Asymmetric Stochastic Volatility Models: Properties and Estimation

Xiuping Mao\textsuperscript{a}, Esther Ruiz\textsuperscript{a,b,*}, Helena Veiga\textsuperscript{a,b,c}, Veronika Czellar\textsuperscript{d}

\textsuperscript{a}Department of Statistics, Universidad Carlos III de Madrid, Spain.
\textsuperscript{b}Instituto Flores de Lemus, Universidad Carlos III de Madrid, Spain.
\textsuperscript{c}Financial Research Center/UNIDE, Portugal.
\textsuperscript{d}Department of Economics, Finance and Control, EMLYON Business School, France.

Abstract

In this paper, we derive the statistical properties of a general family of Stochastic Volatility (SV) models with leverage effect which capture the dynamic evolution of asymmetric volatility in financial returns. We provide analytical expressions of moments and autocorrelations of power-transformed absolute returns. Moreover, we analyze and compare the finite sample performance of two estimation procedures. The first method is an Approximate Bayesian Computation (ABC) filter-based Maximum Likelihood (ML), a technique similar to indirect inference as it requires simulation of pseudo-observations which are weighted according to their distance to the true observations. The second estimation method is a Markov Chain Monte Carlo (MCMC) estimator implemented in BUGS, a user-friendly and freely available software package that automatically calculates the full conditional posterior distribution. We show that the ABC filter-based ML estimator has better finite sample properties when estimating the parameters of a very general specification of the log-volatility with standardized returns following the Generalized Error Distribution (GED). The results are illustrated by analyzing a series of daily S&P500 returns.

Keywords: ABC filtering, BUGS, Leverage effect, SV models, Value-at-Risk.

JEL: C11, C51, C58

\textsuperscript{*}We received helpful comments from Enrique Sentana, George Tauchen, Jun Yu, Mike Wiper and participants at the 2014 Conference on Indirect Estimation Methods in Finance and Economics held in Konstanz and the Bayesian study group of the Department of Statistics at UC3M. We gratefully acknowledge the financial support from the Spanish Ministry of Education and Science, research project ECO2012-32401, and the computer support from EUROFIDAI. The third author is also grateful for project MTM2010-17323.

\textsuperscript{C/} Madrid, 126, 28903, Getafe, Madrid (Spain), Tel: +34 916249851, Fax: +34 916249849, Email: ortega@est-econ.uc3m.es. Corresponding Author.
1. Introduction

Stochastic Volatility (SV) models, originally proposed by Taylor (1982, 1986) and popularized by Harvey et al. (1994), specify the volatility of financial returns as a latent stochastic process. Asymmetric SV models incorporating the fact that volatility responds asymmetrically to past standardized returns of different signs (leverage effect) have been introduced by Taylor (1994) and Harvey and Shephard (1996). Even if normality of standardized returns is still a common assumption, SV models with heavy-tailed and asymmetric distributions such as symmetric or skewed Student-t, Generalized Error Distribution (GED) and mixtures of normals are increasingly popular (Liesenfeld and Jung, 2000; Chen et al., 2008; Wang et al., 2011; Tsotas, 2012).

Although the literature on asymmetric SV models is rather extensive, the statistical properties such as existence and analytical expressions of moments are known only in specific cases. For instance, in an asymmetric SV model with Gaussian standardized returns proposed by Harvey and Shephard (1996), Taylor (1994, 2007) provide expressions for the variance and kurtosis of returns and the autocorrelation function of squared observations. Ruiz and Veiga (2008b) and Pérez et al. (2009) derive the autocorrelation and cross-correlation functions of the power-transformed absolute returns of an asymmetric autorregressive long memory SV model with Gaussian standardized returns that nests the model proposed by Harvey and Shephard (1996). Asai and McAleer (2011) derive the first and second order moments of returns generated by their general specification of the volatility assuming normality of standardized returns. Eventually, Yu (2012a) also assumes normality to derive the moments of returns and the conditions for stationarity, strict stationarity and ergodicity of an asymmetric SV model with time-varying correlations.

However, the statistical properties are unknown in general asymmetric SV models in which the standardized returns are not normally distributed. The first objective of this paper is to fill this gap by deriving the properties of a very general family of asymmetric SV models with non-Gaussian

---


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 distribution when $\nu = \infty$. The GED distribution has heavy tails if $\nu < 2$.

4 See also Cappuccio et al. (2004), Asai (2008), Choy et al. (2008), Asai and McAleer (2011), Nakajima and Omori (2012) and Wang et al. (2013).
standardized returns, from now on referred to as Generalized Asymmetric SV (GASV) family. In particular, we provide general conditions for stationarity and for the existence of integer moments of returns and absolute returns. In addition, we derive analytical expressions of the marginal variance and kurtosis, the autocorrelations of power-transformed absolute returns and the cross-correlations between returns and future power-transformed absolute returns for a given general specification of the volatility with GED standardized returns.

Due to the intractability of the likelihood, parameter estimation in GASV models is a rather difficult task.\(^5\) SV models with Gaussian errors have been estimated via a variety of methods including the generalized method of moments (Taylor, 1986; Melino and Turnbull, 1990), indirect inference (Gourieroux et al., 1993; Monfardini, 1998; Dhaene, 2004), quasi-maximum likelihood (Harvey et al., 1994; Ruiz, 1994) and MCMC methods (Jacquier et al., 1994; Chib et al., 2002).\(^6\) However, to our knowledge, the estimation of a general specification of the GASV family with GED standardized returns has not been covered in the literature yet.

The second goal of this paper is to provide estimation methods for this general setting. We use an Approximate Bayesian Computation (ABC) filter-based Maximum Likelihood (ML) technique proposed by Calvet and Czellar (2014). This procedure shares similarities with indirect inference as it requires simulation of pseudo-observations which are weighted according to their distance to the true observations. The hidden Markov states are sequentially simulated and resampled using a Multinomial distribution. Due to the resampling step, the original ABC-filter estimated likelihood is highly discontinuous which challenges maximum likelihood estimation. We smoothen the ABC filtered likelihood by adopting Malik and Pitt (2011)’s continuous sampling method. This novel continuous ABC filter greatly facilitates estimation via ML in complex state-space models such as GASV models.

Our second estimation method is a MCMC procedure (Jacquier et al., 1994; Kim et al., 1998),

\(^5\)See, for example, the surveys on the estimation of SV models by Broto and Ruiz (2004) and Yu (2012b).
which is very popular to estimate parameters in SV models.\footnote{See Chib (2001) and Asai (2005) for good surveys and the contributions of Watanabe and Omori (2004); Omori et al. (2007); Omori and Watanabe (2008); Nakajima and Omori (2009); Abanto-Valle et al. (2010) and Tsiotas (2012).} In this paper, we use the MCMC estimator implemented in the user-friendly and freely available BUGS software described in 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 (2012a) and Wang et al. (2013). The MCMC estimator implemented in BUGS is appealing because it can handle non-Gaussian level disturbances without much programming effort.

We compare the finite sample properties of the ABC filter-based ML and MCMC estimators. We show that the MCMC estimator has difficulties while the ABC filter-based ML estimator is able to estimate all parameters of the general framework paying a price in terms of programming and computation time.

We also illustrate the performance of both estimation methods on a series of daily S&P500 returns observed between September 1998 and July 2014. We provide Value-at-Risk forecasts for the out-of-sample period of August-December 2014.

The rest of the paper is organized as follows. In Section 2, we establish notation by describing the general framework for SV models with leverage effect and derive its statistical properties. Section 3 presents the popular asymmetric SV models that are contained in the GASV family and derives analytical expressions of their moments. Section 4 describes the two alternative estimators and reports the finite sample comparison of these estimators in Monte Carlo simulations. Section 5 presents an empirical application to daily S&P500 returns. Finally, the main conclusions and some guidelines for future research are summarized in Section 6.

2. Statistical properties of GASV models

2.1. Model description

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

\[
y_t = \exp(h_t/2)\epsilon_t, \quad t = 1, \ldots, T
\]
\[ h_t - \mu = \phi (h_{t-1} - \mu) + f(\varepsilon_{t-1}) + \eta_{t-1}, \]  

where the log-volatility disturbance \( \eta_{t-1} \) is a Gaussian white noise with variance \( \sigma^2_{\eta} \) and \( f(\varepsilon_{t-1}) \) is any function of \( \varepsilon_{t-1} \) independent of \( \eta_{t-1} \) for all leads and lags. The scale parameter \( \mu \) determines the marginal variance of returns, while \( \phi \) controls the rate of decay towards zero of the autocorrelations of power-transformed absolute returns, hence, the persistence in volatility shocks. Note that, in equations (1) and (2), the standardized 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 standardized 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 general, it rules out models in which the persistence, \( \phi \), and/or the variance of the volatility noise, \( \sigma^2_{\eta} \), are time-varying.\(^9\) Finally, note that the only assumption made about the distribution of the level disturbance, \( \varepsilon_t \), is that it is an IID sequence with mean zero and variance one. As a consequence, \( \varepsilon_t \) is strictly stationary.

