Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Asymmetric Stochastic Volatility Models

Autor:

XIUPING MAO

Director/es:

ESTHER RUIZ

HELENA VEIGA

DEPARTAMENTO DE ESTADÍSTICA

Getafe, Octubre 2014

Universidad Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Asymmetric Stochastic Volatility Models

Autor: *XIUPING MAO*

Director/es: **ESTHER RUIZ & HELENA VEIGA**

Firma del Tribunal Calificador:

Firma

Presidente: (Nombre y apellidos)

Vocal: (Nombre y apellidos)

Secretario: (Nombre y apellidos)

Calificación:

Getafe, de de

Universidad Carlos III

PH.D. THESIS

Asymmetric Stochastic Volatility Models

Author:

Xiuping Mao

Advisor:

Esther Ruiz & Helena Veiga

DEPARTMENT OF STATISTICS

Getafe, Madrid, January 9, 2015

To my family and friends!

Acknowledgements

I would like to use this place to express my gratitude and love to all the people in my life, who were there for me and were a part of the process during which this thesis came to life.

First and foremost, I would like to express my deepest gratitude to my advisors Prof. Esther Ruiz and Prof. Helena Veiga for their excellent guidance, patience, enthusiasm, immense knowledge and providing me continuous support of my Ph.D study and research. The good advice, support and friendship of my advisors has been invaluable on both an academic and a personal level, for which I am extremely grateful. I could not have imagined having a better advisor and mentor for my Ph.D study.

Also, I would like to express my gratitude to my colleagues and friends from the Department of Statistics of Universidad Carlos III de Madrid. Especially Juan Miguel Marín and for his suggestions on Byesian estimation. Audra Virbickaitė, João Henrique and Guillermo Carlomagno for being great officemates over the years. And Diego Fresoli and Jorge E. Galán for their help in teaching.

Last but not the least, I would like to thank my family: my parents Shili Mao and Shuying chen, for giving birth to me at the first place and supporting me spiritually throughout my life. My parents and my younger brother, Yongqiang Mao, have provided the loving and stimulating environment I needed to continue. To end with, I would like to express my special thankfulness to my boyfriend, Yanyun Zhao, for his patience, love, never-ending support and all the laughter he gave me throughout all these years. Also for always staying with me in good and bad times. I love you all very much!

Abstract

This dissertation focuses on the analysis of Stochastic Volatility (SV) models with leverage effect. We propose a general family of asymmetric SV (GASV) models and consider in detail two particular specifications within this family. The first one is the Threshold GASV (T-GASV) model which nests some of the most famous asymmetric SV models available in the literature with the errors being either Normal or GED. We also propose score driven GASV models with different assumptions about the error distribution, namely the Normal, Student-t or GED distributions, where the volatility is driven by the score of the lagged return distribution conditional on the volatility. Closed-form expressions of some statistical moments of interest of these two GASV models are derived and analyzed. We show that some of the parameters of these models cannot be properly identified by the moments usually considered when describing the stylized facts of financial returns, namely, excess kurtosis, autocorrelations of squares and cross-correlations between returns and future squared returns. As a byproduct, we obtain the statistical properties of those nested popular asymmetric SV models, some of which 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. Analyzing the SNIS, we find that the asymmetric impact of the level disturbance on the volatility can be different depending on the volatility disturbance.

Finally, we analyze the finite sample properties of a MCMC estimator of the parameters and volatilities of some restricted GASV models. Furthermore, estimating the restricted T-GASV model using this MCMC estimator, we show that 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 models to both daily and weekly financial returns.

Contents

List of Figures

Chapter 1

Introduction

This dissertation focuses on asymmetric Stochastic Volatility models for modelling the financial returns. It has been commonly accepted that, although the returns are usually uncorrelated, the second order moment of the conditional distribution of financial returns is time-varying. There are two main well known features of the time-varying volatility of financial returns, namely *volatility clustering* and *leverage effect*.

Volatility clustering refers to large (small) absolute returns tending to be followed by large (small) absolute returns. This behavior is reflected in the fact that power transformed absolute returns display a positive and slowly decaying autocorrelation function. As an illustration, consider a series of daily S&P500 returns observed from September 1, 1998 to July 25, 2014 with $T = 4000$ observations. The returns are computed as $y_t = 100 \times \triangle \log P_t$, where P_t is the adjusted close price from yahoo.finance on day t. The raw prices together with their corresponding returns are plotted in Figure 1.1, which suggests the presence of volatility clustering. It is also supported by the positive and significant sample autocorrelations of both squared and absolute returns plotted in the last two panels of Figure 1.2. However, the sample autocorrelations of returns plotted in the first panel of Figure 1.2 are not significant indicating that they are uncorrelated.

When modeling the second order dynamics of univariate financial returns, it is often observed that volatility increases are larger in response to negative than to positive past returns of the same

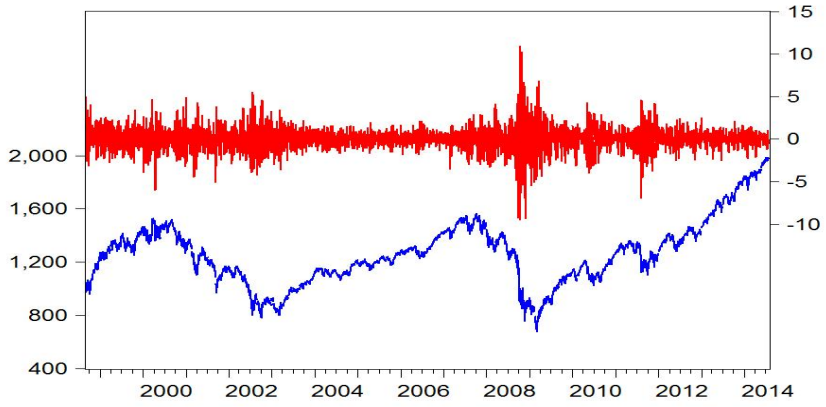

Figure 1.1: S&P500 daily prices (bottom line) and returns (top line) observed from September 1, 1998 up to July 25, 2014.

magnitude; see Bollerslev et al. (2006) for a comprehensive list of references and Hibbert et al. (2008) for a behavioral explanation. After Black (1976), this asymmetric response of volatility is popularly known as *leverage effect* in the related literature. This effect is due to the impact of negative shocks on the value of a firm. In particular, bad news tends to decrease the stock price, and consequently, increase the financial leverage or the debt-to-equity ratio of a firm. On the other hand, this leads to an increase of the risk and to raising the future expected volatility of the stock return. The leverage effect is also reflected by the negative and significant cross-correlations between returns and future absolute or squared returns. Looking at the S&P 500 prices and returns in Figure 1.1, we can observe episodes of large volatilities in returns associated with periods of negative movements in prices. Furthermore, this association can also be observed in the negative sample cross-correlations between returns and future squared and absolute returns plotted in Figure 1.3.

Modeling volatility clustering with asymmetries has led to an enormous literature. Two main alternative families of models are usually implemented. The first family is based on the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model of Bollerslev (1986), with the volatilities specified as a function of past returns and, consequently, observable one-step ahead; see Engle (1995), Giraitis et al. (2007) and Teräsvirta (2009) for comprehensive reviews

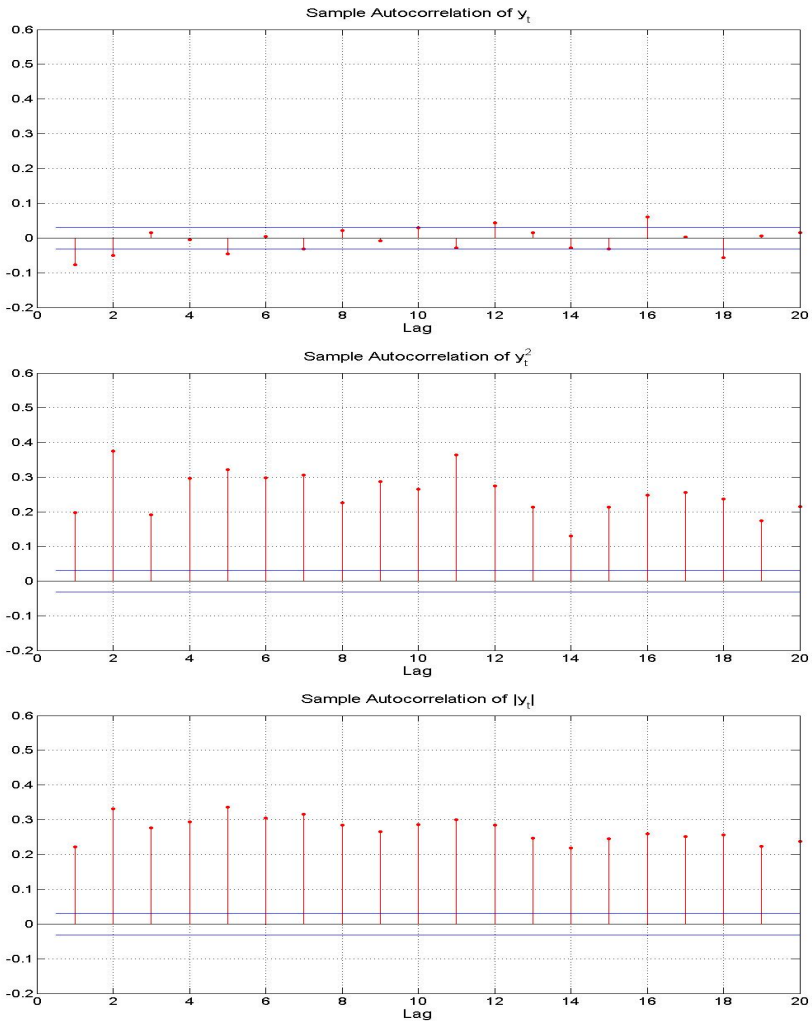

Figure 1.2: Sample autocorrelations of the returns (top panel), the squared returns (middle panel) and the absolute returns (bottom panel) of the S&P500 daily returns.

on GARCH models. Alternatively, the second family includes Stochastic Volatility (SV) models, which specify the volatility as a latent variable that is not directly observable; see Ghysels et al. (1996) and Cavaliere (2006) for reviews on SV models and their applications.

