1.520</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.442)</td>
<td>(0.010)</td>
<td>(0.073)</td>
<td>(0.041)</td>
<td>(0.020)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.779</td>
<td>0.009</td>
<td>0.078</td>
<td>0.042</td>
<td>0.017</td>
</tr>
<tr>
<td>T=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.082</td>
<td>0.977</td>
<td>0.072</td>
<td>-0.082</td>
<td>0.053</td>
<td>1.528</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.278)</td>
<td>(0.006)</td>
<td>(0.0560)</td>
<td>(0.031)</td>
<td>(0.013)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.422</td>
<td>0.006</td>
<td>0.057</td>
<td>0.030</td>
<td>0.011</td>
</tr>
<tr>
<td>T=2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Monte Carlo results of MCMC estimator of the parameters of the RT-GASV model fitted to series simulated from different asymmetric SV models with Gaussian return errors. Reported are the values of the Monte Carlo average and standard deviation (in parenthesis) of the posterior means together with the Monte Carlo average of the posterior standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\phi$</th>
<th>$\alpha$</th>
<th>$\gamma_1$</th>
<th>$\sigma^2_\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>-0.042</td>
<td>0.965</td>
<td>0.003</td>
<td>-0.082</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.475)</td>
<td>(0.046)</td>
<td>(0.095)</td>
<td>(0.057)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.854</td>
<td>0.014</td>
<td>0.103</td>
<td>0.056</td>
</tr>
<tr>
<td>T=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.090</td>
<td>0.973</td>
<td>0.004</td>
<td>-0.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.391)</td>
<td>(0.045)</td>
<td>(0.067)</td>
<td>(0.039)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.648</td>
<td>0.009</td>
<td>0.073</td>
<td>0.039</td>
</tr>
<tr>
<td>T=1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.304</td>
<td>0.963</td>
<td>0.094</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.517)</td>
<td>(0.019)</td>
<td>(0.112)</td>
<td>(0.062)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.919</td>
<td>0.016</td>
<td>0.112</td>
<td>0.060</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\phi$</th>
<th>$\alpha$</th>
<th>$\gamma_1$</th>
<th>$\sigma^2_\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>0.009</td>
<td>0.974</td>
<td>0.078</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.593)</td>
<td>(0.010)</td>
<td>(0.076)</td>
<td>(0.042)</td>
</tr>
<tr>
<td></td>
<td>s.d.</td>
<td>1.756</td>
<td>0.009</td>
<td>0.079</td>
<td>0.041</td>
</tr>
<tr>
<td>T=2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\( \epsilon_t \sim N(0,1) \) \hspace{1cm} \( \epsilon_t \sim GED(\nu) \)

<table>
<thead>
<tr>
<th></th>
<th>RT-GASV</th>
<th>A-ARSV</th>
<th>RT-SV</th>
<th>RT-GASV</th>
<th>A-ARSV</th>
<th>RT-SV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T=500</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>-0.018</td>
<td>-0.024</td>
<td>-0.020</td>
<td>-0.022</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.017)</td>
<td>(0.018)</td>
<td>(0.040)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>s.d.</strong></td>
<td>0.219</td>
<td>0.222</td>
<td>0.218</td>
<td>0.240</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T=1000</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>-0.020</td>
<td>-0.022</td>
<td>-0.022</td>
<td>-0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.013)</td>
<td>(0.026)</td>
<td>(0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>s.d.</strong></td>
<td>0.212</td>
<td>0.216</td>
<td>0.213</td>
<td>0.232</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Monte Carlo results of the relative volatility prediction errors. Reported are the values of the time average and standard deviation (in parenthesis) of \( m_t = \sum_{i=1}^{R} \frac{e_{t}^{(i)}}{R} \) together with the time average of \( s_t = \sqrt{ \sum_{i=1}^{R} (e_{t}^{(i)} - m_t)^2 / (R - 1) } \), where \( e_{t}^{(i)} = (\sigma_t^{(i)} - \hat{\sigma}_t^{(i)}) / \sigma_t^{(i)} \).