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.** Consider a GASV process \( y_t \) defined in equations (1) and (2). The process \( y_t \) is strictly stationary if \( |\phi| < 1 \). Furthermore, if \( \varepsilon_t \) follows a distribution such that both \( E(\exp(0.5cf(\varepsilon_t))) \) and \( E(|\varepsilon_t|^c) \) exist and are finite for some positive integer \( c \), then \( \{y_t\} \) and \( \{|y_t|\} \) have finite, time-invariant moments of order \( c \) given by

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

\(^8\)The normal assumption for \( \eta_t \) when \( f(\varepsilon_{t-1}) = 0 \) is justified in Andersen et al. (2001b) and Andersen et al. (2001a, 2003).

\(^9\)This excludes from the GASV family the time-varying specification of Yu (2012a).
and
\[ E(|y_t|^c) = \exp \left( \frac{c \mu}{2} \right) E(|\epsilon_t|^c) \exp \left( \frac{c^2 \sigma_\eta^2}{8(1-\phi^2)} \right) P(0.5c, \phi), \] (4)

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

**Proof.** See Appendix A.1. \( \Box \)

**Theorem 2.1** establishes the strict stationarity of \( y_t \) if \( |\phi| < 1 \); note that this is the same condition derived by Yu (2012a) in his time-varying correlation model. The existence of the expectation of \( y_t^2 \) is guaranteed if, in addition, \( E(\exp(f(\epsilon_t))) < \infty \). Consequently, under these two conditions, \( y_t \) is also weakly stationary.

Note that according to expression (3), if \( \epsilon_t \) has a symmetric distribution, then all odd moments of \( y_t \) are zero. Furthermore, from expression (4), 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 (4) as follows
\[ \sigma_y^2 = \exp \left( \mu + \frac{\sigma_\eta^2}{2(1-\phi^2)} \right) P(1, \phi). \] (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 (4) with \( c = 4 \) and \( c = 2 \) as follows
\[ \kappa_y = \kappa_\epsilon \exp \left( \frac{\sigma_\eta^2}{1-\phi^2} \right) \frac{P(2, \phi)}{(P(1, \phi))^2}, \] (6)

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

The kurtosis of the basic symmetric Autoregressive SV (ARSV) model considered by Taylor (1982, 1986) and Harvey et al. (1994) is given by \( \kappa_\epsilon \exp \left( \frac{\sigma_\eta^2}{1-\phi^2} \right) \). Therefore, this kurtosis is multiplied by the factor \( r = \frac{P(2, \phi)}{P(1, \phi)} \) in the GASV family.

Note that, the expression of \( E(|y_t|^c) \) in (4) depends on the distribution of \( \epsilon_t \) and \( f(\cdot) \). 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 specify $f(\epsilon_t)$. We will provide analytical expressions for some popular distributions and specifications in Section 3. Note that even in cases in which the function $f(\cdot)$ and/or the distribution of $\epsilon_t$ are such that they do not lead to obtain closed-form expressions for the moments, expression (4) can always be used to simulate them as far as they are finite.

2.3. Dynamic dependence

It is easy to see that returns generated from (1) and (2) 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 in equations (1) and (2) with $|\phi| < 1$. If $\epsilon_t$ follows a distribution such that $E(\exp(0.5cf(\epsilon_t))) < \infty$ and $E(|\epsilon_t|^c) < \infty$ for some positive integer $c$, then the $\tau$-th order autocorrelation of $|y_t|^c$ is finite and given by

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

where $T_n(a, b) \equiv \prod_{i=1}^{n-1} E(\exp(ab^{i-1}f(\epsilon_{t-i})))$ if $n > 1$ while $T_1(a, b) \equiv 1$.

**Proof.** See Appendix A.2. \qed

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

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 in 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

$$P_{c1}(\tau) = \frac{E(|\varepsilon_t|^c) \exp \left( \frac{2(1+c)}{(1-c)^2} \right) E(\varepsilon_t, \exp(0.5\phi(1-c) f(\varepsilon_t))) \exp(0.5(1+c)\phi^2) \sigma_t^2 \eta \frac{P(c,\phi)}{2\phi^2}}{\sqrt{E(|\varepsilon_t|^2) \exp \left( \frac{-2c\phi^2}{(1-c)^2} \right) P(c,\phi) - \left[ E(|\varepsilon_t|^c) \exp(0.5(1+c)\phi^2) \right]^2}}$$

(8)

Proof. See Appendix A.3.

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. Note that when referring to conditional variances, we refer to the variance of returns at time $t$ conditional on past returns observed up to and including time $t-1$. 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 (2012a) who proposes a function that relates the conditional variance to the lagged return innovation, $\varepsilon_{t-1}$, 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. The second attempt is due to Takahashi et al. (2013) who specify the news impact function for SV models in the spirit of Yu (2012a) as the volatility at time $t+1$ conditional on returns at time $t$. However, in order to obtain a U-shaped NIC, Takahashi et al. (2013) propose 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, $\text{Var}(y_t|y_1, \cdots, y_{t-1})$. 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, $\sigma_t^2$, 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 the conditional variance. We propose the Stochastic News Impact Surface (SNIS) as the surface that relates $\sigma^2_t$ to $\epsilon_{t-1}$ 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

$$\text{SNIS}_t = \exp((1 - \phi)\mu \sigma^2_y) \exp(f(\epsilon_{t-1}) + \eta_{t-1}). \quad (9)$$

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 $\eta_{t-1}$. It is important to point out that, given that $\epsilon_t$ and $\eta_t$ are mutually independent, the relation between $h_t$ and $\epsilon_t$ is measured just by the covariance between $f(\epsilon_t)$ and $\epsilon_t$. However, when looking at the more interesting relation between the volatility, $\sigma_t$, and $\epsilon_t$, we observe that the leverage effect depends on $\eta_{t-1}$.

In order to illustrate the SNIS, we consider the following specification of $f(\cdot)$

$$f(\epsilon_t) = \alpha I(\epsilon_t < 0) + \gamma_1 \epsilon_t + \gamma_2 (|\epsilon_t| - E|\epsilon_t|), \quad (10)$$

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); see Asai et al. (2012) for this specification allowing for long-memory. The T-GASV model nests several popular models previously proposed in the literature to represent asymmetric volatility 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. When $\gamma_2 = 0$, the specification is very similar to the asymmetric SV model proposed by Asai and McAleer (2006). Finally, when $\gamma_1 = \gamma_2 = 0$, 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 function, it is possible to allow the leverage
effect to be different depending on the size of $\varepsilon_t$.

The top panel of Figure 2 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 $\exp((1 - \phi)\mu)\sigma^2_\gamma = 1$. These parameter values are chosen to resemble those often obtained when asymmetric SV models are fitted to real financial data. We can observe that the SNIS shows a discontinuity due to the presence of the indicator function in (10). The most important feature of the SNIS plotted in Figure 2 is that it illustrates that the leverage effect of SV models is different depending on the values of the volatility shock. It is very clear when the volatility shock is positive. However, 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 2. 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 A-ARSV model of Harvey and Shephard (1996). For this purpose, we consider the same parameters as above with $\alpha = \gamma_2 = 0$ and plot the corresponding SNIS in the second panel of Figure 2. In this case, we observe that the leverage effect is very weak for negative log-volatility shocks.

Finally, in the bottom panel of Figure 2, we illustrate the shape of the SNIS of the E-SV model with $\alpha = 0$. In this case, we can observe that there is no discontinuity but the effect of $\varepsilon_{t-1}$ 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 $\eta_t = 0$.

3. Alternative Asymmetric SV models

Appropriate choices of the function $f(\cdot)$ and of the distribution of $\varepsilon_t$ allow us to obtain closed-form expressions of the moments of returns. In this section, we derive and illustrate these moments for the T-GASV model when $\varepsilon_t$ 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.