Both GARCH and SV models have been extended to represent the dynamic evolution of conditionally heteroscedastic time series with leverage effect. Among the GARCH family, the main proposals are: the Exponential GARCH (EGARCH) model of Nelson (1991), the Glosten-Jagannathan-Runkle (GJR) model of Glosten et al. (1993), the Asymmetric Power ARCH (APARCH)

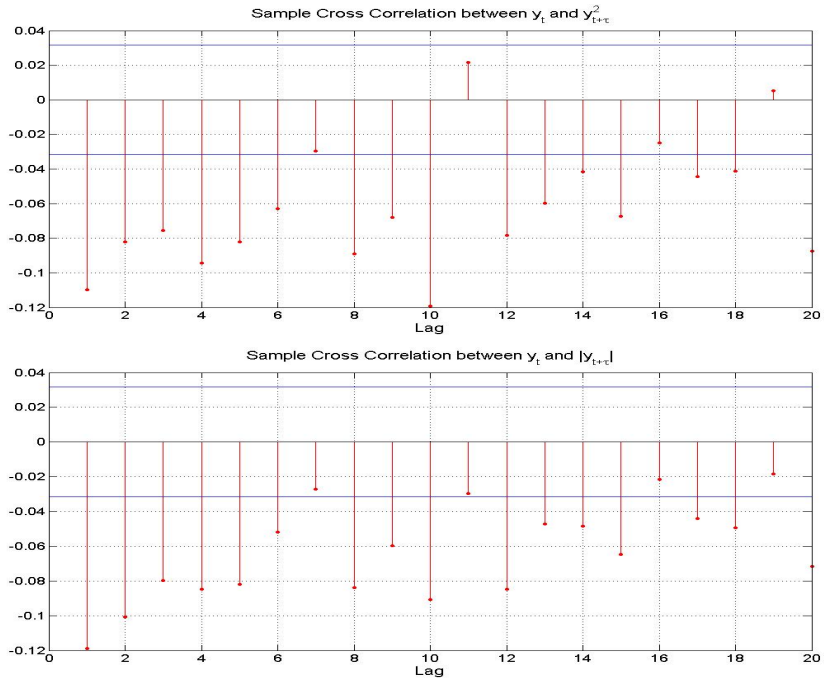

Figure 1.3: Sample cross-correlations between returns and future squared returns (top panel) and the future absolute returns (bottom panel) of the S&P500 daily returns.

model of Ding et al. (1993), the Threshold GARCH (TGARCH) of Zakoian (1994) and the Generalized Quadratic GARCH (GQARCH) of Sentana (1995). The similarities and differences among these asymmetric GARCH models have been described by Rodríguez and Ruiz (2012) who show that, among them, the EGARCH specification is the most flexible while the GJR and GQARCH models may have important limitations to represent the volatility dynamics often observed in real financial returns if their parameters are restricted to guarantee the positivity, stationarity and finite kurtosis restrictions. Furthermore, their empirical study shows that the conditional standard deviations estimated by the TGARCH and EGARCH models are almost identical and very similar to those estimated by the APARCH model, while the estimates of the GQARCH and GJR models differ among them and with respect to the other three specifications.

SV models have also being extended to cope with leverage effect. Extensions of the simple discrete time model, due to Taylor (1986), have been proposed, among others, by Wiggins (1987), Chesney and Scott (1989), Harvey and Shephard (1996) and So et al. (2002). Consequently, 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 the 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. 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 these asymmetric SV models 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 for explaining the empirical properties of financial returns.

In this thesis, we focus on the asymmetric SV models. First of all, the SV models are shown to be more flexible than GARCH models to represent the properties often observed in real financial returns; see Carnero et al. (2004). Second, 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. Third, even though models within the GARCH family have been extensively analyzed in the literature, the advantages and limitations of the alternative asymmetric SV models have not been previously analyzed. Knowing the moments of returns implied by different specifications can be important when estimating the parameters using estimators based on the Method of Moments (MM) as those proposed, for example, by Bollerslev and Zhou (2002) and Garcia et al. (2011). Furthermore, knowing the moments of the alternative specifications, we can compare them to see which one is more adequate to explain the empirical properties often observed when dealing with real data, namely, leptokurtosis, positive and persistent autocorrelations of power-transformed absolute returns and negative cross-correlations between returns and future power-transformed absolute returns. We propose a family of asymmetric SV models that we call generalized asymmetric SV

(GASV) and derive its properties. The GASV family is rather general including as particular cases some of the most popular asymmetric SV models. The analytical expressions of their statistical properties are obtained, so that we are able to point out the advantages and limitations of each of the restricted specifications.

Besides volatility clustering and leverage effect, another important and well documented empirical feature of standardized financial returns is the fact that they are heavy-tailed distributed; see, for instance, Liesenfeld and Jung (2000), Jacquier et al. (2004) and Chen et al. (2008) among many others. In order to capture this latter feature, both GARCH and SV models have been extended by assuming fat-tailed return errors. Two examples are the GARCH-t model of Bollerslev (1987) and the asymmetric SV model with Student-t distribution of Asai and McAleer (2011). Nonetheless, these traditional models often specify the asymmetric volatility as being driven by past return errors. Consequently, they can suffer from a potential drawback since a large realisation of the return error, which could be due to the heavy-tailed nature of its distribution, will be attributed to an increase in volatility. Therefore, in the GARCH context, Creal et al. (2013) and Harvey (2013) have recently proposed models in which the dynamic of volatility is driven by the lagged score of the conditional distribution of returns to automatically correct for influential observations. This gives rise to the Generalised Autoregressive Score (GAS) models which are also known as dynamic conditional score (DCS) models. We extend the GAS idea to asymmetric SV models by specifying the unobserved volatility to be driven by lagged scores. Given that the conditional distribution of returns does not have an analytical expression, the score is computed with respect to the distribution of returns conditional on the volatilities. We show that this type of models lays in the GASV family. We denote the new models as GAS-GASV (GAS^{2}V) and consider three alternative GAS^{2}V models depending on the assumed distribution of the return errors, namely, Normal, Student-t and Generalised Error Distribution (GED). Closed-form expressions of several relevant statistics of these models are derived to analyse their ability to represent the main empirical features often observed in financial returns. It is important to point out that analytical expressions of these moments of the GAS^{2}V model with Student-t errors can be

derived, in opposition to the traditional specifications of the SV models in which their derivation is hardly possible when the errors are Student-t. Moreover, we show that the GAS^{2}V model with Student-t errors generates returns with very similar properties to those generated by the GAS^{2}V model with GED errors as far as the parameters of both distributions are chosen to have the same kurtosis. Therefore, this could indicate the existence of difficulties in identifying the parameters of the GAS^{2}V model when looking at the moments.

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 thesis, 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).[1] 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. Examples of procedures based on the Monte Carlo likelihood evaluation are the simulated Maximum Likelihood (MCL) procedure of Durbin and Koopman (1997) and the Efficient

[1] The SNIS proposed in this thesis 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).

Importance Sampling (EIS) procedure of Liesenfeld and Richard (2003) and Richard and Zhang (2007); see also Asai and McAleer (2011) for the implementation of the latter procedure for estimating their exponential SV model and Koopman et al. (2014) for an extension. Alternatively, 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 BUGS 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, for example, by Yu (2012) and Wang et al. (2013). The MCMC estimator implemented by BUGS 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 restricted T-GASV and GAS^{2}V models in situations similar to those encountered when analyzing time series of real financial returns. Furthermore, we show that the nested specifications of the restricted T-GASV model can be adequately identified when the parameters are estimated using the BUGS software. Therefore, in empirical applications, researchers will be better off by fitting the general model proposed in this thesis and letting the data to choose the preferred specification of the volatility instead of choosing a particular ad hoc specification.

The rest of this dissertation is organized as follows. Chapter 2 proposes the GASV family and derives its statistical properties. Moreover, we propose the T-GASV model which is included in the GASV family and incorporates some of the most famous asymmetric SV models previously available. We consider a MCMC estimator of the restricted T-GASV model and conduct Monte Carlo experiments to analyze its finite sample properties. An empirical application to daily S&P500 returns is presented. In Chapter 3, we propose the GAS^{2}V model and fit it to both daily and weekly financial returns. Finally, Chapter 4 concludes the thesis and proposes possible lines

of future research.

Chapter 2

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

2.1 Introduction

A variety of alternative SV models are available to choose among for modeling the financial returns with leverage effect, such as the asymmetric autoregressive SV (A-ARSV) model of Taylor (1994) and Harvey and Shephard (1996), the Exponential SV (E-SV) model of Demos (2002) and Asai and McAleer (2011) and the Threshold SV (T-SV) model of Breidt (1996) and So et al. (2002) among many others. Although these models are often implemented to present the dynamic dependence of volatilities, their statistical properties are either partially known or completely unknown.

In this chapter, we propose a general family of asymmetric SV models, named as GASV family, and derive the general expression of its statistical properties. This GASV family is rather general including as particular cases some of the asymmetric SV models mentioned above. Moreover, we propose further a specification, called T-GASV model, with the motivation that it nests some

of the most popular asymmetric volatility specifications previously available in the literature. The closed-form expressions of its statistical properties are obtained. As a marginal outcome of this analysis, we also obtained the statistical properties of the models nested within the T-GASV model, some of which were previously unknown in the literature and, hence, we are able to point out the advantages and limitations of each of the restricted specifications. We also propose a useful tool, SNIS, to describe the asymmetric response of volatility to positive and negative past returns. It is a surface relating the conditional volatility with the level and volatility disturbances. 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 considered as competitive alternatives to GARCH models, their implementation is always limited due to the intractable likelihood. In this chapter, we consider a MCMC estimator of the GASV models implemented by the user-friendly and free software, BUGS. We carry out extensive Monte Carlo experiments to analyze its finite sample performance when estimating both the parameters and the underlying volatilities of the restricted T-GASV model. Moreover, we also find that, by fitting our restricted T-GASV model to the series generated from those nested asymmetric SV models, it is able to identify the true Data Generating Process (DGP). Finally, the MCMC estimator is implemented to estimate the volatilities and forecast the Value at Risk (VaR) of the daily S&P500 return series after fitting all the asymmetric SV models considered in this chapter.

The rest of this chapter is organized as follows. Section 2.2 defines the GASV family and derives its statistical properties. Section 2.3 proposes the SNIS to describe the asymmetric response of volatility. The properties of the T-GASV are analyzed and compared in Section 2.4. In Section 2.5, we analyze and compare different asymmetric SV models contained in the GASV family. Section 2.6 conducts Monte Carlo experiments to analyze the finite sample properties of the MCMC estimator of the parameters and underlying volatilities of the restricted T-GASV model and presents an empirical application to daily S&P500 returns. Finally, the main conclusions and some guidelines for future research are summarized in Section 2.7.

2.2 The GASV family and its statistical properties

In this section, we define the GASV family 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.

2.2.1 Model description

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

$$y_t = \exp(h_t/2)\epsilon_t, \qquad t = 1, \cdots, T \tag{2.1}$$

$$h_t - \mu = \phi(h_{t-1} - \mu) + f(\epsilon_{t-1}) + \eta_{t-1}, \tag{2.2}$$

where $f(\epsilon_{t-1})$ is any function of ϵ_{t-1} for which no restrictions are imposed further than being a function of ϵ_{t-1} but not of the other disturbance in the model, η_{t-1}. Therefore, given ϵ_t, $f(\epsilon_t)$ is observable. The volatility noise, η_t, is a Gaussian white noise with variance σ_η^2.[1] It is assumed to be independent of ϵ_t for all leads and lags. The scale parameter, μ, is related with the marginal variance of returns, while ϕ 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 (2.1) and (2.2), 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 (2.2) is rather general,

[1] The normality of η_t when $f(\epsilon_{t-1}) = 0$ has been justified by, for example, Andersen et al. (2001a) and Andersen et al. (2001b, 2003).

it rules out models in which the persistence, ϕ, and/or the variance of the volatility noise, σ_η^2, are time-varying. Finally, note that the only assumption made about the distribution of the level disturbance, ϵ_t, is that it is an IID sequence with mean zero and variance one. As a consequence, ϵ_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 ϵ_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)[2]; 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 ϵ_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).