<table>
<thead>
<tr>
<th>( \epsilon_t \sim GED(\nu) )</th>
<th>( \epsilon_t \sim N(0,1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RT-GASV</td>
</tr>
<tr>
<td><strong>( \mu )</strong></td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>(-0.019,0.079)</td>
</tr>
<tr>
<td><strong>( \phi )</strong></td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>(0.974,0.993)</td>
</tr>
<tr>
<td><strong>( \alpha )</strong></td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>(-0.019,0.079)</td>
</tr>
<tr>
<td><strong>( \gamma_1 )</strong></td>
<td>-0.129</td>
</tr>
<tr>
<td></td>
<td>(-0.155,-0.103)</td>
</tr>
<tr>
<td><strong>( \sigma_n^2 )</strong></td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.000,0.018)</td>
</tr>
<tr>
<td><strong>( \nu )</strong></td>
<td>1.359</td>
</tr>
<tr>
<td></td>
<td>(1.237,1.382)</td>
</tr>
<tr>
<td><strong>DIC</strong></td>
<td>6497</td>
</tr>
<tr>
<td><strong>Log-Likelihood</strong></td>
<td>-5.728</td>
</tr>
</tbody>
</table>

Table 4: MCMC estimates of the parameters of alternative asymmetric SV models for S&P500 daily returns. The values reported are the mean and 95% credible interval (in parenthesis) of the posterior distributions.
<table>
<thead>
<tr>
<th>𝜖ₜ ∼ \text{GED}(ν)</th>
<th>Failure Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-ARSV</td>
<td>0.056</td>
</tr>
<tr>
<td>RT-SV</td>
<td>0.085</td>
</tr>
<tr>
<td>RT-GASV</td>
<td>0.049</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>𝜖ₜ ∼ \text{N}(0, 1)</th>
<th>Failure Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-ARSV</td>
<td>0.092</td>
</tr>
<tr>
<td>RT-SV</td>
<td>0.099</td>
</tr>
<tr>
<td>RT-GASV</td>
<td>0.070</td>
</tr>
</tbody>
</table>

Table 5: Failure rates.
Appendix A. Proof of Theorems

Appendix A.1. Proof of Theorem 2.1

Consider $y_t$, which, according to equation (1), is given by $y_t = \epsilon_t \exp(h_t/2)$. From equation (2), $h_t$ can be written as

$$h_t - \mu = \sum_{i=1}^{\infty} \phi^{i-1}(f(\epsilon_{t-i}) + \eta_{t-i}). \quad (A.1)$$

First, note that if $|\phi| < 1$ and $x = (x_1, x_2, \cdots) \in \mathbb{R}_\infty$, then $\Psi(x) = \sum_{i=1}^{\infty} \phi^{i-1} x_i$ is a measurable function. Given that for any $x_0$ and $\forall \varsigma > 0$, we can find a value of $\delta = \sqrt{1 - \phi^2} \varsigma$, such that $\forall x$ satisfying $|x - x_0| = \sqrt{\sum_{i=1}^{\infty} (x_i - x_0^i)^2} < \delta$, we have $|\Psi(x) - \Psi(x_0)| = |\sum_{i=1}^{\infty} \phi^{i-1} (x_i - x_0^i)|$. Using the Cauchy-Schwarz inequality, it follows that $|\Psi(x) - \Psi(x_0)| \leq \sqrt{\sum_{i=1}^{\infty} \phi^{2i-2}} \sqrt{\sum_{i=1}^{\infty} (x_i - x_0^i)^2} < \frac{\delta}{\sqrt{1 - \phi^2}} = \varsigma$. Therefore, $\Psi(x)$ is continuous, and consequently, measurable.