Consider the T-GASV model defined in equations (1), (2) and (10) with $\varepsilon_t \sim GED(\nu)$. If $\nu > 1$, then the conditions in Theorem 2.1 are satisfied and a closed-form expression of $E(|\gamma_t|^c)$ 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

$$P(1, \phi) = \prod_{i=1}^{\infty} \left\{ \sum_{k=0}^{\infty} \left( \frac{\Gamma(1/\nu)}{\Gamma(3/\nu)} \right)^{k/2} \frac{\Gamma((k+1)/\nu)}{2\Gamma(1/\nu)k!} \phi^{k(i-1)} \left[ (\gamma_1 + \gamma_2)^k + \exp(\alpha \phi^{i-1})(\gamma_2 - \gamma_1)^k \right] \right\},$$

(11)

where $\Gamma(\cdot)$ is the Gamma function. Note that in order to compute $P(\cdot, \cdot)$, 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. We can calculate the kurtosis in (6) in a similar way.

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

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

(12)

where $\Phi(\cdot)$ is the Gaussian cumulative distribution function. The expression of the marginal variance derived by Taylor (1994) for the classical asymmetric SV model of Harvey and Shephard (1996), by Ruiz and Veiga (2008a) for the asymmetric long-memory SV model and by Asai and McAleer (2011) for their general specification can be obtained as particular cases of expression (5) with $P(1, \phi)$ defined as in (12).

When $\nu < 1$, we cannot obtain analytical expressions of $E(|y_t|^c)$. However, in Appendix B.1, we show that, in this case, $E(|y_t|^c)$ in equation (4) 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 (4) are $\gamma_2 + \gamma_1 < 2\sqrt{c}/c$ and $\gamma_2 - \gamma_1 < 2\sqrt{c}/c$.

The expectations needed to obtain closed-form expressions of the autocorrelations in expression

10These expectations are derived by Demos (2002) when $p = 0$ or 1 and $\alpha = 0$.
11The 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.
(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 Gaussian 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. 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 returns, namely, $\rho_{21}(1)$, is plotted in the first panel of Figure 3 as a function of the leverage parameters, $\gamma_1$ and $\alpha$. We observe that the autocorrelations of squared returns are larger for larger values of $\gamma_2$. However, both surfaces are rather flat and, consequently, the leverage parameters do not have large effects on the first order autocorrelations of squares.

In the second panel of Figure 3, we illustrate the effect of the parameters on the cross-correlations between $y_t$ and $y^2_{t+1}$ denoted by $\rho_{21}(1)$. Observe that the first order cross-correlations between returns and future squared returns are indistinguishable for the two values of $\gamma_2$ considered in Figure 3. $|\gamma_1|$ drags $\rho_{21}(1)$ in an approximately linear way while the effect of $\alpha$ is non-linear.

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. For this purpose, the first column of Figure 4 illustrates the acf of squared returns (top panel) and the ccf between returns and future squared returns (bottom panel) in a T-GASV model 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 acf of $y^2_t$ has an exponential decay. Furthermore, fatter tails of $\epsilon_t$ imply smaller autocorrelations of squared returns; see Carnero et al. (2004) for similar conclusions in the context of symmetric SV models. The ccf plotted in the bottom panel of the first column shows that the parameter $\nu$ of the GED distribution has a very mild influence on the cross-correlations.

The second column in Figure 4 illustrates the acfs and ccfs of the A-ARSV model with the same parameters as above and $\alpha = \gamma_2 = 0$. We can observe that the autocorrelations of squared returns and the absolute cross-correlations are slightly smaller than those of the corresponding T-GASV model. Therefore, including $\gamma_2$ and $\alpha$ in the T-GASV model allows for stronger volatility clustering.
and leverage effect.

The restriction $\alpha = 0$ corresponds to the E-SV model. The parameter $\gamma_2$ measures the dependence of $h_t$ on past absolute return disturbances in the same form as in the EGARCH model. The third column of Figure 4 plots the autocorrelations and cross-correlations for an E-SV model with the same parameter values as 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 $|\varepsilon_{t-1}|$ into the A-ARSV model generates larger autocorrelations of squares returns. 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 shows that, given a particular E-SV model, we may find an A-ARSV model with almost the same autocorrelations and cross-correlations. For instance, the autocorrelations in the E-SV model with $\nu = 2$ are very similar to those in the A-ARSV model with $\nu = 2.5$. Furthermore, the cross-correlations are indistinguishable in any case. Nevertheless, these two models generate returns with different kurtoses. Therefore, if the parameter $\nu$ is free, it could be difficult identify the parameters $\gamma_2$ and $\sigma^2_\eta$ 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.

The autocorrelations in the T-GASV and E-SV models are almost identical. Including $\alpha$ only has a paltry effect on volatility clustering. 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 $f(\varepsilon_t) = \alpha I(\varepsilon_t < 0)$. 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 to the more general Threshold SV model. Hereafter, we refer to this model as restricted Threshold SV (RT-SV). The last column of Figure 4 illustrates the shape of the autocorrelations of squared returns and the cross-correlations between returns and future squared returns for a RT-SV model with the same values of the parameters.
as those considered above and $\gamma_1 = \gamma_2 = 0$. Comparing the autocorrelations of squared and absolute returns in the T-GASV and the RT-SV models represented in the top panel, we observe that the latter are slightly smaller than the former. However, the absolute cross-correlations in the RT-SV model are the smallest among all the models considered. The presence of $\alpha$ in the T-GASV model seems to reinforce the role of the leverage parameter $\gamma_1$.

4. Model estimation

As shown in the previous section, GASV models are attractive because of their flexibility to mimic the dynamic properties of financial return data. The downside to adopting an appealing GASV model is the difficulty involved with parameter estimation and volatility forecasts. The intractability of the likelihood is inherited from standard SV models, see Broto and Ruiz (2004) and Yu (2012b) for surveys on estimation procedures for SV models.

In this paper, we consider two estimation methods for GASV models. The first one is an ABC filter-based ML procedure proposed by Calvet and Czellar (2014), which closely relates to indirect inference in the sense that it estimates the conditional distribution of the hidden-state, by a set of particles which are able to generate pseudo-observations close to the real observations. The second is the MCMC estimator described by Meyer and Yu (2000) who propose to estimate the A-ARSV model using the user-friendly and freely available BUGS software. This estimator is attractive because it reduces the computational effort. Furthermore, note that Bayesian MCMC does not rely on asymptotic approximations to conduct inference.

4.1. ABC filter-based maximum likelihood estimation

Consider a set of observations available at date $t$, denoted by $y_{1:t} = (y_1, \ldots, y_t)$ and assume that they have been generated from the T-GASV model described in equations (1), (2) and (10). The model can be written in state space form via the following hidden Markov state process:

\[
x_t = \begin{pmatrix} x_{t,1} \\ x_{t,2} \end{pmatrix} \equiv \begin{pmatrix} h_{t+1} \\ h_t \end{pmatrix}, \quad \text{with} \quad \begin{cases} h_{t+1} = \mu + \phi (h_t - \mu) + f(\epsilon_t) + \eta_t \\ f(\epsilon_t) = \alpha I(\epsilon_t < 0) + \gamma_1 \epsilon_t + \gamma_2 (|\epsilon_t| - E(|\epsilon_t|)) \\ \epsilon_t \sim IID(0,1) \text{ and } \eta_t \sim N(0, \sigma^2) \end{cases}
\]
The conditional density of \( y_t = \exp(h_t/2)e_t \) then depends on \( x_{t,2} \) via \( h_t \) and on \( x_{t,1} \) via \( e_t \). Denote the transition kernel by \( A(x_{t}|x_{t-1}) \) and define the observation density as the density of \( y_t \) conditional on the hidden state

\[
y_t|x_t \sim f_{Y|X}(y_t|x_t). \tag{14}
\]

The observation density is not available in closed form, as \( x_{t,1} \) is partially revealing for \( e_t \). An exception is the symmetric SV model with Gaussian errors. The availability of the observation density is a necessary condition for standard particle filter-based methods (Gordon et al., 1993) in which the importance weight of a particle is provided by the observation density at the given particle. This makes particle filter-based ML estimation challenging. However, sampling from the joint distribution:

\[
(x_t, y_t)|(x_{t-1}, y_{1:t-1}) \sim f_{X,Y}(x_t, y_t|x_{t-1}, y_{1:t-1}) = f_Y(y_t|x_t, y_{1:t-1})A(x_t|x_{t-1}), \tag{15}
\]

is straightforward.\footnote{Equality in (15) is due to the fact that under the T-GASV model the density of \( x_t \) conditional on \((x_{t-1}, y_{1:t-1})\) coincides with the transition kernel \( A(x_{t}|x_{t-1}) \), i.e. \( y_{1:t-1} \) does not provide any additional information for \( x_t \) than does \( x_{t-1} \).}

Hence, one can use the accurate ABC particle filter of Calvet and Czellar (2014) for ML estimation and forecasting purposes.