2.2.2 Moments of returns

We now derive the statistical properties of the GASV family in equations (2.1) and (2.2). 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 (2.1) and (2.2). The process $\{y_t\}$ is strictly stationary if $|\phi| < 1$. Further, if ϵ_t follows a distribution such that both $E(\exp(0.5cf(\epsilon_t)))$ 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 order c which are given by*

$$E(|y_t|^c) = \exp\left(\frac{c\mu}{2}\right) E(|\epsilon_t|^c) \exp\left(\frac{c^2\sigma_\eta^2}{8(1-\phi^2)}\right) P(0.5c\phi^{i-1}) \tag{2.3}$$

[2]The GED distribution with parameter ν 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

$$E(y_t^c) = \exp\left(\frac{c\mu}{2}\right) E(\epsilon_t^c) \exp\left(\frac{c^2\sigma_\eta^2}{8(1-\phi^2)}\right) P(0.5c\phi^{i-1}), \tag{2.4}$$

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

Proof. See Appendix A.1.1. □

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 (2.4), if ϵ_t has a symmetric distribution, then all odd moments of y_t are zero. Furthermore, from expression (2.3), 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* (2.3) *as follows*

$$\sigma_y^2 = \exp\left(\mu + \frac{\sigma_\eta^2}{2(1-\phi^2)}\right) P(\phi^{i-1}). \tag{2.5}$$

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* (2.3) *with* $c = 4$ *and* $c = 2$ *as follows*

$$\kappa_y = \kappa_\epsilon \exp\left(\frac{\sigma_\eta^2}{1-\phi^2}\right) \frac{P(2\phi^{i-1})}{(P(\phi^{i-1}))^2}, \tag{2.6}$$

where κ_ϵ *is the kurtosis of* ϵ_t.

The kurtosis of the basic symmetric Autoregressive SV (ARSV) model considered by Harvey et al. (1994) is given by $\kappa_\epsilon \exp\left(\frac{\sigma_\eta^2}{1-\phi^2}\right)$. 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 (2.3) depends on $f(\cdot)$ and on the distribution of ϵ_t. Therefore, in order to obtain closed-form expressions of the variance and kurtosis of returns, one needs to assume a particular distribution of ϵ_t and a specification of $f(\epsilon_t)$. We will particularize these expressions for some popular distributions and specifications in Section 2.4. Also, it is important to note that even for those cases in which the function $f(\cdot)$ and/or the distribution of ϵ_t are such that they do not allow to obtain closed-form expressions of the moments, expression (2.3) can always be used to simulate them as far as they are finite.

2.2.3 Dynamic dependence

Looking at the dynamic dependence of returns when they are defined as in (2.1) and (2.2), 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 (2.1) and (2.2) with $|\phi| < 1$. If ϵ_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 τ-th order autocorrelation of $|y_t|^c$ is finite and given by*

$$\rho_c(\tau) = \frac{E(|\epsilon_t|^c)E(|\epsilon_t|^c \exp(0.5c\phi^{\tau-1}f(\epsilon_t)))\exp\left(\frac{\phi^\tau c^2\sigma_\eta^2}{4(1-\phi^2)}\right)P(0.5c(1+\phi^\tau)\phi^{i-1})T(\tau,0.5c\phi^{i-1})-[E(|\epsilon_t|^c)P(0.5c\phi^{i-1})]^2}{E(|\epsilon_t|^{2c})\exp\left(\frac{c^2\sigma_\eta^2}{4(1-\phi^2)}\right)P(c\phi^{i-1})-[E(|\epsilon_t|^c)P(0.5c\phi^{i-1})]^2}, \tag{2.7}$$

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.1.2. □

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 (2.7) 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 (2.1) and (2.2) with $|\phi| < 1$. If ϵ_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 τ-th order cross-correlation between y_t and $|y_{t+\tau}|^c$ for $\tau > 0$ is finite and given by*

$$\rho_{c1}(\tau) = \frac{E(|\epsilon_t|^c)\exp\left(\frac{2c\phi^\tau - 1}{8(1-\phi^2)}\sigma_\eta^2\right)E(\epsilon_t \exp(0.5c\phi^{\tau-1}f(\epsilon_t)))P(0.5(1+c\phi^\tau)\phi^{i-1})\frac{T(\tau, 0.5c\phi^{i-1})}{\sqrt{P(\phi^{i-1})}}}{\sqrt{E(|\epsilon|^{2c})\exp\left(\frac{c^2\sigma_\eta^2}{4(1-\phi^2)}\right)P(c\phi^{i-1}) - [E(|\epsilon_t|^c)P(0.5c\phi^{i-1})]^2}}. \tag{2.8}$$

Proof. See Appendix A.1.3. □

2.3 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. It 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, ϵ_{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 ϵ_t on σ_{t+1} for different values of η_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, σ_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 and the conditional variance are different objects. Therefore, in this thesis, we propose measuring the effect of past shocks, ϵ_{t-1} and η_{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 σ_t^2 with ϵ_{t-1} and η_{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, ϵ_{t-1}, and volatility shocks, η_{t-1}, on the volatility at time t. Therefore, the SNIS is given by

$$\text{SNIS}_t = \exp((1-\phi)\mu)\sigma_y^{2\phi} \exp\left(f(\epsilon_{t-1}) + \eta_{t-1}\right). \tag{2.9}$$

Note that the shape of SNIS does not depend on the type of the distribution of ϵ_t as it is a function of $f(\epsilon_{t-1})$ and η_{t-1}.

For illustrating the SNIS, we consider the following specification of $f(\cdot)$

$$f(\epsilon_t) = \alpha I(\epsilon_t < 0) + \gamma_1 \epsilon_t + \gamma_2 |\epsilon_t|, \tag{2.10}$$

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 (2.1), (2.2) and (2.10) as Threshold GASV

(T-GASV).[3] 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 ϵ_t follows a Gaussian distribution, we obtain the 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 (2.10) 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 ϵ_t.

Figure 2.1 plots the SNIS of the T-GASV model with $\{\phi, \sigma_\eta^2, \alpha, \gamma_1, \gamma_2\} = \{0.98, 0.05, 0.07, -0.08, 0.1\}$ and μ is chosen such that $\exp((1-\phi)\mu)\sigma_y^{2\phi} = 1$. These parameter values are chosen to resemble those often obtained when the 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 (2.10). The leverage effect is very clear when the volatility shock is positive. The most important feature of the SNIS plotted in Figure 2.1 is that it shows that the leverage effect of SV models is different depending on the values of the volatility shock. In practice, when η_{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 2.1. It is important to observe that by introducing η_t in the T-GASV model, more flexibility is added to represent the leverage effect.

Summarizing, Figure 2.1 shows that, for the T-GASV model and the particular parameter values considered, given a value of the lagged volatility shock, η_{t-1}, the response of volatility is stronger when ϵ_{t-1} is negative than when it is positive with the same magnitude. Furthermore, this asymmetric response depends on the log-volatility noise, η_{t-1}. The leverage effect is clearly stronger when η_{t-1} is positive and large than when it is negative.

[3] 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 (2.10) with long-memory. However, they do not develop further the statistical properties of the model.

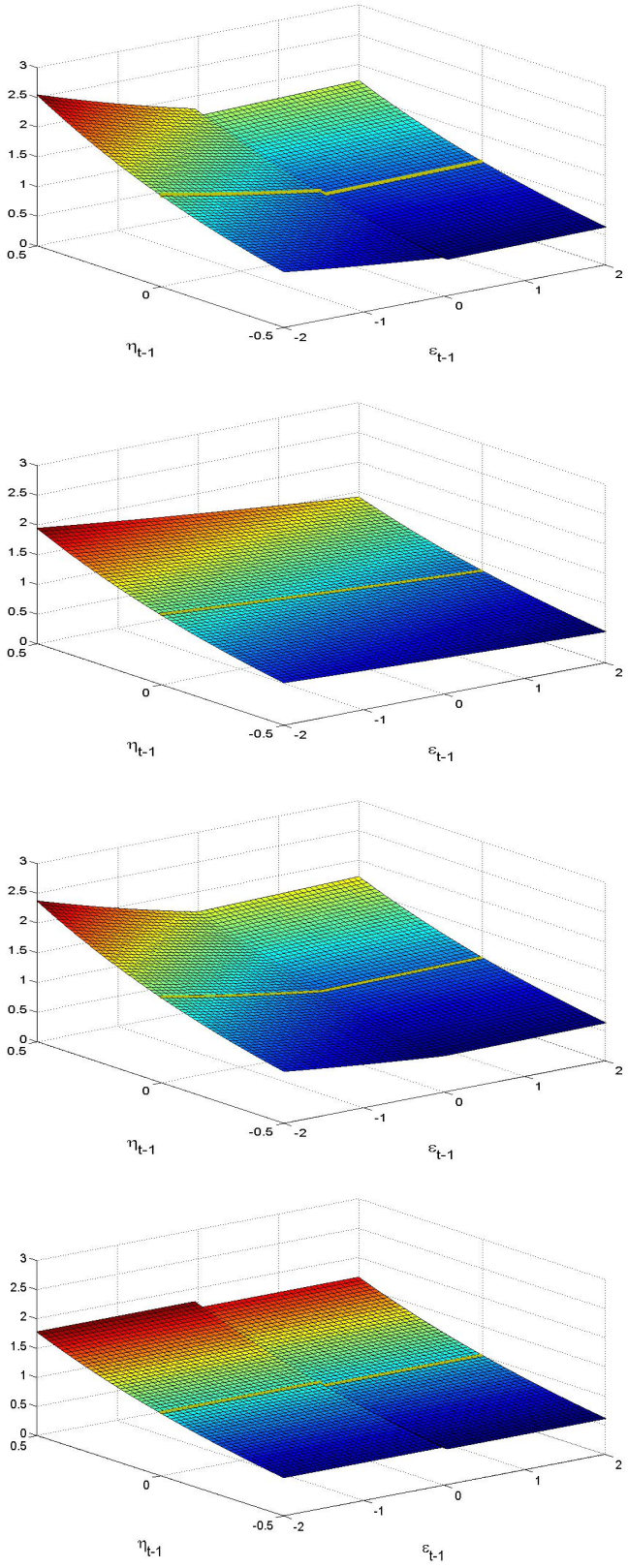

Figure 2.1: SNIS of different GASV models with $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ and $\exp((1-\phi)\mu)\sigma_y^{2\phi} = 1$. The parameter values are $\{\alpha, \gamma_1, \gamma_2\} = \{0.07, -0.08, 0.1\}$. Top panel corresponds to the SNIS of T-GASV, the second panel corresponds to the SNIS of A-ARSV model; the third panel is the SNIS of E-SV model and bottom panel is the SNIS of RT-SV model.