Second, given that $\epsilon_t$ and $\eta_t$ are both IID and mutually independent for any lag and lead, then $\{f(\epsilon_t) + \eta_t\}$ is also an IID sequence. Lemma 3.5.8 of Stout (1974) states that an IID sequence is always strictly stationary. Therefore, in (A.1), if $|\phi| < 1$, $h_t$ is expressed as a measurable function of a strictly stationary process and, consequently, according to Theorem 3.5.8 of Stout (1974), $h_t$ is strictly stationary. As $\sigma_t$ is a continuous function of $h_t$, $\sigma_t$ is also strictly stationary. The level noise $\epsilon_t$ is independent of $\sigma_t$ and strictly stationary by definition. Therefore, it is easy to show that $y_t = \sigma_t \epsilon_t$ is strictly stationary.

When $|\phi| < 1$, $y_t$ and $\sigma_t^2$ are strictly stationary and, consequently, any existing moments are time invariant. Next we show that $\sigma_t$ has finite moments of arbitrary positive order $c$ when $\epsilon_t$ follows a distribution such that $E(\exp(0.5 c f(\epsilon_t))) < \infty$.

From expression (A.1), the power-transformed volatility can be written as follows

$$\sigma_t^c = \exp(0.5 c \mu) \exp \left(0.5 c \sum_{i=1}^{\infty} \phi^{i-1}(f(\epsilon_{t-i}) + \eta_{t-i})\right). \quad (A.2)$$

Given that $\epsilon_t$ and $\eta_t$ are mutually independent for all lags and leads, the following expression
is obtained after taking expectations on both sides of equation (A.6)

\[ E(\sigma_i^2) = \exp(0.5c\mu)E\left\{ \exp\left(0.5c \sum_{i=1}^{\infty} \phi^{i-1} f(\epsilon_{t-i})\right) \right\} E\left\{ \exp\left(0.5c \sum_{i=1}^{\infty} \phi^{i-1} \eta_{t-i}\right) \right\} . \] (A.3)

As \( \eta_t \) is Gaussian, the last expectation in (A.3) can be evaluated using the expression of the moments of the Log-Normal. Furthermore, given that \( \eta_t \) and \( \epsilon_t \) are both IID sequences, it is easy to show that (A.3) becomes

\[ E(\sigma_i^2) = \exp(0.5c\mu) \exp\left(\frac{c^2 \sigma^2_{\eta}}{8 (1 - \phi^2)}\right) \prod_{i=1}^{\infty} E\left\{ \exp\left(0.5c\phi^{i-1} f(\epsilon_{t-i})\right) \right\} . \] (A.4)

We need to show that \( P(0.5c\phi^{i-1}) = \prod_{i=1}^{\infty} E\left\{ \exp\left(0.5c\phi^{i-1} f(\epsilon_{t-i})\right) \right\} \) is finite when \( E\left( \exp\left(0.5c f(\epsilon_{t-i})\right) \right) < \infty \). In general, we are going to prove that when \( \sum_{i=1}^{\infty} |b_i| < \infty \) and \( E(\exp(b_i f(\epsilon_{t-i}))) < \infty \), then \( P(b_i) = \prod_{i=1}^{\infty} E\left\{ \exp\left(b_i f(\epsilon_{t-i})\right) \right\} \) is always finite.

Define \( a_i = E(\exp(b_i f(\epsilon_{t-i}))) \). As \( 0 < a_i < \infty \), according to Section 0.25 of Ryzhik et al. (2007), the sufficient and necessary condition for the infinite product \( \prod_{i=1}^{\infty} a_i \) to converge to a finite, nonzero number is that the series \( \sum_{i=1}^{\infty} (a_i - 1) \) converge. Expanding \( a_i \) in Taylor series around \( b_i = 0 \), we have

\[ a_i - 1 = O(b_i) \quad \text{as} \quad b_i \to 0. \]

Consequently, for some \( \zeta > 0 \), there exist a finite \( M \) independent of \( i \) such that

\[ \sup_{|b_i| < \zeta, b_i \neq 0} |O(b_i)| < M|b_i|. \]

\( \sum_{i=1}^{\infty} |b_i| < \infty \) implies \( \sum_{i=1}^{\infty} |a_i - 1| < \infty \), therefore \( \sum_{i=1}^{\infty} (a_i - 1) < \infty \). Thus \( P(b_i) = \prod_{i=1}^{\infty} a_i < \infty \).