The idea of ABC filtering proposed by Calvet and Czellar (2011) and Jasra et al. (2012) closely relates to indirect inference in the sense that it estimates the conditional distribution of the hidden-state, \( \lambda(x_t|y_{1:t}) \), by a set of particles \( \{\tilde{x}_t^{(n)}\}_{n=1}^N \) which are able to generate pseudo-observations \( \{\tilde{y}_t^{(n)}\}_{n=1}^N \) close to the real observation \( y_t \). The distance is measured by a strictly positive kernel \( K: \mathbb{R} \to \mathbb{R} \) integrating to unity. For any \( h_t > 0 \), define \( K_{h_t}(y) = K\left(\frac{y}{h_t}\right)/h_t \). The recursive method is as follows. Assume that at date \( t = 0 \), particles \( \{x_0^{(n)}\}_{n=1}^N \) are generated from a prior density \( \lambda_0 \). Then, for \( t = 1, \ldots, T \) the algorithm below, which is reproduced from Calvet and Czellar (2011, 2014) and is a variant of Jasra et al. (2012)'s method.\footnote{Jasra et al. (2012) consider a uniform kernel \( K_{h_t}(y) = I(\|y\| < h_t) \) which is not strictly positive.}

**ABC Particle Filter:**

**Step 1** (State-observation sampling): For every \( n = 1, \ldots, N \), simulate a state-observation pair \((\tilde{x}_t^{(n)}, \tilde{y}_t^{(n)})\) from \( f_{X,Y}(\cdot | x_{t-1}^{(n)}, y_{1:t-1}) \).
Step 2 (Correction): Given the new data point $y_t$, compute

$$w_t^{(n)} = K_{h_t}(y_t - \tilde{y}_t^{(n)}), \quad n \in \{1, \ldots, N\}.$$ 

Step 3 (Selection): For every $n = 1, \ldots, N$, draw $x_t^{(n)}$ from $\tilde{x}_t^{(1)}, \ldots, \tilde{x}_t^{(N)}$ with importance weights $p_t^{(1)}, \ldots, p_t^{(N)}$ and $p_t^{(n)} = w_t^{(n)}/\sum_{n'=1}^{N} w_t^{(n')}$. 

An accurate ABC filter is based on bandwidth and kernel conditions which guarantee that the kernel estimator

$$\hat{f}_Y(y_t|y_{1:t-1}) = \frac{1}{N} \sum_{n=1}^{N} w_t^{(n)},$$

converges in mean square error to the conditional density of $y_t$ given past observations; see Calvet and Czellar (2014) for kernel and bandwidth conditions and mean square convergence result.

Hence, a consistent estimator of the loglikelihood is

$$\frac{1}{T} \sum_{t=1}^{T} \log \hat{f}_Y(y_t|y_{1:t-1})$$

and a simulated ML estimation is thus enabled by the accurate ABC filter. When the filter size $N$ goes to infinity, the simulated ML is equivalent to the ML estimator. As suggested by Calvet and Czellar (2014), we use a quasi-Cauchy kernel and the optimal bandwidth defined as

$$K(u) = (1 + \pi^2 u^2 / 4)^{-2} \quad \text{and} \quad h_t^* = \hat{\sigma}_t \left( \frac{5 \pi^9/2}{48N} \right)^{1/5}$$

where $\hat{\sigma}_t = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (\tilde{y}_t^{(n)} - \overline{\tilde{y}}_t)^2}$ and $\overline{\tilde{y}}_t = N^{-1} \sum_{n=1}^{N} \tilde{y}_t^{(n)}$.

The only numerical challenge is that even when fixing the seed when evaluating the loglikelihood in (17) leads to a discontinuous function due to the multinomial resampling in Step 3 of the ABC filter. Augmenting the number of particles is a way to smoothen the objective function and alleviate the problems encountered in the minimization process. In this paper, we adopt an alternative solution and use the continuous resampling by Malik and Pitt (2011) for standard particle filters. Despite the fact that our hidden state is bivariate, we can use a one dimensional resampling. Indeed, when ABC-filtering in model (13), $\{\tilde{x}_{t,2}^{(n)}\}$ are necessary for the simulation of the pseudo-observations.
\( \{ \tilde{y}^{(n)}_t \} \) in Step 2, but \( \{ \tilde{x}^{(n)}_{t,2} \} \) are ignored when generating \( \{ \tilde{x}^{(n)}_{t,1} \} \) in step 1. Hence we can base our resampling Step 3 on the first components \( \{ \tilde{x}^{(n)}_{t,1} \} \) solely. The continuous resampling algorithm by Malik and Pitt (2011) along with Malmquist (1950) ordered uniforms is summarized in Appendix C.

The resulting ABC log-likelihood is continuous in the parameters as is illustrated in Figure 1 displaying the standard and continuous ABC filtered loglikelihoods in the case of a T-GASV model for \( T = 500 \) and \( N = 200,000 \).\(^{14}\)

4.2. MCMC estimator

Next, we describe briefly the MCMC estimator implemented in BUGS. Let \( p(\theta) \) be the joint prior distribution of the unknown parameters \( \theta = \{ \mu, \phi, \alpha, \gamma_1, \gamma_2, \sigma^2_\eta, \nu \} \). Following Meyer and Yu (2000), the prior densities of \( \phi \) and \( \sigma^2_\eta \) are \( \phi = 2\phi^* - 1 \) with \( \phi^* \sim \text{Beta}(20, 1.5) \) and \( \sigma^2_\eta = 1/\tau^2 \) with \( \tau \sim \text{IG}(2.5, 0.025) \), respectively, where \( \text{IG}(\cdot, \cdot) \) is the inverse Gaussian distribution.\(^{15}\) The remaining prior densities are chosen to be uninformative, that is, \( \mu \sim \text{N}(0, 10) \), \( \alpha \sim \text{N}(0.05, 10) \), \( \gamma_1 \sim \text{N}(-0.05, 10) \), \( \gamma_2 \sim \text{N}(0.05, 10) \) and \( \nu \sim \text{U}(0, 4) \). These priors are assumed to be independent. The joint prior density of \( \theta \) and \( h \) is

\[
p(\theta, h) = p(\theta)p(h_0) \prod_{t=1}^{T+1} p(h_t|h_{t-1}, \theta). \tag{19}\]

The likelihood function is then given by

\[
p(y|\theta, h) = \prod_{t=1}^{T} p(y_t|h_t, \theta). \tag{20}\]

Note that the conditional distribution of \( y_t \) given \( h_t \) and \( \theta \) is \( y_t|h_t, \theta \sim \text{GED}(\nu) \). 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

\[
y_t|u, h_t \sim U \left( \frac{\exp(h_t/2)}{\sqrt{2\Gamma(3/\nu)}/\Gamma(1/\nu)} u^{1/\nu}, \frac{\exp(h_t/2)}{\sqrt{2\Gamma(3/\nu)}/\Gamma(1/\nu)} u^{1/\nu} \right), \tag{21}\]

\(^{14}\)The random number generators for a fixed seed need to provide continuously changing numbers in the parameters.

\(^{15}\)Although the prior of \( \phi^* \) is very informative, when it is changed to \( \text{Beta}(1, 1) \), the results are very similar.
where \( u|\nu \sim \text{Gamma}(1 + 1/\nu, 2^{-\nu/2}) \). 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:

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

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

\[
p(\theta, h|y) \propto p(\theta)p(h_0) \prod_{t=1}^{T+1} p(h_t | h_{t-1}, y, \theta) \prod_{t=1}^{T} p(y_t | h_t, \theta). \tag{22}
\]

4.3. Finite sample performance of the parameter estimators

In this section, we carry out extensive Monte Carlo experiments to analyze the finite sample performance of the ABC filter-based ML and MCMC estimators when estimating the parameters of the T-GASV model. We consider two designs for the Monte Carlo experiments. First, we treat \( \nu \) as known and estimate all other parameters in the model. In this case, \( R = 500 \) replicates are generated by the T-GASV model with normal errors, parameters \((\mu, \phi, \alpha, \gamma_1, \gamma_2, \sigma_n^2) = (0, 0.98, 0.07, -0.08, 0.1, 0.05)\) and sample sizes \( T = 500, 1000 \) and 2000. Second, \( R \) replicates are generated by the T-GASV model with GED standardized returns with \( \nu = 1.5 \). In this case, the parameter \( \nu \) is also estimated. For computational efficiency, we choose to estimate \( \mu \cdot (1 - \phi) \) instead of \( \mu \) in our estimation procedures. The top panel of Table 1 reports the Monte Carlo averages of the estimates and the asymptotic standard deviations of the ABC filter-based ML estimates. The lower panel of Table 1 reports the Monte Carlo averages and standard deviations of the posterior means of the MCMC estimates of the parameters.
Consider first the results obtained with the ABC filter-based ML estimator. We can observe that regardless of whether the distribution of the standardized returns is assumed to be known, or not, all the parameters are accurately estimated, even when the sample size is as small as $T = 500$. The Monte Carlo averages of the asymptotic standard deviations of the parameter estimates decay approximately at rate $\sqrt{T}$. These Monte Carlo experiments show that the ABC filter-based ML estimator does a remarkably good job in estimating the parameters of the general asymmetric SV model considered in this paper.