2.4 Threshold GASV model

As mentioned above, appropriate choices of the function $f(\cdot)$ and of the distribution of ϵ_t allow obtaining closed-form expressions of the moments of returns. In this section, we derive these expressions for the T-GASV model when ϵ_t follows a GED distribution.

Consider the T-GASV model defined in equations (2.1), (2.2) and (2.10) 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 A.2.1 for the corresponding expectations. In particular, the marginal variance of y_t is given by equation (2.5) with

$$P(b_i) = \prod_{i=1}^{\infty} \left\{ \sum_{k=0}^{\infty} \left(\left(\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)} \right)^{k/2} \frac{\Gamma((k+1)/\nu)}{2\Gamma(1/\nu)k!} b_i^k \left[(\gamma_1 + \gamma_2)^k + \exp(\alpha b_i)(\gamma_2 - \gamma_1)^k \right] \right) \right\}, \quad (2.11)$$

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 in expression (2.6) with $P(\phi^{i-1})$ and $P(2\phi^{i-1})$ as in expression (2.11)

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 A.2.2 for the corresponding expectations. In particular, the marginal variance is given by expression (2.5) while the kurtosis is given by expression (2.6) with

$$P(b_i) = \prod_{i=1}^{\infty} \left\{ \exp\left(\alpha b_i + \frac{b_i^2(\gamma_1 - \gamma_2)^2}{2} \right) \Phi(b_i(\gamma_2 - \gamma_1)) + \exp\left(\frac{b_i^2(\gamma_1 + \gamma_2)^2}{2} \right) \Phi(b_i(\gamma_2 + \gamma_1)) \right\}, \quad (2.12)$$

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

When $\nu < 1$, we cannot obtain analytical expressions of $E(|y_t|^c)$. However, in A.2.1, we show that $E(|y_t|^c)$ in equation (2.3) is finite if $\gamma_2 + \gamma_1 \leq 0$ and $\gamma_2 - \gamma_1 \leq 0$.[4] Finally, if $\nu = 1$, the

[4]The same conditions should be satisfied for the finiteness of $E(|y_t|^c)$ when ϵ_t follows a Student-t distribution with $d > 2$ degrees of freedom.

conditions for the existence of $E(|y_t|^c)$ in equation (2.3) are $\gamma_2 + \gamma_1 < 2\sqrt{2}/c$ and $\gamma_2 - \gamma_1 < 2\sqrt{2}/c$.

As mentioned in Section 2.2, 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 2.2 which plots it as a function of the leverage parameters α and γ_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 α, γ_1 and γ_2 on the kurtosis of returns are very different depending on the persistence. The kurtosis increases with α, $|\gamma_1|$ and γ_2. However, their effects are only appreciable when ϕ is close to 1.

The expectations needed to obtain closed-form expressions of the autocorrelations in expression (2.7) and cross-correlations in (2.8) have been derived in A.2.1 for the T-GASV model with parameter $\nu > 1$ and in A.2.2 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 (2.7) 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 $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ 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 2.3 as functions of the leverage parameters, γ_1 and α. In the top left panel of Figure 2.3, which corresponds to the autocorrelations of squares, we can observe that they are larger, the larger is γ_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 2.3. They are also larger the larger is the parameter γ_2. However, we can observe that the autocorrelations of absolute returns increase with the threshold parameter α. The effect of γ_1

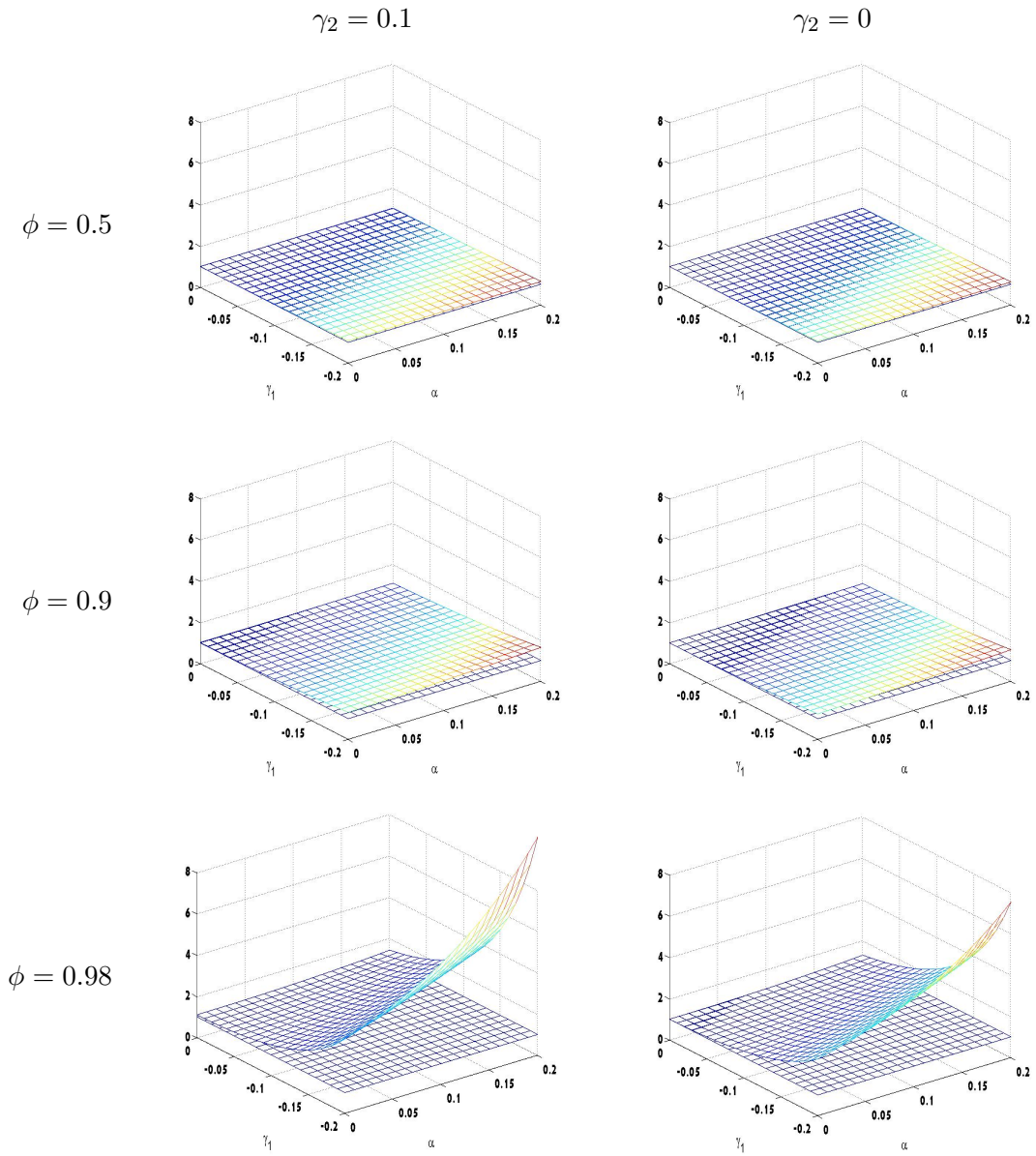

Figure 2.2: Ratio between the kurtoses of the T-GASV model and the symmetric ARSV model with Gaussian errors when $\gamma_2 = 0.1$ (left column) and 0 (right column) for three different values of the persistence parameter, $\phi = 0.5$ (first row), $\phi = 0.9$ (middle row) and $\phi = 0.98$ (bottom row).

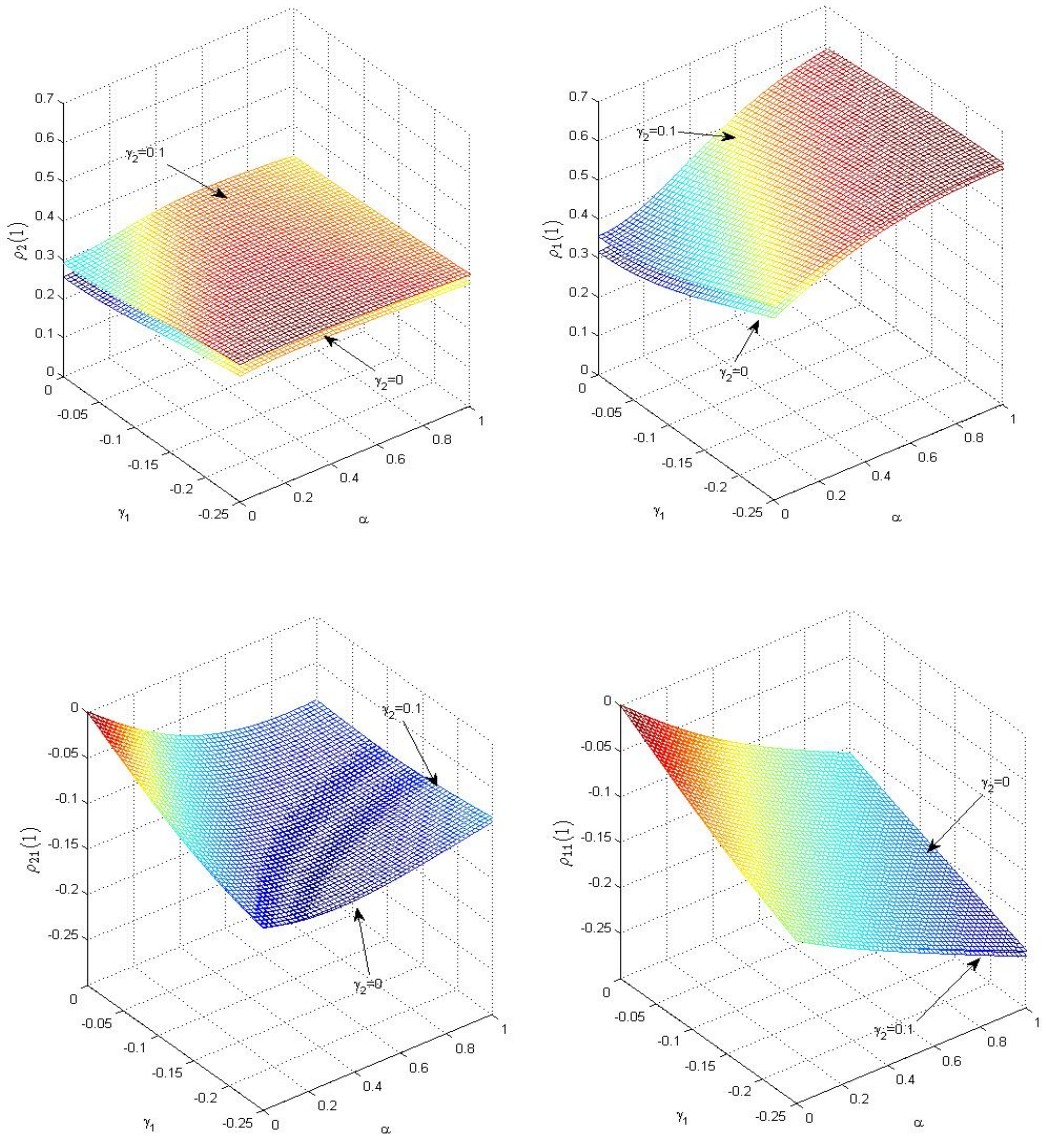

Figure 2.3: First order autocorrelations of squares (top left panel), first order autocorrelations of absolute returns (top right panel), first order cross-correlations between returns and future squared returns (bottom left panel) and first order cross-correlations between returns and future absolute returns (bottom right panel) of different Gaussian T-GASV models with parameters $\phi = 0.98$ and $\sigma_\eta^2 = 0.05$.