Here \( b_i = 0.5c\phi^{i-1} \). Therefore, if \( |\phi| < 1 \), then \( \sum_{i=1}^{\infty} |b_i| = \frac{0.5c}{1 - \phi} < \infty \). Thus, the product \( \prod_{i=1}^{\infty} E(\exp(0.5c\phi^{i-1} f(\epsilon_{t-i}))) \) and, consequently, \( E(\sigma_i^2) \) are finite when \( E(\exp(0.5c\phi^{i-1} f(\epsilon_{t-i}))) < \infty \). Note that when \( |\phi| < 1 \), \( E(\exp(0.5c f(\epsilon_i))) < \infty \) guarantees that \( E(\exp(0.5c\phi^{i-1} f(\epsilon_{t-i}))) < \infty \) for any positive integer \( i \). Therefore, if \( |\phi| < 1 \) and \( E(\exp(0.5c f(\epsilon_i))) < \infty \), \( E(\sigma_i^2) \) is finite.
Finally, consider $y_t$, which, according to equation (1), is given by $y_t = \sigma_t \epsilon_t$. Therefore, given that $\sigma_t$ and $\epsilon_t$ are contemporaneously independent, the following expressions are obtained

$$E(|y_t|^c) = E(\sigma_t^c)E(|\epsilon_t|^c), \quad (A.5)$$

$$E(y_t^c) = E(\sigma_t^c)E(\epsilon_t^c). \quad (A.6)$$

Replacing formula (A.4) into (A.5) yields the following required expression

$$E(|y_t|^c) = \exp(0.5c\mu)E(|\epsilon_t|^c) \exp\left(\frac{c^2\sigma_n^2}{8(1 - \phi^2)}\right) P(0.5c\phi^{i-1}), \quad (A.7)$$

where $P(b_i) \equiv \prod_{i=1}^\infty E(\exp(b_i f(\epsilon_t - i)))$. Therefore, if further $\epsilon_t$ follows a distribution such that $E(\epsilon_t^c) < \infty$, which is equivalent to $E(|\epsilon_t|^c) < \infty$, then $|y_t|$ has finite moments of arbitrary order $c$. On the other hand, following the same steps, we obtain

$$E(y_t^c) = \exp(0.5c\mu)E(\epsilon_t^c) \exp\left(\frac{c^2\sigma_n^2}{8(1 - \phi^2)}\right) P(0.5c\phi^{i-1}). \quad (A.8)$$

Thus, $E(y_t^c) < \infty$ if $|\phi| < 1$, $E(\epsilon_t^c) < \infty$ and $E(\exp(0.5cf(\epsilon_t))) < \infty$.

**Appendix A.2. Proof of Theorem 2.2**

Consider $y_t$ as given in equations (1) and (2). We first compute the $\tau$-th order auto-covariance of $|y_t|^c$ which is given by

$$E(|\epsilon_t|^c | \sigma_t^c | \epsilon_{t-\tau} | \sigma_{t-\tau}^c) - [E(|y_t|^c)]^2. \quad (A.9)$$