Consider now the results regarding the MCMC estimator which are based on a total number of 20,000 iterations after a burn-in of 10,000.\(^{16}\) When the distribution of the standardized returns is known, we can observe that the Monte Carlo averages of the posterior means are, in general, further from the true parameters than are the ABC filter-based ML estimates. Standard deviations are also larger. Obviously, as the sample size increases the differences between the ABC filter-based ML and MCMC estimators disappear. However, even when $T = 2000$ the Monte Carlo standard errors of the MCMC estimator are still much larger than those of the ABC filter-based ML estimator mainly for the constant, $\mu \cdot (1 - \phi)$, and the threshold, $\alpha$, parameters. When the data are generated with GED standardized returns (right panel), the Monte Carlo averages of the estimates of $\mu \cdot (1 - \phi)$, $\phi$, $\alpha$ and $\gamma_1$ are very similar to those obtained with known $\nu = 2$ (left panel). However, the estimates of $\gamma_2$, $\sigma^2_\eta$ and $\nu$ suffer from biases which do not disappear with the sample size. Moreover, the standard deviations do not decrease with the sample size as expected. Both $\sigma^2_\eta$ and $\nu$ are underestimated while $\gamma_2$ is overestimated. The correlation between the estimates of $\gamma_2$ and $\sigma^2_\eta$ is almost -0.7 while the correlation between the estimates of $\nu$ and $\gamma_2$ is as high as -0.8. When compared with the standard deviations of the ABC filter-based ML estimator, the Monte Carlo standard deviations of the posterior means of the MCMC estimator are approximately three times larger when $T = 2000$. So the MCMC estimator is less successful in estimating the parameters $\gamma_2$, $\sigma^2_\eta$ and $\nu$.\(^{17}\) However, the MCMC estimator has a significant computational advantage. While an ABC filter-based ML

\(^{16}\)The random number generators for a fixed seed need to provide continuously changing numbers in the parameters. For instance, the inverse cdf method satisfies this property.

\(^{17}\)Skaug and Yu (2014) use 100000 draws while Omori et al. (2007) show that even few hundred thousand draws cannot be enough. For some particular cases, we increase the number of iterations to 75000 obtaining very similar results.
estimation of a T-GASV model with $T = 500$ takes on average 2438 minutes to compile\textsuperscript{18}, the MCMC took only 11 minutes\textsuperscript{19}. Hence, computation time is the price to pay for the accuracy gain obtained via ABC filter-based ML estimator.

5. Empirical application

5.1. Data description and estimation results

In this section, we fit the T-GASV model to a series of daily S&P500 returns observed between September 1, 1998 and December 17, 2014. We split the observation period into an in-sample and an out-of-sample periods. The in-sample period ends on July 25, 2014 and contains $T = 4000$ observations. The out-of-sample period starts on July 28, 2014 and represents $H = 101$ observations. The returns are computed as $y_t = 100 \times \Delta \log P_t$, where $P_t$ is the adjusted close price at time $t$ obtained from finance.yahoo.com. The adjusted close prices together with their corresponding returns are plotted in Figure 5 for the in-sample period which suggests the presence of volatility clustering with episodes of large volatility associated with periods of negative movements in prices. Furthermore, this association between large volatility and negative returns is observed in the negative sample cross-correlations between returns and future squared returns also plotted in Figure 5.

We estimate the T-GASV model with GED standardized returns using both ABC filter-based ML and MCMC estimators. Table 2 reports the estimates obtained using the observations corresponding to the in-sample period. For the ABC filter-based ML estimates, we report asymptotic standard errors in parentheses. Regarding the MCMC estimation, we report the posterior mean and the posterior standard deviation of the MCMC estimator of each parameter. The plug-in ABC-filter estimated loglikelihoods are also reported. Heteroskedasticity and autocorrelation consistent Vuong test comparing the MCMC specification to the ABC filtered specification are reported in parentheses below the log-likelihood value. The ABC filter-based loglikelihood is smaller for MCMC but the difference is not significant at the 5\% level according to the Vuong test. The estimates of $\nu$ are

\textsuperscript{18}The computation was performed on an Intel(R) Xeon(R) CPU E5-2630 0 @ 2.30GHz.
\textsuperscript{19}The computation was performed on an Intel(R) Xeon(R) CPU E5-2665 0 @ 2.40GHz.
less than the value two for both estimation methods which indicates that the standardized returns follow a fat-tailed distribution.

When using the ABC filter-based ML estimator, all parameters including the threshold are statistically significant. Furthermore, except for $\phi$, the ABC filter-based ML and MCMC estimates are rather different with the asymptotic standard deviations of the ABC filter-based estimates being smaller than the posterior standard deviations of the MCMC estimates. In concordance with the Monte Carlo results, the MCMC estimate of $\gamma_2$, seems to be overestimated while $\sigma_\eta^2$ and $\nu$ are underestimated.

Finally, Figure 5 plots the plug-in moments implied by the two estimators together with the corresponding sample moments. The plug-in moments given by the ABC filter-based ML estimator are slightly closer to the sample moments comparing with those of the corresponding MCMC estimator.

5.2. Forecasting performance

In this subsection, we perform an out-of-sample comparison of the ability of the two alternative estimation methods considered in this paper to calculate the one-step-ahead VaR of the daily S&P500 returns using the T-GASV model with GED standardized returns. Given the extremely heavy computations involved in the estimation of the ABC filter-based ML we use estimates provided in Table 2 for the VaR forecasts in the entire out-of-sample period. We compute one-day-ahead VaR forecasts from July 28, 2014, to December 17, 2014 as

$$VaR^\alpha_{t+1} = -\hat{q}_{t,\alpha},$$

with $\alpha = 1\%, 5\%$ and $10\%$ and $\hat{q}_{t,\alpha}$ are the quantiles of the distribution of $y_{t+1}|y_{1:t}$, which is conveniently estimated by the ABC filter. We provide in total 101 one-step-ahead VaR forecasts.

In order to evaluate the adequacy of the VaR forecasts with both estimation methods, Table 2 reports the failure rates of the 101 VaR forecasts. The failure rates of the T-GASV model with GED standardized returns are quite similar across the estimation methods and are not significantly different from their theoretical values at the 5\% significance level. For the 5\% VaR, the empirical failure rate of the MCMC estimator is closer to the theoretical value. Figure 6 plots the negative
of the VaR forecasts computed with the two estimators for the out-of-sample period. For 5%, the MCMC forecasts are slightly higher than those computed with the ABC filter-based ML estimator for the last part of the out-of-sample period, but for 1% and 10% significance levels the VaR forecasts are quite similar. This indicates that the out-of-sample performance of the model is not affected by the in-sample estimation biases obtained when using the MCMC estimator.

6. Conclusions and further research

In this paper, we derive the statistical properties of a general family of asymmetric SV models denoted 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 standardized returns 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 in the GASV family. Closed-form expressions of the statistical properties of these models are obtained when the errors are GED. We also show that the response of volatility to positive and negative return shocks is different depending on the size and sign of the volatility shocks.

Moreover, we analyze the finite sample properties of two estimators of the parameters. The first is an ABC filter-based ML estimator closely related to indirect inference and the second, the MCMC estimator implemented in the BUGS software. We show that estimating a general specification belonging to the GASV family with the ABC filter-based ML estimator leads to accurate estimates of the parameters even when the sample size is small. The same does not happen with the MCMC estimator when the standardized returns follow a GED distribution. In this case, we find parameter estimates biases. However, using BUGS reduces substantially the coding effort and the MCMC estimator is significantly less computational and time-consuming than the ABC filter-based ML estimator.