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 α or γ_1.

In the second row of Figure 2.3, 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 γ_2 considered in Figure 2.3. Second, for a given value of γ_2, it is obvious that increasing the leverage parameters α 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 α is non-linear. On the other hand, the absolute cross-correlations between returns and future absolute returns have an approximately linear relationship with γ_1 and α 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 2.3 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 ϵ_t on the acf and ccf, Figure 2.4 plots the acf of squared and absolute returns and the ccf between returns and future squared and absolute returns for the 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 acfs of y_t^2 and $|y_t|$ in the first two panels have an exponential decay. Furthermore, fatter tails of ϵ_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 ν of the GED distribution has a very mild influence on the cross-correlations, especially for $\rho_{11}(\tau)$.

To put it briefly, both ν and γ_2 increase the flexibility of the T-GASV model to represent the volatility clustering while have little influence on the leverage effect. On the other hand, γ_1 affects

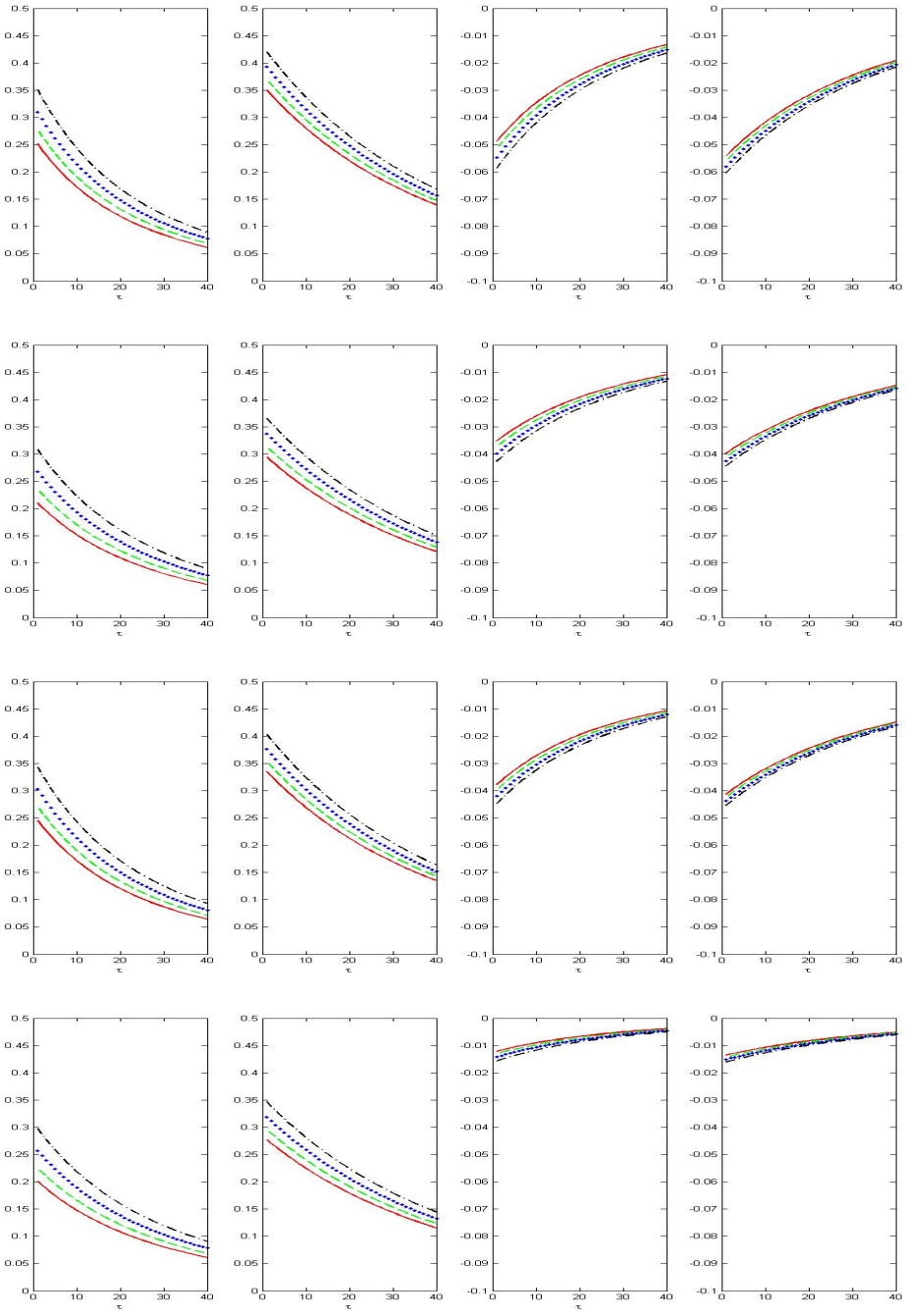

Figure 2.4: First forty orders of the autocorrelations of squares (first column), autocorrelations of absolute returns (second column), cross-correlations between returns and future squared returns (third column) and cross-correlations between returns and future absolute returns (fourth column) for different specifications of asymmetric SV models when $\phi = 0.98$ and $\sigma_\eta^2 = 0.05$. The first row corresponds to a T-GASV model with $\alpha = 0.07, \phi = 0.98, \sigma_\eta^2 = 0.05$, $\gamma_1 = -0.08$, $\gamma_2 = 0.1$ 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 $\gamma_1 = \gamma_2 = 0$.

the leverage effect and this effect is reinforced by the inclusion of α, which could influence slightly the autocorrelations of absolute returns.

2.5 Famous Asymmetric SV models included in the GASV family

In this section, we analyze some of the most popular asymmetric SV models in the literature which are included in the GASV family and can be nested by the T-GASV model, namely, A-ARSV, E-SV and restricted T-SV (RT-SV) models. We obtain the closed-form expressions of their statistical properties from those of the T-GASV model derived in Section 2.4. Some of these properties were previously unknown in the literature. These models are extended by assuming that the return errors follow a GED distribution and compared with one and another in terms of their statistical properties in order to identify their limitations and advantages when used to represent the dynamic properties of the financial returns.

2.5.1 A-ARSV model

One of the most popular SV specifications with leverage effect is the Gaussian A-ARSV model originally proposed by Taylor (1994) and Harvey and Shephard (1996) which specifies the volatility as follows

$$h_t - \mu = \phi(h_{t-1} - \mu) + \eta^*_{t-1}, \tag{2.13}$$

with η^*_t and ϵ_t in the return equation (2.1) being jointly Normal with zero means, variances $\sigma^2_{\eta*}$ and 1, respectively, and correlation δ; see Bartolucci and De Luca (2003), Yu et al. (2006) and Tsiotas (2012)[5] among many others for empirical applications. Define γ_1 and σ^2_η as $\gamma_1 = \delta\sigma_{\eta^*}$ and $\sigma^2_\eta = (1-\delta^2)\sigma^2_{\eta^*}$. Then, the A-ARSV model is equivalent to the following restricted volatility

[5] Tsiotas (2012) allows the return disturbance to follow several asymmetric and fat-tailed distributions.

specification of equation (2.2)

$$h_t - \mu = \phi(h_{t-1} - \mu) + \gamma_1 \epsilon_{t-1} + \eta_{t-1}, \tag{2.14}$$

which is obtained from T-GASV model when ϵ_t is Gaussian and $\alpha = \gamma_2 = 0$; see Asai and McAleer (2011) and Yu (2012) for the equivalence of these two specifications. However, it is important to note that the equivalence between the specifications in (2.13) and (2.14) can only be established when ϵ_t is Normal if the volatility is assumed to be Log-Normal. In this chapter, we focus on the A-ARSV model defined by the equation (2.1) and (2.14) and extend it to allow for fat tails of ϵ_t by assuming that $\epsilon_t \sim GED(\nu)$ distribution.

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). Particularly, the marginal variance and kurtosis of y_t, given in (2.5) and (2.6), reduce to $\sigma_y^2 = \exp(\mu)\exp\left(\frac{\sigma_\eta^2+\gamma_1^2}{2(1-\phi^2)}\right)$ and $k_y = k_\epsilon \exp\left(\frac{\sigma_\eta^2+\gamma_1^2}{1-\phi^2}\right)$, respectively. Note that $\sigma_\eta^2+\gamma_1^2 = \sigma_{\eta*}^2$. As a consequence, several authors conclude that, in the basic Gaussian A-ARSV model, the variance and kurtosis of y_t do not depend on whether there is leverage effect or not; see Taylor (1994), Ghysels et al. (1996) and Harvey and Shephard (1996). One can always find a symmetric model with a larger variance of the errors that has the same variance and kurtosis as a given asymmetric model.

By using the expressions of the statistical properties of the T-GASV model in the previous section, we can also obtain closed-form expressions of the moments of the A-ARSV model when the return errors are GED. As an illustration, Figure 2.4 plots the acfs and ccfs of the A-ARSV model for the same parameter values of the T-GASV model 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 γ_2 and α 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, ϵ_t, are fatter. Once more, the thickness of the tails has very mild influence on the cross-correlations and,

therefore, on the leverage effect.

Next, we consider SNIS of the A-ARSV model which is obtained from (2.9) with $\alpha = \gamma_2 = 0$. The second panel of Figure 2.1 illustrates the SNIS of an A-ARSV model with the same parameters as in the illustration of SNIS of the T-GASV model, i.e., $\{\phi, \gamma_1, \sigma_\eta^2\} = \{0.98, -0.08, 0.05\}$ and $\exp((1-\phi)\mu)\sigma_y^{2\phi} = 1$. Given η_{t-1}, the SNIS$_t$ is an exponential function with exponent γ_1. Thus, bad news generates a higher impact on volatility than good news of the same size. The magnitude of this difference increases with η_{t-1}. Moreover, it is magnified (mitigated) by positive (negative) η_{t-1}. Hence the leverage effect is very weak for negative log-volatility shocks. However, for the particular model considered in Figure 2.1, the leverage effect is very mild when compared with that of the T-GASV model.

2.5.2 Exponential SV model

Consider now the following specification of h_t proposed by Demos (2002) and Asai and McAleer (2011) based on the EGARCH model with an added noise

$$h_t - \mu = \phi(h_{t-1} - \mu) + \gamma_1 \epsilon_{t-1} + \gamma_2 |\epsilon_{t-1}| + \eta_{t-1}, \tag{2.15}$$

where all the parameters and processes are defined and interpreted as in the T-GASV model in (2.10). The model specified by (2.1) and (2.15), denoted as E-SV, can also be obtained by assuming Normality of ϵ_t and $\alpha = 0$ in the T-GASV model.[6] The parameter γ_2 measures the dependence of h_t on past absolute return disturbances in the same form as in the EGARCH model. 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.[7]

Using the results of the T-GASV model from the previous section, we can obtain the properties of the E-SV model when the return errors have a GED distribution. The second row of Figure 2.4 plots the autocorrelations and cross-correlations for an E-SV model with the same parameter

[6] Asai and McAleer (2011) also consider an E-SV model with Student-t return errors.