Note that from equation (2), $\sigma_t^c = \exp\{0.5c\mu\} \exp\{0.5c \sum_{i=1}^\tau \phi^{i-1}(f(\epsilon_t - i) + \eta_{t-i})\}$ can be written as follows

$$\sigma_t^c = \exp\{0.5c\mu(1 - \phi^\tau)\} \exp\left\{0.5c \sum_{i=1}^\tau \phi^{i-1}(f(\epsilon_t - i) + \eta_{t-i})\right\} \sigma_{t-\tau}^{c\phi^\tau}. \quad (A.10)$$
The following expression of the auto-covariance is obtained after substituting (A.7) and (A.10) into (A.9)

\[
\text{cov}(|y_t|^c, |y_{t-\tau}|^c) = \\
E \left( |\epsilon_t|^c |\epsilon_{t-\tau}|^c \exp(0.5c\mu(1 - \phi^\tau)) \exp \left( \sum_{i=1}^{\tau} 0.5c\phi^{i-1}(f(\epsilon_{t-i} + \eta_i)) \right) \right) \sigma_{t-\tau}^{c(\phi^\tau+1)} - \left\{ \exp(0.5c\mu)E(|\epsilon_i|^c) \exp \left( \frac{c^2\sigma_\eta^2}{8(1 - \phi^2)} \right) P(0.5c\phi^{i-1}) \right\}^2.
\]

(A.11)

Given that \(\epsilon_t\) and \(\eta_i\) are IID sequences mutually independent for any lag and lead and that \(\sigma_{t-\tau}\) only depends on lagged disturbances, substituting the time-invariant moment of \(\sigma_t\) in (A.4), equation (A.11) can be written as follows

\[
\text{cov}(|y_t|^c, |y_{t-\tau}|^c) = \\
\exp(c\mu)E(|\epsilon_i|^c) \exp \left( \frac{1 + \phi^\tau}{4(1 - \phi^2)} c^2\sigma_\eta^2 \right) E \left( |\epsilon_i|^c \exp(0.5c\phi^{i-1}f(\epsilon_i)) \right) \prod_{i=1}^{\tau-1} E \left( \exp(0.5c\phi^{i-1}f(\epsilon_{t-i})) \right) \\
\prod_{i=1}^{\infty} E \left( \exp(0.5c(1 + \phi^\tau)\phi^{i-1}f(\epsilon_{t-i})) \right) - \exp(c\mu)(E(|\epsilon_i|^c))^2 \exp \left( \frac{c^2\sigma_\eta^2}{4(1 - \phi^2)} \right) P(0.5c\phi^{i-1})^2.
\]

The required expression of \(\rho_c(\tau)\) follows directly from \(\rho_c(\tau) = \frac{\text{cov}(|y_t|^c, |y_{t-\tau}|^c)}{E(|y_t|^c)^2 - E(|y_t|^c)^2}\), where the denominator can be obtained from (A.7).

**Appendix A.3. Proof of Theorem 2.3**

The calculation of the cross-covariance between \(|y_t|^c\) and \(y_{t-\tau}\) is obtained following the same steps as in Appendix A.2. That is

\[
\text{cov}(|y_t|^c, y_{t-\tau}) = \exp(0.5(c + 1)\mu)E(|\epsilon_i|^c) \exp \left( \frac{1 + c^2 + 2c\phi^\tau}{8(1 - \phi^2)} \right) E \left( \epsilon_i \exp(0.5c\phi^{i-1}f(\epsilon_i)) \right) \\
\prod_{i=1}^{\infty} E \left( \exp(0.5(c + \phi^\tau)\phi^{i-1}f(\epsilon_{t-i})) \right) \prod_{i=1}^{\tau-1} E \left( \exp(0.5c\phi^{i-1}f(\epsilon_{t-i})) \right).
\]

(A.12)

Finally, \(\rho_{c1}(\tau) = \frac{\text{cov}(|y_t|^c, y_{t-\tau})}{\sqrt{E(|y_t|^c)^2} - E^2(|y_t|^c)}\) together with (A.7) and (A.12) yields the required equation (8).
Appendix B. Expectations

Appendix B.1. Expectations needed to compute $E(|y_t|^c)$, $\text{corr}(|y_t|^c, |y_{t+\tau}|^c)$ and $\text{corr}(y_t, |y_{t+\tau}|^c)$ when $\epsilon \sim \text{GED}(\nu)$ and $f(\epsilon) = \alpha I(\epsilon < 0) + \gamma_1 \epsilon + \gamma_2 |\epsilon|$