Finally, a general SV specification is fitted to estimate the volatilities of S&P500 daily returns. For this particular data set, when estimating the VaRs with the two estimation methods, the failure
rates are similar which indicates that the out-of-sample performance of the model is not affected by the estimation in-sample biases of the MCMC estimator.

Several possible extensions of this paper could be of interest for future research. 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 one is closer to the empirical properties often observed when dealing with financial returns. Comparing the properties of the GASV models with those of the best candidates within the GARCH family is left for further research. Finally, analyzing the behavior of the estimation procedure proposed by Skaug and Yu (2014) and the Empirical Characteristic Function (ECF) estimator of Knight et al. (2002) in the context of the general asymmetric SV models considered in this paper is also in our research agenda.
Figure 1: Loglikelihood via ABC Filtering for T-GASV with Normal Standardized Returns. This figure illustrates the loglikelihood functions estimated by the standard (black line) and continuous (red line) accurate ABC filters for a simulated sample from a T-GASV with Normal standardized returns with $T = 500$ and $N = 200000$ particles. The true parameter set is $\mu = 0$, $\phi = 0.98$, $\alpha = 0.07$, $\gamma_1 = -0.08$, $\gamma_2 = 0.1$ and $\sigma_\eta^2 = 0.05$. 

$T$-GASV with $T=500$

![Loglikelihood via ABC Filtering](image)

Table and Figures
Figure 2: SNIS of different GASV models with $\phi = 0.98$, $\sigma^2 = 0.05$ and $\exp((1 - \phi)\mu)\sigma^2 = 1$. Top panel $f(\varepsilon) = \alpha I(\varepsilon < 0) + \gamma_1 \varepsilon + \gamma_2 (|\varepsilon| - E(|\varepsilon|))$; middle panel $f(\varepsilon) = \gamma_1 \varepsilon$; and bottom panel $f(\varepsilon) = \gamma_1 \varepsilon + \gamma_2 (|\varepsilon| - E(|\varepsilon|))$. The parameter values are $\{\alpha, \gamma_1, \gamma_2\} = \{0.07, -0.08, 0.1\}$. 
Figure 3: First order autocorrelations of squares (left panel) and first order cross-correlations between returns and future squared returns (right panel) of different Gaussian T-GASV models with parameters $\phi = 0.98$ and $\sigma^2_\eta = 0.05$.

Figure 4: First forty orders of the autocorrelations of squares (first row) and cross-correlations between returns and future squared returns (second row) for different specifications of asymmetric SV models. The first column corresponds to a T-GASV model with $\alpha = 0.07, \phi = 0.98, \sigma^2_\eta = 0.05, \gamma_1 = -0.08$, $\gamma_2 = 0.1$ and $\nu = 1.5$ (solid lines), $\nu = 1.7$ (dashed lines), $\nu = 2$ (dotted lines) and $\nu = 2.5$ (dashdot lines). The second column corresponds to the A-ARSV with $\alpha = \gamma_2 = 0$. The third column matches along with the E-SV model with $\alpha = 0$. Finally, the last column corresponds to the RT-SV model with $\gamma_1 = \gamma_2 = 0$. 

26
Table 1

Monte Carlo results of the ABC filter-based Maximum Likelihood and MCMC estimators of the T-GASV parameters. Regarding the ABC estimator, values of the Monte Carlo averages and asymptotic standard deviations (in parenthesis) are reported and for the MCMC, values of the Monte Carlo averages and standard deviations (in parenthesis) of the posterior means are reported.

<table>
<thead>
<tr>
<th></th>
<th>T-GASV $\epsilon_t \sim N(0, 1)$</th>
<th>T-GASV $\epsilon_t \sim GED(0, 1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu \cdot (1 - \phi)$ $\phi$ $\alpha$ $\gamma_1$ $\gamma_2$ $\sigma^2_\eta$</td>
<td>$\mu \cdot (1 - \phi)$ $\phi$ $\alpha$ $\gamma_1$ $\gamma_2$ $\sigma^2_\eta$ $\nu$</td>
</tr>
<tr>
<td><strong>True</strong></td>
<td>0 0.98 0.07 -0.08 0.1 0.05</td>
<td>0 0.98 0.07 -0.08 0.1 0.05 1.5</td>
</tr>
<tr>
<td><strong>ABC</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T=500</strong></td>
<td>Mean 0.012 0.973 0.069 -0.087 0.111 0.045</td>
<td>Mean 0.004 0.975 0.073 -0.085 0.115 0.045 1.501</td>
</tr>
<tr>
<td></td>
<td>Std.dev (0.044) (0.015) (0.076) (0.053) (0.103) (0.024)</td>
<td>Std.dev (0.032) (0.010) (0.057) (0.046) (0.095) (0.025) (0.127)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T=1000</strong></td>
<td>Mean 0.004 0.977 0.072 -0.082 0.107 0.048</td>
<td>Mean 0.001 0.978 0.072 -0.081 0.108 0.047 1.495</td>
</tr>
<tr>
<td></td>
<td>Std.dev (0.023) (0.008) (0.042) (0.030) (0.074) (0.016)</td>
<td>Std.dev (0.017) (0.006) (0.028) (0.029) (0.070) (0.016) (0.081)</td>
</tr>
<tr>
<td><strong>T=2000</strong></td>
<td>Mean 0.002 0.978 0.071 -0.082 0.098 0.049</td>
<td>Mean 0.002 0.979 0.069 -0.083 0.101 0.049 1.492</td>
</tr>
<tr>
<td></td>
<td>Std.dev (0.013) (0.005) (0.023) (0.020) (0.052) (0.011)</td>
<td>Std.dev (0.011) (0.004) (0.017) (0.020) (0.055) (0.013) (0.066)</td>
</tr>
<tr>
<td><strong>MCMC</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T=500</strong></td>
<td>Mean 0.015 0.965 0.096 -0.074 0.145 0.041</td>
<td>Mean 0.014 0.961 0.095 -0.075 0.184 0.024 1.462</td>
</tr>
<tr>
<td></td>
<td>Std.dev (0.065) (0.019) (0.113) (0.061) (0.168) (0.026)</td>
<td>Std.dev (0.062) (0.023) (0.110) (0.064) (0.234) (0.018) (0.323)</td>
</tr>
<tr>
<td><strong>T=1000</strong></td>
<td>Mean 0.004 0.974 0.083 -0.077 0.139 0.045</td>
<td>Mean 0.002 0.973 0.082 -0.077 0.181 0.027 1.427</td>
</tr>
<tr>
<td></td>
<td>Std.dev (0.040) (0.010) (0.076) (0.040) (0.106) (0.019)</td>
<td>Std.dev (0.040) (0.011) (0.076) (0.042) (0.181) (0.017) (0.216)</td>
</tr>
<tr>
<td><strong>T=2000</strong></td>
<td>Mean 0.000 0.977 0.078 -0.078 0.119 0.048</td>
<td>Mean 0.000 0.973 0.086 -0.074 0.210 0.023 1.390</td>
</tr>
<tr>
<td></td>
<td>Std.dev (0.029) (0.006) (0.055) (0.028) (0.064) (0.011)</td>
<td>Std.dev (0.040) (0.011) (0.078) (0.044) (0.152) (0.016) (0.209)</td>
</tr>
</tbody>
</table>
Table 2

Parameter estimates (left panel) of the T-GASV model using S&P 500 returns between September 9, 1998 and July 25, 2014, and failure rates of the 1%, 5% and 10% VaR forecasts for the out-of-sample period July 28, 2014 - December 17 2014. Standard errors of parameter estimates are reported in parentheses. Regarding the ABC estimator, asymptotic standard errors are reported and for the MCMC, posterior standard errors are reported. ABC-filter estimated loglikelihoods are also reported. Heteroskedasticity and autocorrelation consistent Vuong test comparing the MCMC specification to the ABC filtered specification are reported in parenthesis below the log-likelihood value of the MCMC loglikelihood.