[7] 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 ϵ_{t-1} in the volatility equation; see Danielsson (1998) and Meyer and Yu (2000). In this case, although the estimation of the parameters is usually easier, the derivation of the properties is harder.

values of the T-GASV model considered in Figure 2.4 except that $\alpha = 0$. Comparing the plots of the A-ARSV and E-SV models in Figure 2.4, we can observe that adding $|\epsilon_{t-1}|$ into the A-ARSV model generates larger autocorrelations of squares and absolute returns but not a 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 2.4 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 ϵ_t, ν, 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. Nevertheless, these two models generate returns with different kurtoses. Therefore, if the parameter ν is a free parameter, we cannot identify the parameters γ_2 and σ_η^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 from Figure 2.4, 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 2.3. Including α 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, α allows for a more flexible pattern of the leverage effect.

Finally, 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 third panel of Figure 2.1. In this case, we can observe that there is not any discontinuity but the effect of ϵ_{t-1} on σ_t still depends on η_{t-1}. Comparing the SNIS of the E-SV model with that of A-ARSV model, we can observe that these two surfaces are similar. We can identify the important role of α in the

response of volatility by comparing the SNIS of the E-SV and T-GASV models. As before, we also plot the NIC of the EGARCH model of Nelson (1991) by considering $\eta_t = 0$.

2.5.3 Threshold SV model

The third popular specification for the volatility considered in this chapter is the Threshold SV (T-SV) model proposed by Breidt (1996) and So et al. (2002) which captures the leverage effect by allowing the parameters of the log-volatility equation to be different depending on the sign of lagged returns. Although its statistical properties are unknown, the T-SV model is rather popular; see, for example, Asai and McAleer (2004, 2005), Muñoz et al. (2007), Chen et al. (2008), Smith (2009), Montero et al. (2010) and Elliott et al. (2011).

In this subsection, we analyze the ability of T-SV models to explain the empirical properties of financial returns. We use simulated data to show that the T-SV model captures asymmetric conditional heteroscedasticity when the constant of the log-volatility equation changes with the sign of lagged returns. However, changes in the persistence parameter and/or in the variance of the log-volatility noise do not guarantee leverage. Therefore, we consider a restricted version of the T-SV model, called Restricted T-SV (RT-SV), in which only the constant changes. We derive its statistical properties and compute the SNIS. Moreover, we extend this RT-SV model by assuming that the return errors follow a GED distribution. The statistical properties of this extended RT-SV models are also analyzed.

Threshold SV model

The T-SV model of Breidt (1996) is given by

$$h_t = \begin{cases} \alpha_1 + \phi_1 h_{t-1} + \sigma_{\eta 1}\eta_{t-1}, & y_{t-1} \geq 0, \\ \alpha_2 + \phi_2 h_{t-1} + \sigma_{\eta 2}\eta_{t-1}, & y_{t-1} < 0, \end{cases} \tag{2.16}$$

where η_t is a standardized Gaussian white noise processes that independent of ϵ_t. The T-SV model introduces the leverage effect by allowing the parameters to change depending on the sign of past

returns. So et al. (2002) consider a T-SV model with $\sigma_{\eta 1} = \sigma_{\eta 2}$.

Deriving the statistical properties of the T-SV model in equations (2.1) and (2.16) is a difficult task.[8] Consequently, we use simulated data to analyse the role that each parameter plays in explaining the relevant statistical properties of financial returns. We consider nine specifications that can be classified into three scenarios. The first scenario includes models with $\phi_1 = \phi_2 = 0.98$ and $\sigma^2_{\eta 1} = \sigma^2_{\eta 2} = 0.05$ while the constant is allowed to change according to the following combinations $\{\alpha_1, \alpha_2\} = \{-0.12, 0.08\}, \{-0.07, 0.05\}$, and $\{-0.14, 0.1\}$. These models are denoted by M1, M2, and M3, respectively. The second category includes models in which $\alpha_1 = \alpha_2 = 0$ and $\sigma^2_{\eta 1} = \sigma^2_{\eta 2} = 0.05$, while the persistence parameter changes according to the following combinations, $\{\phi_1, \phi_2\} = \{0.9, 0.98\}$, $\{0.95, 0.98\}$ and $\{0.6, 0.9\}$. The corresponding models are denoted as M4, M5 and M6, respectively. Finally, the third scenario includes models with $\alpha_1 = \alpha_2 = 0$ and $\phi_1 = \phi_2 = 0.98$ while the variance of log-volatility noise changes according to the following combinations, $\{\sigma^2_{\eta 1}, \sigma^2_{\eta 2}\} = \{0.02, 0.01\}$, $\{0.05, 0.04\}$ and $\{0.05, 0.02\}$. These models are denoted as M7, M8 and M9, respectively. The parameters have been chosen to represent those usually estimated in empirical applications; see So et al. (2002), Asai and McAleer (2005), Muñoz et al. (2007) and Chen et al. (2008). After simulating $R = 1000$ series of size $T = 10000$ from each of the nine models considered, the sample kurtosis, the τ-th order autocorrelations of squares, $\rho_2(\tau)$, and cross-correlations between levels and future squares, $\rho_{21}(\tau)$, are obtained. Table 2.1 reports the corresponding Monte Carlo means and standard deviations for $\tau = 1$.

Consider first the results for models M1, M2 and M3 in which the constant changes. We can observe that the moments of the series generated by these models are close to those often observed when dealing with real financial returns. Moreover, Figure 2.5, that plots the Monte Carlo averages and the 5% and 95% percentiles of the sample autocorrelations of squares and the cross-correlations for the first twenty lags, displays patterns similar to those observed in real data. Particularly, the autocorrelations of squares are all positive and significantly different from zero

[8] The results in Section 2.2 cannot be used because the GASV family does not include models in which ϕ and σ^2_η change.

and if the difference between α_2 and α_1 is large enough, the cross-correlations are significant and negative. Therefore, T-SV models with changes in the constant are able to generate asymmetric conditional heteroscedasticity.

Next, consider the results reported in Table 2.1 for models M4, M5 and M6 in which the autoregressive parameter changes. Observe that the kurtoses and autocorrelations of squares are clearly smaller than before. The magnitude of the first order cross-correlations is also too small to represent the leverage effect often observed in real financial returns; see also Figure 2.6 that illustrates further that when ϕ changes, the generated series do not show significant cross-correlations. Furthermore, in model M6, in which $\phi_1 = 0.6$ and $\phi_2 = 0.9$, even the autocorrelations of squares are barely larger than zero. Therefore, when ϕ changes, the series generated by the T-SV model presents volatility clustering without leverage effect. Moreover, depending on the particular values of ϕ, the series could even be without the volatility clustering.

Finally, consider the results reported in Table 2.1 for models M7, M8 and M9 in which σ_η^2 changes. Observe that these models generate significant autocorrelations of squares and, consequently, volatility clustering. The values of the kurtoses are also rather realistic. However, the cross-correlations are not significantly different from zero; see also Figure 2.7. Therefore, changes in σ_η^2 seem to generate conditionally heteroscedasticity without leverage effect.

Summing up, changes in ϕ and/or σ_η^2 do not pick up the leverage effect, while the threshold in the constant of the log-volatility equation enables the T-SV model to capture conditional heteroscedasticity with leverage effect.

Model	α_1	α_2	ϕ_1	ϕ_2	$\sigma^2_{\eta 1}$	$\sigma^2_{\eta 2}$	Kurtosis	$\rho_2(1)$	$\rho_{21}(1)$
M1	-0.12	0.08	0.98		0.05		13.327 (4.111)	0.266 (0.036)	-0.039 (0.017)
M2	-0.07	0.05	0.98		0.05		11.423 (3.643)	0.257 (0.032)	-0.025 (0.016)
M3	-0.14	0.1	0.98		0.05		14.965 (6.173)	0.271 (0.040)	-0.046 (0.017)
M4	0		0.9	0.98	0.05		4.685 (0.171)	0.141 (0.012)	-0.012 (0.007)
M5	0		0.95	0.98	0.05		6.240 (0.450)	0.195 (0.018)	-0.008 (0.009)
M6	0		0.6	0.9	0.05		3.393 (0.040)	0.040 (0.006)	-0.012 (0.005)
M7	0		0.98		0.02	0.01	4.383 (0.160)	0.133 (0.011)	0.001 (0.006)
M8	0		0.98		0.05	0.04	9.243 (1.587)	0.243 (0.026)	0.001 (0.013)
M9	0		0.98		0.05	0.02	7.214 (0.869)	0.219 (0.022)	0.003 (0.011)

Table 2.1: Monte Carlo means and standard deviations (in parenthesis) of the sample kurtosis, first order autocorrelation of squares and first order cross-correlation between returns and future squared returns.

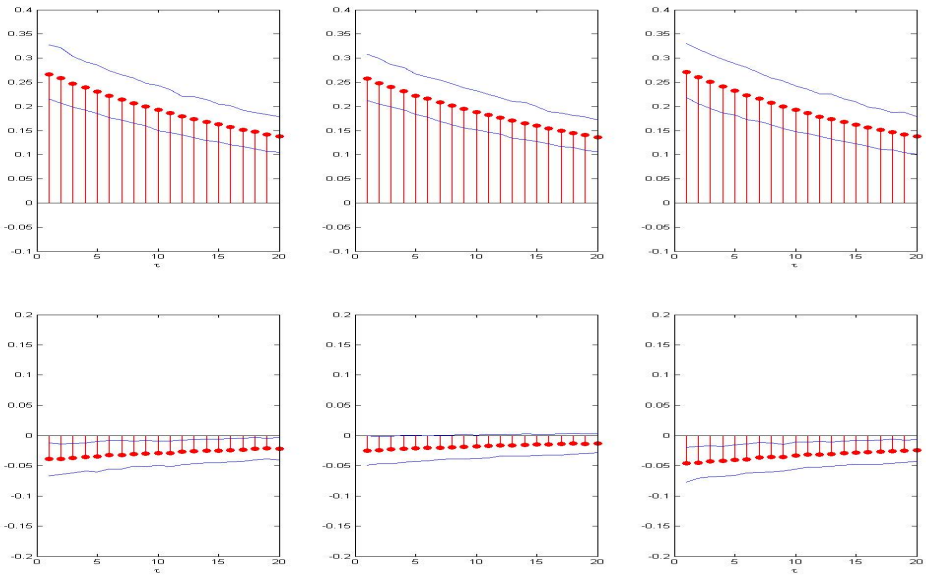

Figure 2.5: Monte Carlo averages and 5% and 95% percentiles (blue lines) of the autocorrelations of squares (first row) and cross-correlations between returns and future squared returns (second row) of models M1 (first column), M2 (middle column) and M3 (last column).