If $\epsilon$ has a centered and standardized GED distribution, with parameter $0 < \nu \leq \infty$, then, the density function of $\epsilon$ is given by $\psi(\epsilon) = C_0 \exp\left(-\frac{|\epsilon|^{\nu}}{2\lambda}\right)$, where $C_0 = \frac{\nu \lambda^{\frac{1}{2}}}{\Gamma\left(\frac{1}{\nu}\right)}$ and $\lambda \equiv \left(\frac{2^{-2/\nu} \Gamma\left(1/\nu\right)}{\Gamma\left(3/\nu\right)}\right)^{1/2}$, with $\Gamma(\cdot)$ being the Gamma function. Thus, given that the distribution of $\epsilon$ is symmetric with support $(-\infty, \infty)$, if $p$ is a nonnegative finite integer, then

$$E(|\epsilon|^p) = C_0 \int_{-\infty}^{+\infty} |\epsilon|^p \exp\left(-\frac{|\epsilon|^{\nu}}{2\lambda}\right) d\epsilon$$

$$= 2C_0 \int_{0}^{+\infty} \epsilon^p \exp\left(-\frac{\epsilon^{\nu}}{2\lambda}\right) d\epsilon. \quad \text{(B.1)}$$

On the other hand,

$$E(|\epsilon|^p \exp(bf(\epsilon))) = \int_{-\infty}^{+\infty} |\epsilon|^p \exp(b\alpha I(\epsilon < 0) + b\gamma_1 \epsilon + b\gamma_2 |\epsilon|) C_0 \exp\left(-\frac{|\epsilon|^{\nu}}{2\lambda}\right) d\epsilon$$

$$= C_0 \left[ \int_{-\infty}^{0} (-\epsilon)^p \exp(b\alpha) \exp(b(\gamma_1 - \gamma_2)\epsilon) \exp\left(-\frac{(-\epsilon)^{\nu}}{2\lambda}\right) d\epsilon \right.$$

$$+ \left. \int_{0}^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \exp\left(-\frac{\epsilon^{\nu}}{2\lambda}\right) d\epsilon \right].$$

Integrating by substitution with $s = -\epsilon$ in the first integral, we obtain

$$E(|\epsilon|^p \exp(bf(\epsilon))) = C_0 \left[ \int_{0}^{+\infty} s^p \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1) s) \exp\left(-\frac{s^{\nu}}{2\lambda}\right) ds \right.$$

$$+ \left. \int_{0}^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \exp\left(-\frac{\epsilon^{\nu}}{2\lambda}\right) d\epsilon \right] \quad \text{(B.2)}$$

$$= C_0 \int_{0}^{+\infty} \epsilon^p \exp\left(-\frac{\epsilon^{\nu}}{2\lambda}\right) \left[ \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\epsilon) + \exp(b(\gamma_1 + \gamma_2)\epsilon) \right] d\epsilon.$$
We can rewrite the previous equation by replacing $\epsilon$ with $\lambda (2y)^{1/\nu}$ as follows

\[
E(|\epsilon|^p \exp(bf(\epsilon))) = C_0 \frac{\lambda^{p+1} 2^{1+\frac{p}{\nu}}}{\nu} \int_0^{+\infty} y^{-1+\frac{1+p}{\nu}} \exp(-y) \left[ \exp(b \alpha) \exp(b(\gamma_2 - \gamma_1) \lambda 2^{\frac{1}{2}} y^{\frac{1}{\nu}}) + \exp(b(\gamma_1 + \gamma_2) \lambda 2^{\frac{1}{2}} y^{\frac{1}{\nu}}) \right] dy.
\]