<table>
<thead>
<tr>
<th></th>
<th>ABC Log-Likelihood</th>
<th>VaR forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>µ*(1−ϕ)</td>
<td>ϕ</td>
</tr>
<tr>
<td>ABC</td>
<td>-0.0478</td>
<td>0.9814</td>
</tr>
<tr>
<td></td>
<td>(0.0036)</td>
<td>(0.0017)</td>
</tr>
<tr>
<td>MCMC</td>
<td>0.0001</td>
<td>0.9795</td>
</tr>
<tr>
<td></td>
<td>(0.0162)</td>
<td>(0.0037)</td>
</tr>
</tbody>
</table>
Figure 5: S&P500 daily prices (top row, second line) and returns (top row, first line) observed from September 1, 1998 up to July 25, 2014. Sample autocorrelations of squares (first panel, bottom row) and cross-correlations of returns and future squared returns (second panel, bottom row) together with the corresponding plug-in moments obtained after fitting the T-GASV to the daily S&P500 returns. The dashed lines correspond to the moments implied by the model estimated with the ABC filter-based Maximum Likelihood estimator and the continuous lines correspond to the model estimated with the MCMC estimator.
Figure 6: VaR forecasts computed with the ABC filter-based Maximum Likelihood (continuous line) and MCMC (dotted line) estimators. Out-of-sample period ranges from July 28, 2014 to December, 2014.
Appendix A. Proof of Theorems

Appendix A.1. Proof of Theorem 2.1

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

\[
h_t - \mu = \sum_{i=1}^{\infty} \phi^{i-1} (f(\varepsilon_{t-i}) + \eta_{t-i}).
\]  

(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 \( \forall \varepsilon > 0 \), we can find a value of \( \delta = \sqrt{1 - \phi^2} \varepsilon \), such that \( x \) 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 \( |\Psi(x) - \Psi(x_0)| \leq \sqrt{\sum_{i=1}^{\infty} \phi^{2i-2}} \sqrt{\sum_{i=1}^{\infty} (x_i - x_0^i)^2} < \frac{\delta}{\sqrt{1 - \phi^2}} = \varepsilon \). Therefore, \( \Psi(x) \) is continuous, and consequently, measurable.

Second, given that \( \varepsilon_t \) and \( \eta_t \) are both IID and mutually independent for any lag and lead, then \( \{f(\varepsilon_t) + \eta_t\} \) is also an IID sequence. Lemma 3.5.8 of William (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 William (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 \( \varepsilon_t \) is independent of \( \sigma_t \) and strictly stationary by definition. Therefore, it is easy to show that \( y_t = \sigma_t \varepsilon_t \) 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 \( \varepsilon_t \) follows a distribution such that \( E(\exp(0.5 c f(\varepsilon_t))) < \infty \).

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

\[
\sigma_t^c = \exp(0.5 c \mu) \exp \left( 0.5 c \sum_{i=1}^{\infty} \phi^{i-1} (f(\varepsilon_{t-i}) + \eta_{t-i}) \right).
\]  

(A.2)

Given that \( \varepsilon_t \) 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)

\[
E(\sigma_t^c) = \exp(0.5 c \mu) \left[ \exp \left( 0.5 c \sum_{i=1}^{\infty} \phi^{i-1} f(\varepsilon_{t-i}) \right) \right] \left[ \exp \left( 0.5 c \sum_{i=1}^{\infty} \phi^{i-1} \eta_{t-i} \right) \right].
\]  

(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

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

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

Define $a_i = E(\exp(b_if(\epsilon_{t-i})))$. 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 $\prod_{i=1}^{\infty} a_i < \infty$.

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, $E(\sigma^c_t)$ are finite when $E(\exp(0.5c\phi^{i-1}f(\epsilon_{t-i}))) < \infty$. Note that when $|\phi| < 1$, $E(\exp(0.5cf(\epsilon_t))) < \infty$ guarantees that $E(\exp(0.5c\phi^{i-1}f(\epsilon_{t-i}))) < \infty$ for any positive integer $i$. Therefore, if $|\phi| < 1$ and $E(\exp(0.5cf(\epsilon_t))) < \infty$, $E(\sigma^c_t)$ 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

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

$$E(y_t^c) = E(\sigma^c_t)E(\epsilon_t^c). \quad (A.6)$$
Replacing formula (A.4) into (A.5) yields the following required expression

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

(A.7)

where \( P(x, y) \equiv \prod_{i=1}^{\infty} E(\exp(xy^{i-1}f(\epsilon_{t-i})))). \) 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 \(|y_t|^c\) has finite moments of arbitrary order \( c \). On the other hand, following the same steps, we obtain

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

(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

\[ E(|\epsilon_t|^c \sigma_{t-\tau}^c \sigma_{t-\tau}^c) - [E(|y_t|^c)]^2. \]  

(A.9)

Note that from equation (2), \( \sigma_t^c = \exp \{0.5ch_t\} \) can be written as follows

\[ \sigma_t^c = \exp \{0.5c(1 - \phi^\tau) \} \exp \left\{ 0.5c \sum_{i=1}^{\tau} \phi^{i-1}(f(\epsilon_{t-i}) + \eta_{t-i}) \right\} \sigma_{t-\tau}^{c\phi^\tau}. \]  

(A.10)

The following expression of the auto-covariance is obtained after substituting (A.7) and (A.10) into (A.9)

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

(A.11)

Given that \( \epsilon_t \) and \( \eta_t \) 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
\[
\text{cov}(|y_t|^c, |y_{t-\tau}|^c) = \\
\exp(c\mu)E(|\epsilon|^c)\exp\left(\frac{1 + c^2 + 2c\phi^\tau}{8(1 - \phi^2)}\sigma_\eta^2\right)E\left(|\epsilon|^c\exp\left(0.5c\phi^{i-1}f(\epsilon_i)\right)\right) \prod_{i=1}^{\tau-1}E\left(\exp\left(0.5c\phi^{i-1}f(\epsilon_i)\right)\right)
\]
\[
\cdot \prod_{i=1}^{\infty}E\left(\exp\left(0.5c(1 + \phi^\tau)\phi^{i-1}f(\epsilon_{i-\tau})\right)\right) - \exp(c\mu)(E(|\epsilon|^c))\exp\left(\frac{c^2\sigma_\eta^2}{4(1 - \phi^2)}\right)[P(0.5c, \phi)]^2.
\]

The required expression of \(\rho_c(\tau)\) follows directly from \(\rho_c(\tau) = \frac{\text{cov}(|y_t|^c, |y_{t-\tau}|^c)}{E(|y_t|^c)E(|y_{t-\tau}|^c)}\), 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
\[
\text{cov}(|y_t|^c, y_{t-\tau}) = \exp(0.5(c + 1)\mu)E(|\epsilon|^c)\exp\left(\frac{1 + c^2 + 2c\phi^\tau}{8(1 - \phi^2)}\sigma_\eta^2\right)E\left(|\epsilon|^c\exp\left(0.5c\phi^{i-1}f(\epsilon_i)\right)\right) \prod_{i=1}^{\tau-1}E\left(\exp\left(0.5c\phi^{i-1}f(\epsilon_i)\right)\right).
\] (A.12)

Finally, \(\rho_c(\tau) = \frac{\text{cov}(|y_t|^c, y_{t-\tau})}{\sqrt{E(|y_t|^c)E(y_{t-\tau})/E(y_t^c)}}\) 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 \(f(\epsilon) = \alpha I(\epsilon < 0) + \gamma_1\epsilon + \gamma_2(|\epsilon| - E|\epsilon|)\)

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 \(C_0 \equiv \frac{\nu}{\lambda^{2^{\nu/2}(\nu/2)+1}(\Gamma(\nu/2))^{1/2}}\) and \(\lambda \equiv (2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/\nu))^{1/2}\), 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
\[
E(|\epsilon|^p) = C_0 \int_{-\infty}^{+\infty} |\epsilon|^p \exp\left(-\frac{|\epsilon|^\nu}{2\lambda}\right) d\epsilon
\]
\[
= 2C_0 \int_{0}^{+\infty} e^p \exp\left(-\frac{e^\nu}{2\lambda}\right) d\epsilon.
\]
Substituting \(s = \frac{e^\nu}{2\lambda}\) and solving the integral yields
\[
E(|\epsilon|^p) = 2\frac{\bar{\epsilon}}{\lambda^p} \Gamma((p+1)/\nu)/\Gamma(1/\nu).
\] (B.1)
On the other hand,

\[
E(|e|^{p} \exp(bf(e)))
\]

\[
= \int_{-\infty}^{+\infty} |e|^{p} \exp(b\alpha I(e < 0) + b\gamma_1 e + b\gamma_2(|e| - E|e|)) C_0 \exp \left( -\frac{|e|^{\nu}}{2\lambda^{\nu}} \right) de
\]

\[
= C_0 \exp(-b\gamma_2 E|e|) \left[ \int_{-\infty}^{0} (-e)^{p} \exp(b\alpha) \exp(b(\gamma_1 - \gamma_2)e) \exp \left( -\frac{(-e)^{\nu}}{2\lambda^{\nu}} \right) de \\
+ \int_{0}^{+\infty} e^{p} \exp(b(\gamma_1 + \gamma_2)e) \exp \left( -\frac{e^{\nu}}{2\lambda^{\nu}} \right) de \right].
\]