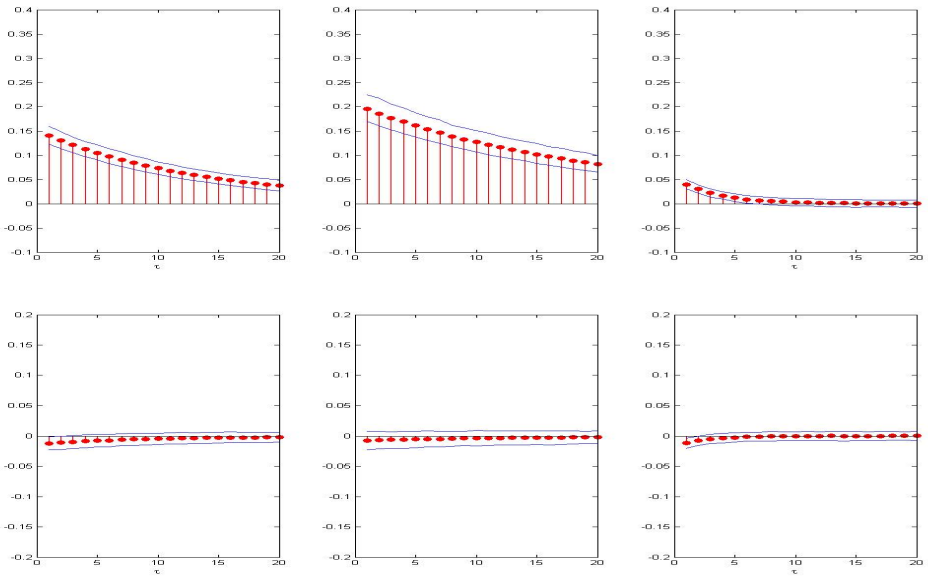

Figure 2.6: Monte Carlo averages and 5% and 95% percentiles (blue lines) of the autocorrelations of squares (first row) and cross-correlations between returns and future squared returns (second row) of models M4 (first column), M5 (middle column) and M6 (last column).

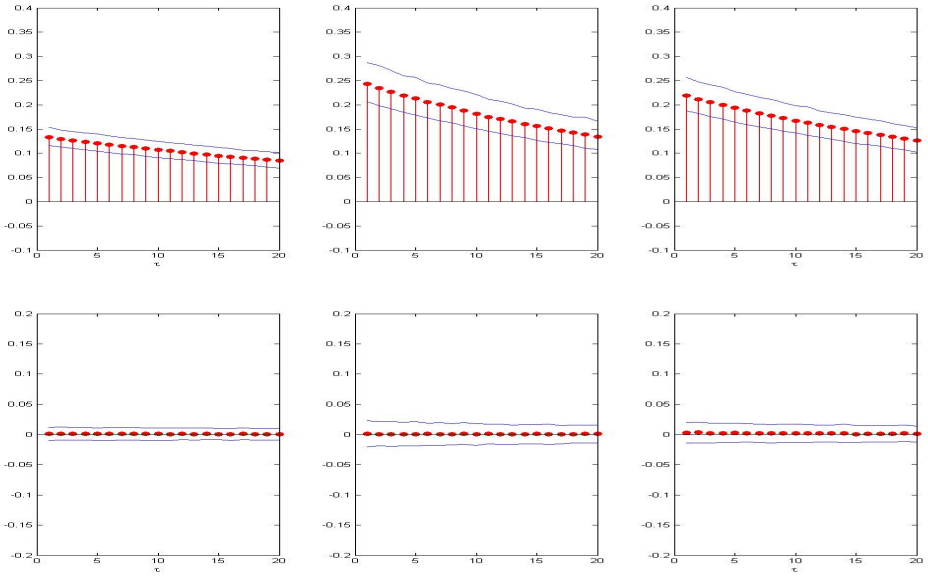

Figure 2.7: Monte Carlo averages and 5% and 95% percentiles (blue lines) of the autocorrelations of squares (first row) and cross-correlations between returns and future squared returns (second row) of models M7 (first column), M8 (middle column) and M9 (last column).

The restricted Threshold SV model

Since, only the threshold in the constant in equation (2.16) allows the model to generate asymmetric conditional heteroscedasticity with volatility clustering, we focus our analysis on the following specification of the volatility, denoted as RT-SV

$$h_t = \mu^* + \alpha I(\epsilon_{t-1} < 0) + \phi h_{t-1} + \sigma_\eta \eta_{t-1}, \tag{2.17}$$

in which the autoregressive parameter and the variance of the log-volatility noise are constant and $\mu^* = \alpha_1$ and $\mu^* + \alpha = \alpha_2$. This specification is included in the GASV family and has been previously considered by Asai and McAleer (2006).

The Stochastic News Impact Surface According to the definition of SNIS in Section 2.3, the SNIS of the RT-SV is given by

$$SNIS_t = \exp(\mu^*)\sigma_y^{2\phi} \exp(\alpha I(\epsilon_{t-1} < 0) + \sigma_\eta \eta_{t-1}), \tag{2.18}$$

where σ_y^2 is the marginal variance of y_t which can be easily obtained from the equation (2.5) given that $\mu = \frac{\mu^*}{1-\phi}$ and

$$P(b_i) \equiv \prod_{i=1}^{\infty} \frac{1}{2} \left[\exp(b_i \alpha) + 1\right]. \tag{2.19}$$

Figure 2.1 plots at the bottom panel the SNIS of RT-SV model with parameters $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ and $\alpha = 0.7$. The value of μ^* is chosen such that $\exp(\mu^*)\sigma_y^{2\phi} = 1$. The main characteristic of the SNIS is its discontinuity with respect to ϵ_{t-1}. Furthermore, it represents different leverage effects depending on the value of the log-volatility disturbance.

RT-SV model with Gaussian errors The statistical properties of the Gaussian RT-SV model can be obtained from the results of the T-GASV model in Section 2.4. Note that, given that ϵ_t is

Gaussian, the conditions for Theorem 2.1, Theorem 2.2 and Theorem 2.3 in Section 2.2 are satisfied so that $E(\exp(0.5c\alpha I(\epsilon_t < 0))) = \frac{1}{2}(1 + \exp(0.5c\alpha)) < \infty$, $E(\epsilon_t^c) < \infty$ and $E(\epsilon_t^{2c}) < \infty$ for any positive integer c. Therefore, when $|\phi| < 1$, the RT-SV model is stationary and the moments of $|y_t|$, the autocorrelations of $|y_t|^c$ and the cross-correlations between y_t and $|y_{t+\tau}|^c$ for $\tau > 0$ are all finite. Furthermore, the odd moments of y_t are always zero.

All its closed-form statistical properties can be obtained from those of the T-GASV model derived in Section 2.4 by restricting that $\mu = \frac{\mu^*}{1-\phi}$ and $\gamma_1 = \gamma_2 = 0$, given that, when $\epsilon_t \sim N(0,1)$, $E|\epsilon_t|^c = \frac{2^{c/2}}{\sqrt{\pi}}\Gamma\left(\frac{c+1}{2}\right)$, $P(b_i)$ given in the equation (2.19) and $T(n, b_i) \equiv \prod_{i=1}^{n-1} \frac{1}{2}[\exp(b_i\alpha) + 1]$ if $n > 1$ while $T(1, b_i) \equiv 1$.

In order to illustrate the shape of the autocorrelations of the squared returns generated by the RT-SV model, the left panel of Figure 2.8 plots them for RT-SV models with parameters $\mu^* = 0$, $\phi = 0.98$, $\sigma_\eta^2 = 0.05$ and $\alpha = 0, 0.1$ and 0.2. Observe that the value of α barely has influence on the autocorrelations of squares which are very similar to those in Figure 2.5 for the simulated data.

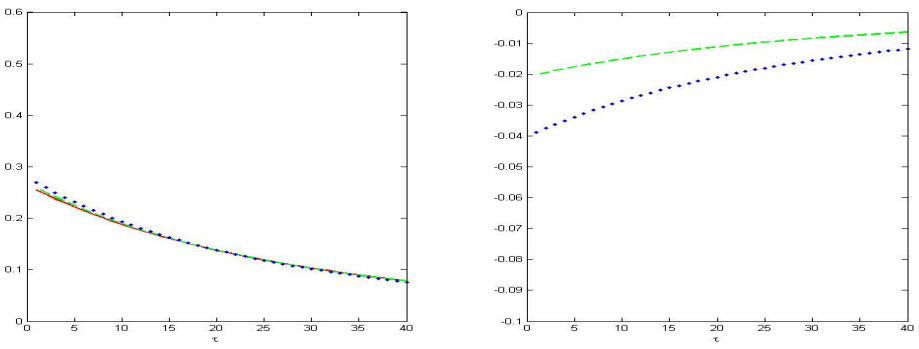

Figure 2.8: Autocorrelations of squares (left column) and cross-correlations between returns and future squared returns (right column) for Gaussian RT-SV models with $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\mu^* = 0$ and $\alpha = 0$ (solid lines), $\alpha = 0.1$ (dashed lines) and $\alpha = 0.2$ (dotted lines).

The cross-correlations between returns and future squared returns for the same RT-SV models considered above are plotted in the right panel of Figure 2.8. We observe that larger values of α generate returns with larger leverage effect. Furthermore, the magnitude of the cross-correlations is very close to that of the simulated ones plotted in Figure 2.5 for models M1 and M2. This

confirms the Monte Carlo results about changes on α capturing the leverage effect without destroying the volatility clustering.

RT-SV model with GED error The RT-SV model considered above can be extended to assume a GED distribution for the return errors. Once more, the statistical properties of the RT-SV model with GED errors can be obtained using the results in Section 2.4. The last row of Figure 2.4 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 ϕ, σ_η^2 and ν as those considered for the T-GASV model in Figure 2.4. Comparing the autocorrelations of squares and absolute returns of the T-GASV model represented in the top panel of Figure 2.4 and those of the RT-SV model with GED return errors, 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 α in the T-GASV model is reinforcing the role of the leverage parameter γ_1.

2.6 MCMC estimation and empirical results for GASV models

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.[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, 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 BUGS 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. There are two main versions of BUGS, namely WinBUGS and OpenBUGS. WinBUGS is an established and stable, stand-alone version, which is not further developed. In this thesis, we adopt OpenBUGS that is still being updated.

In this section, we describe briefly the algorithm of the MCMC estimator for estimating the T-GASV model with restriction $\gamma_2 = 0$ and $\epsilon_t \sim GED(\nu)$, denoted as RT-GASV. Recall that in Section 2.5, we show that one possible problem is the parameter identification. For a T-GASV model with parameter $\nu = \nu_0$, we may find another model with $\nu \neq \nu_0$ and different parameter values that represents the same dynamics of $|y_t|^c$. This might be due to the fact that the parameters γ_2 and ν do the same job that allow the T-GASV model to capture more volatility clustering. Therefore, we focus on the RT-GASV model where $\gamma_2 = 0$.

[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 based on the second order Taylor expansion of the density function. Finally Liesenfeld and Richard (2003) propose a Maximum Likelihood approach based upon an efficient importance sampling.