Expanding the expression within the square brackets in a Taylor series and substituting $C_0$, the following expression is obtained

\[
E(|\epsilon|^p \exp(bf(\epsilon))) = \frac{\lambda^p 2^{\frac{1}{2}-1}}{\Gamma\left(\frac{1}{\nu}\right)} \int_0^{+\infty} \sum_{k=0}^{+\infty} \left[ \exp(b \alpha) \left( b \lambda 2^{\frac{1}{2}} (\gamma_2 - \gamma_1) \right)^k + \left( b \lambda 2^{\frac{1}{2}} (\gamma_1 + \gamma_2) \right)^k \right] y^{-1+\frac{1+p+k}{\nu}} \frac{\exp(-y)}{k!} dy.
\]

Define $\Delta = \max \{|b \lambda 2^{1/\nu}(\gamma_1 + \gamma_2)|, \max(\exp(b \alpha), 1)|b \lambda 2^{1/\nu}(\gamma_2 - \gamma_1)|\}$ . Then, we can use the results in Nelson (1991) to show that if $\nu > 1$ then the summation and integration in (B.3) can be interchanged. Further, applying Formula 3.381 #4 of Ryzhik et al. (2007) yields the following required expression\(^\text{13}\)

\[
E(|\epsilon|^p \exp(bf(\epsilon))) = 2^p \frac{\lambda^p}{\nu} \sum_{k=0}^{\infty} \left( 2^{1/\nu} \lambda b \right)^k \left[ (\gamma_1 + \gamma_2)^k + \exp(b \alpha) (\gamma_2 - \gamma_1)^k \right] \frac{\Gamma((p+k+1)/\nu)}{2 \Gamma(1/\nu) k!} < \infty. \tag{B.4}
\]

Following the same steps, the following required expression is obtained when $\nu > 1$,

\[
E(\epsilon^p \exp(bf(\epsilon))) = 2^p \frac{\lambda^p}{\nu} \sum_{k=0}^{\infty} \left( 2^{1/\nu} \lambda b \right)^k \left[ (\gamma_1 + \gamma_2)^k + (-1)^p \exp(b \alpha) (\gamma_2 - \gamma_1)^k \right] \frac{\Gamma((p+k+1)/\nu)}{2 \Gamma(1/\nu) k!} < \infty. \tag{B.5}
\]

Note that the expectations (B.4) and (B.5) are only valid when $\nu > 1$. When $0 < \nu \leq 1$, it is not possible to obtain closed-form expression of the required expectations. In this case, we can only obtain the conditions for the expectations to be finite. When $0 < \nu < 1$, it is very easy to verify that $E(|\epsilon|^p \exp(bf(\epsilon))) < \infty$ if and only if the both integrals in (B.2)

\(^{13}\)See Nelson (1991) for the proof of finiteness of the formula.
are finite, which holds if and only if \( b(\gamma_2 - \gamma_1) \leq 0 \) and \( b(\gamma_2 + \gamma_1) \leq 0 \). When \( \nu = 1 \), similarly, the sufficient and necessary conditions for the infinity of \( E(|e|^p \exp(bf(e))) \) are 
\[
 b(\gamma_2 - \gamma_1) < \frac{1}{2^\nu} \quad \text{and} \quad b(\gamma_2 + \gamma_1) < \frac{1}{2^\nu} .
\]
That is \( b(\gamma_2 - \gamma_1) < \sqrt{2} \) and \( b(\gamma_2 + \gamma_1) < \sqrt{2} \). The conditions for the infinity of \( E(e^p \exp(bf(e))) \) are the same as those for \( E(|e|^p \exp(bf(e))) \) \( 0 < \nu \leq 1 \).

Finally, when \( \epsilon \sim \) Student-t with \( d \) degrees of freedom \( (d > 2) \) and is normalized to satisfy \( E(\epsilon) = 0, \text{var}(\epsilon) = 1 \), then

\[
 E(|\epsilon|^p \exp(bf(\epsilon))) = C_1 \left[ \int_0^{+\infty} e^p \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\epsilon) \left( 1 + \frac{\epsilon^2}{d-2} \right)^{-\frac{d+1}{2}} d\epsilon 
+ \int_0^{+\infty} e^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \left( 1 + \frac{\epsilon^2}{d-2} \right)^{-\frac{d+1}{2}} d\epsilon, \right] \quad \text{(B.6)}
\]

where \( C_1 = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{(d-2)\pi \Gamma\left(\frac{d}{2}\right)}} \). We can verify that \( E(|\epsilon|^p \exp(bf(\epsilon))) = \infty \) unless \( b(\gamma_2 - \gamma_1) \leq 0 \) and \( b(\gamma_2 + \gamma_1) \leq 0 \).