Integrating by substitution with \( s = -e \) in the first integral, we obtain

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

\[
= C_0 \exp(-b\gamma_2 E|e|) \int_{0}^{+\infty} e^{p} \exp \left( -\frac{e^{\nu}}{2\lambda^{\nu}} \right) [\exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)e) + \exp(b(\gamma_1 + \gamma_2)e)] de.
\]

We can rewrite the previous equation by replacing \( e \) with \( \lambda(2\gamma)^{1/\nu} \) as follows

\[
E(|e|^{p} \exp(bf(e)))
\]

\[
= C_0 \exp(-b\gamma_2 E|e|) \frac{\lambda^{p+1} 2^{1-p}}{\nu} \\
\cdot \int_{0}^{+\infty} y^{-1 + \frac{1-p}{2}} \exp(-y) \left[ \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\lambda 2^{\frac{1}{\nu}} y^{\frac{1}{\nu}}) + \exp(b(\gamma_1 + \gamma_2)\lambda 2^{\frac{1}{\nu}} y^{\frac{1}{\nu}}) \right] dy.
\]

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

\[
E(|e|^{p} \exp(bf(e)))
\]

\[
= \exp(-b\gamma_2 2^{\frac{1}{\nu}} \lambda \Gamma(2/\nu) / \Gamma(1/\nu)) \frac{\lambda^{p+1} 2^{1-p}}{\nu} \\
\cdot \int_{0}^{+\infty} + \infty \sum_{k=0}^{+\infty} \left[ \exp(b\alpha) \left( b\lambda 2^{\frac{1}{\nu}}(\gamma_2 - \gamma_1) \right)^{k} + \left( b\lambda 2^{\frac{1}{\nu}}(\gamma_1 + \gamma_2) \right)^{k} \right] y^{-1 + \frac{1-p}{2}} \exp(-y) \frac{k!}{
\]

\[
\text{Define } \Delta = \max \{ |b\lambda 2^{1/\nu}(\gamma_1 + \gamma_2)|, \max(\exp(b\alpha), 1)|b\lambda 2^{1/\nu}(\gamma_2 - \gamma_1)| \}. \text{ Then, we can use the results in Nelson (1991) to show that if } \nu > 1 \text{ then the summation and integration in (B.3) can be interchanged. Further, applying Formula 3.381 #4 of Ryzhik et al. (2007) yields the following.}
\]
required expression$^{20}$

$$E(|e|^p \exp(b f(e)))$$

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

(B.4)

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

$$E(e^p \exp(b f(e)))$$

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

(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 $E(|e|^p \exp(b f(e))) < \infty$ if and only if the both integrals in (B.2) 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 $E(|e|^p \exp(b f(e)))$ are $b(\gamma_2 - \gamma_1) < 1/\sqrt{2}$ and $b(\gamma_2 + \gamma_1) < 1/\sqrt{2}$. That is $b(\gamma_2 - \gamma_1) < \sqrt{2}$ and $b(\gamma_2 + \gamma_1) < \sqrt{2}$. The conditions for the infinity of $E(e^p \exp(b f(e)))$ are the same as those for $E(|e|^p \exp(b f(e))) 0 < \nu \leq 1$.

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

$$E(|e|^p \exp(b f(e)))$$

$$= C_1 \left[ \int_{0}^{+\infty} e^p \exp(b \alpha) \exp(b(\gamma_2 - \gamma_1)e) \left( 1 + \frac{e^2}{d-2} \right)^{-\frac{d+1}{2}} d\epsilon 
+ \int_{0}^{+\infty} e^p \exp(b(\gamma_1 + \gamma_2)e) \left( 1 + \frac{e^2}{d-2} \right)^{-\frac{d+1}{2}} d\epsilon \right],$$  

(B.6)

$^{20}$See Nelson (1991) for the proof of finiteness of the formula.
where \( C_1 = \frac{r^{(d+1)}}{(d-2)\pi r^2} \exp\left[ -b\gamma_2 2^{1/\nu} \lambda \Gamma(2/\nu)/\Gamma(1/\nu) \right] \). 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** \( E(|y_t|^c), \text{corr}(|y_t|^c,|y_{t+1}|^c) \) and \( \text{corr}(y_t,|y_{t+1}|^c) \) \( \text{when} \ \epsilon \sim N(0,1) \)

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

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

(B.7)

and

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

(B.8)

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

\[
E(\exp(bf(\epsilon))) = \exp\left( -b\gamma_2 \sqrt{\frac{2}{\pi}} \right) \left( \exp(ba) \exp(\bar{A})\Phi(\bar{C}) + \exp(\bar{B})\Phi(\bar{D}) \right),
\]

\[
E(\epsilon \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \exp\left( -b\gamma_2 \sqrt{\frac{2}{\pi}} \right) \left\{ -\exp(ba) \left[ 1 + \sqrt{2\pi} \bar{C} \exp(\bar{A})\Phi(\bar{C}) \right] + \left[ 1 + \sqrt{2\pi} \bar{D} \exp(\bar{B})\Phi(\bar{D}) \right] \right\},
\]

\[
E(|\epsilon| \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \exp\left( -b\gamma_2 \sqrt{\frac{2}{\pi}} \right) \left\{ \exp(ba) \left[ 1 + \sqrt{2\pi} \bar{C} \exp(\bar{A})\Phi(\bar{C}) \right] + \left[ 1 + \sqrt{2\pi} \bar{D} \exp(\bar{B})\Phi(\bar{D}) \right] \right\},
\]

and

\[
E(|\epsilon|^2 \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \exp\left( -b\gamma_2 \sqrt{\frac{2}{\pi}} \right) \left\{ \exp(ba) \left[ \bar{C} + \sqrt{2\pi}(\bar{C}^2 + 1) \exp(\bar{A})\Phi(\bar{C}) \right] + \left[ \bar{D} + \sqrt{2\pi}(\bar{D}^2 + 1) \exp(\bar{B})\Phi(\bar{D}) \right] \right\},
\]

37
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 $\bar{D} = b(\gamma_1 + \gamma_2)$.

Appendix C. Continuous Resampling of $\{\tilde{x}_{t,1}^{(n)}\}$ using Malik and Pitt (2011) procedure

Step (i) (Sorting): Sort the particles $\{\tilde{x}_{t,1}^{(n)}\}$ in ascending order

$$\tilde{x}_{t,1}^{(1)} \leq \tilde{x}_{t,1}^{(2)} \leq \cdots \leq \tilde{x}_{t,1}^{(N)}.$$  

Denote the importance weight associated with $\tilde{x}_{t,1}^{(n)}$ by $\hat{p}_{i}^{(n)}$.

Step (ii) (Sorted uniforms): Draw sorted $\{u_i\}$ using Malmquist (1950)’s sorted uniforms. That is draw $\{v_i\}$ from a uniform distribution on the interval $(0, 1)$. Set $u_N = v_N^{1/N}$ and for $n = N-1, \ldots, 1$: 

$$u_n = u_{n+1}v_n^{1/n}.$$  

The resulting $u_1 < u_2 < \cdots < u_N$ are ordered uniform variables.

Step (iii) (Index sampling): Define $\pi_t^{(0)} = \hat{p}_t^{(1)}/2$, $\pi_t^{(N)} = \hat{p}_t^{(N)}/2$ and $\pi_t^{(k)} = (\hat{p}_t^{(k+1)} + \hat{p}_t^{(k)})/2$ for $k = 1, \ldots, N-1$, and set $s = 0$, $j = 1$. For $i = 0, \ldots, N$, sample indexes $r_1, \ldots, r_N$ and a new set of uniforms $u_1^*, \ldots, u_N^*$:

$$s = s + \pi_t^{(i)}$$

while ($u_j \leq s$ AND $j \leq N$) {

$$r_j = i$$

$$u_j^* = (u_j - (s - \pi_t^{(i)}))/\pi_t^{(i)}$$

$$j = j + 1$$

}

Step (iv) (Particle drawing): For the selected regions where $r_j = 0$, we set $x_{t,1}^{(j)} = \tilde{x}_{t,1}^{(1)}$ and where $r_j = N$ we set $x_{t,1}^{(j)} = \tilde{x}_{t,1}^{(N)}$. Otherwise, set

$$x_{t,1}^{(j)} = (\tilde{x}_{t,1}^{(r_j+1)} - \tilde{x}_{t,1}^{(r_j)}) \times u_j^* + \tilde{x}_{t,1}^{(r_j)}.$$  

38