**Appendix B.2. Expectations needed to compute** \( E(|y_t|^c), \text{corr}(|y_t|^c, |y_{t+\tau}|^c) \) and \( \text{corr}(y_t, |y_{t+\tau}|^c) \) **when** \( \epsilon \sim N(0,1) \)

Assume that all the parameters are defined as in equations (1) and (2). When \( \epsilon \sim N(0,1) \), using the expression (B.2) and the formula 3.462-1 of Ryzhik et al. (2007), the following expressions for any positive integer \( p \) and any integer \( b \) are derived

\[
 E(|\epsilon|^p \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \left\{ \exp(b\alpha) \Gamma(p+1) \exp\left(\frac{b^2(\gamma_1 - \gamma_2)^2}{4}\right) D_{-p-1}(b(\gamma_1 - \gamma_2)) 
+ \Gamma(p+1) \exp\left(\frac{b^2(\gamma_1 + \gamma_2)^2}{4}\right) D_{-p-1}(-b(\gamma_1 + \gamma_2)) \right\} 
\]

(B.7)
and

\[
E(e^p \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \left\{ (-1)^p \exp(ba) \Gamma(p+1) \exp \left( \frac{b^2(\gamma_1 - \gamma_2)^2}{4} \right) D_{-p-1}(b(\gamma_1 - \gamma_2)) + \Gamma(p+1) \exp \left( \frac{b^2(\gamma_1 + \gamma_2)^2}{4} \right) D_{-p-1}(-b(\gamma_1 + \gamma_2)) \right\},
\]

(B.8)

where \( D_{-\alpha}(\cdot) \) is the parabolic cylinder function. Particularly, when \( p = 0, 1 \) or \( 2 \), the expressions are reduced to

\[
E(\exp(bf(\epsilon))) = \exp(ba) \exp(\bar{A}) \Phi(\bar{C}) + \exp(\bar{B}) \Phi(\bar{D}),
\]

\[
E(\epsilon \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \left\{ -\exp(ba) \left[ 1 + \sqrt{2\pi} \bar{C} \exp(\bar{A}) \Phi(\bar{C}) \right] + \left[ 1 + \sqrt{2\pi} \bar{D} \exp(\bar{B}) \Phi(\bar{D}) \right] \right\},
\]

\[
E(|\epsilon| \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \left\{ \exp(ba) \left[ 1 + \sqrt{2\pi} \bar{C} \exp(\bar{A}) \Phi(\bar{C}) \right] + \left[ 1 + \sqrt{2\pi} \bar{D} \exp(\bar{B}) \Phi(\bar{D}) \right] \right\}
\]

and

\[
E(|\epsilon|^2 \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \left\{ \exp(ba) \left[ \bar{C} + \sqrt{2\pi} (\bar{C}^2 + 1) \exp(\bar{A}) \Phi(\bar{C}) \right] + \left[ \bar{D} + \sqrt{2\pi} (\bar{D}^2 + 1) \exp(\bar{B}) \Phi(\bar{D}) \right] \right\},
\]

where \( \Phi(\cdot) \) is the Normal distribution function, \( \bar{A} = \frac{b^2(\gamma_1 - \gamma_2)^2}{2} \), \( \bar{B} = \frac{b^2(\gamma_1 + \gamma_2)^2}{2} \), \( \bar{C} = -b(\gamma_1 - \gamma_2) \) and \( \bar{D} = b(\gamma_1 + \gamma_2) \).

References