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 the RT-GASV model. Moreover, we also investigate that, by fitting our RT-GASV model to the series generated from those nested asymmetric SV models, whether it is able to identify the true Data Generating Process (DGP).

Finally, the MCMC estimator is implemented to estimate the volatilities and the Value at Risk (VaR) of the series of daily S&P500 returns after fiting all the asymmetric SV models considered in this chapter.

2.6.1 Finite sample performance of a MCMC estimator for Threshold GASV model

Next, we describe briefly the algorithm. Let $p(\boldsymbol{\theta})$ be the joint prior distribution of the unknown parameters $\boldsymbol{\theta} = \{\mu, \phi, \alpha, \gamma_1, \sigma_\eta^2, \nu\}$. Following Meyer and Yu (2000), the prior densities of ϕ and σ_η^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.[10] The remaining prior densities are chosen to be 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 $\boldsymbol{\theta}$ and $\boldsymbol{h}$ is given by

$$p(\boldsymbol{\theta}, \boldsymbol{h}) = p(\boldsymbol{\theta})p(h_0)\prod_{t=1}^{T+1} p(h_t|h_{t-1}, \boldsymbol{\theta}). \tag{2.20}$$

The likelihood function is then given by

$$p(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{h}) = \prod_{t=1}^{T} p(y_t|h_t, \boldsymbol{\theta}). \tag{2.21}$$

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

[10] Although the prior of ϕ^* is very informative, when it is changed to $Beta(1, 1)$, the results are very similar.

given by

$$y_t|u,h_t \sim U\left(-\frac{\exp(h_t/2)}{\sqrt{2\Gamma(3/\nu)/\Gamma(1/\nu)}}u^{1/\nu}, \frac{\exp(h_t/2)}{\sqrt{2\Gamma(3/\nu)/\Gamma(1/\nu)}}u^{1/\nu}\right), \tag{2.22}$$

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

$$\begin{aligned}
&\theta_1^{(1)} \sim p(\theta_1|\theta_2^{(0)},\ldots,\theta_K^{(0)},\boldsymbol{h}^{(0)},\boldsymbol{y});\\
&\vdots\\
&\theta_K^{(1)} \sim p(\theta_1|\theta_2^{(1)},\ldots,\theta_{K-1}^{(1)},\boldsymbol{h}^{(0)},\boldsymbol{y});\\
&h_1^{(1)} \sim p(h_1|\boldsymbol{\theta}^{(1)},h_2^{(0)},\ldots,h_{T+1}^{(0)},\boldsymbol{y});\\
&\vdots\\
&h_{T+1}^{(1)} \sim p(h_{T+1}|\boldsymbol{\theta}^{(1)},h_1^{(1)},\ldots,h_T^{(1)},\boldsymbol{y}).
\end{aligned}$$

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

$$p(\boldsymbol{\theta},\boldsymbol{h}|\boldsymbol{y}) \propto p(\boldsymbol{\theta})p(\boldsymbol{h}_0)\prod_{t=1}^{T+1}p(h_t|h_{t-1},\boldsymbol{y},\boldsymbol{\theta})\prod_{t=1}^{T}p(y_t|h_t,\boldsymbol{\theta}). \tag{2.23}$$

We consider two designs for the Monte Carlo experiments. First, R replicates are generated by the RT-GASV model with parameters $\{\mu,\phi,\alpha,\gamma_1,\sigma_\eta^2,\nu\} = \{0, 0.98, 0.07, -0.08, 0.05, 1.5\}$. 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 2.2 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.

	μ	ϕ	α	γ_1	σ_η^2	ν
True	**0**	**0.98**	**0.07**	**-0.08**	**0.05**	**1.5**
T=500						
Mean	0.268	0.952	0.108	-0.074	0.083	1.581
	(1.445)	(0.063)	(0.126)	(0.077)	(0.068)	(0.369)
s.d.	1.803	0.018	0.121	0.066	0.034	0.184
T=1000						
Mean	0.046	0.974	0.077	-0.082	0.055	1.520
	(1.442)	(0.010)	(0.073)	(0.041)	(0.020)	(0.132)
s.d.	1.779	0.009	0.078	0.042	0.017	0.112
T=2000						
Mean	0.082	0.977	0.072	-0.082	0.053	1.528
	(1.278)	(0.006)	(0.056)	(0.031)	(0.013)	(0.098)
s.d.	1.422	0.006	0.057	0.030	0.011	0.071

Table 2.2: 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.

Second, we also want to check whether by fitting the RT-GASV model we are able to identify the true restricted specifications. With this purpose, we generate $R = 200$ replicates of size $T = 1000$ from each of the restricted models, A-ARSV and RT-SV, with the distribution parameter $\nu = 2$ or $\nu = 1.5$ and fit the RT-GASV model. The results, reported in Table 2.3, provide evidence that it is possible to identify the true data generating process (DGP) by fitting the more general RT-GASV model.

	A-ARSV						RT-SV					
	μ	ϕ	α	γ_1	σ_η^2	ν	μ	ϕ	α	γ_1	σ_η^2	ν
True	**0**	**0.98**	**0**	**-0.08**	**0.05**	**2**	**0**	**0.98**	**0.07**	**0**	**0.05**	**2**
Mean	0.068	0.974	-0.005	-0.082	0.055	2.022	0.145	0.972	0.077	-0.002	0.056	2.076
	(1.377)	(0.009)	(0.069)	(0.037)	(0.018)	(0.200)	(1.531)	(0.011)	(0.080)	(0.042)	(0.019)	(0.201)
s.d.	1.722	0.009	0.078	0.042	0.016	0.200	1.720	0.010	0.080	0.042	0.016	0.188
True	**0**	**0.98**	**0**	**-0.08**	**0.05**	**1.5**	**0**	**0.98**	**0.07**	**0**	**0.05**	**1.5**
Mean	0.140	0.974	-0.008	-0.088	0.053	1.478	0.003	0.972	0.081	0.002	0.058	1.525
	(1.431)	(0.010)	(0.075)	(0.043)	(0.018)	(0.140)	(1.461)	(0.013)	(0.075)	(0.048)	(0.022)	(0.128)
s.d.	1.767	0.009	0.077	0.042	0.017	0.117	1.752	0.010	0.080	0.042	0.018	0.113

Table 2.3: Monte Carlo results of MCMC estimator of the parameters of the RT-GASV model fitted to series simulated from different asymmetric SV models. 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.

Summarizing the Monte Carlo results on the MCMC estimator considered in this chapter, 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 RT-GASV model.

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 estimation 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 2.4 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 $T = 1000$. 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 RT-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.

	RT-GASV	A-ARSV		RT-SV	
		$\nu = 2$	$\nu = 1.5$	$\nu = 2$	$\nu = 1.5$
Mean	-0.027	-0.030	-0.040	-0.022	-0.022
	(0.016)	(0.040)	(0.053)	(0.018)	(0.021)
s.d.	0.232	0.216	0.238	0.213	0.229

Table 2.4: Monte Carlo results of the relative volatility estimation errors. Reported are the values of the time average and standard deviation (in parenthesis) of $m_t = \sum_{i=1}^{R} e_t^{(i)}/R$ together with the time average of $s_t = \sqrt{\sum_{i=1}^{R}(e_t^{(i)} - m_t)^2/(R-1)}$, where $e_t^{(i)} = (\sigma_t^{(i)} - \hat{\sigma}_t^{(i)})/\sigma_t^{(i)}$.

2.6.2 Empirical application

Estimation results

In this subsection, the RT-GASV model is fitted to represent the dynamic dependence of the daily S&P500 returns described in Chapter 1. It is clear that the volatility clustering and leverage effect are present in this series. Consequently, the RT-GASV model is fitted first assuming GED errors 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 ϵ_t. For completeness, we also fit the other two restricted models. All the parameters and volatilities have been estimated implementing the MCMC estimator of BUGS.

To compare two competitive models, saying M_0 and M_1, we consider the Bayes Factor (BF). The BF, which is defined as the ratio of the marginal likelihood values of two competing models, $\frac{p(y|M_0)}{p(y|M_1)}$, where $p(y|M_k)$ is the marginal likelihood of model k with $k = 0, 1$. If the prior odds ratio is 1 by Bayes' theorem, the posterior odds ratio takes the same value as the BF. Jeffreys (1961) gave a scale for the interpretation of BFs. If $ln(BF)$ is less (bigger) than 0, there is evidence in favor of (against) M_1. Moreover, if $ln(BF) \in (0, 1)$, the evidence against M_1 is barely worth mention; if $ln(BF) \in (1, 3)$, the evidence against M_1 is positive; if $ln(BF) \in (1, 3)$ (0r $(3, \infty)$), the evidence against M_1 is strong (or very strong).

Table 2.5 reports the posterior mean and the 95% credible interval of the MCMC estimator of

	$\epsilon_t \sim GED(\nu)$			$\epsilon_t \sim N(0,1)$		
	RT-GASV	A-ARSV	RT-SV	RT-GASV	A-ARSV	RT-SV
μ	-0.035	0.084	-5.638	-0.477	-0.007	-5.647
	(-0.019,0.079)	(-0.017,0.222)	(-6.551,-4.276)	(-0.891,0.519)	(-0.133,0.207)	(-6.910,-4.184)
ϕ	0.982	0.980	0.984	0.979	0.981	0.980
	(0.974,0.993)	(0.973,0.990)	(0.974,0.992)	(0.969,0.987)	(0.973,0.992)	(0.969,0.989)
α	0.035		0.193	0.019		0.236
	(-0.019,0.079)		(0.169,0.224)	(-0.026,0.051)		(0.189,0.271)
γ_1	-0.129	-0.143		-0.145	-0.143	
	(-0.155,-0.103)	(-0.162,-0.126)		(-0.172,-0.117)	(-0.168,-0.125)	
σ_η^2	0.009	0.009	0.012	0.017	0.016	0.021
	(0.000,0.018)	(0.002,0.019)	(0.005,0.022)	(0.005,0.027)	(0.005,0.023)	(0.010,0.032)
ν	1.359	1.391	1.419			
	(1.237,1.382)	(1.309,1.365)	(1.344,1.423)			
Log-Likelihood	-5688.8663	-5689.4757	-5689.9409	-5590.4776	-5590.4774	-5595.0947

Table 2.5: 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.

each parameter. The left panel reports the results of those models with GED errors while the right panel for the models with Normal errors. Checking the results of the models with GED errors, we can observe that when the RT-GASV model is fitted, the credible interval for the threshold parameter α contains the zero. The Monte Carlo experiments in the previous section suggest that fitting the general RT-GASV model with GED errors proposed in this paper, one could identify the true restricted specification of the log-volatilities. Consequently, it seems that the threshold parameter is not needed to represent the conditional heteroscedasticity of the S&P500 returns. Second, the credible interval of the estimate of distribution parameter ν excludes the value 2 which, according to our Monte Carlo results, indicates that models with GED errors outperform the counterparts with Gaussian errors. Finally, the log-Likelihoods of all the three models are very close which indicates similar in-sample performance no matter which distribution is assumed to the return errors. Figure 2.9 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 Gaussian errors.

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

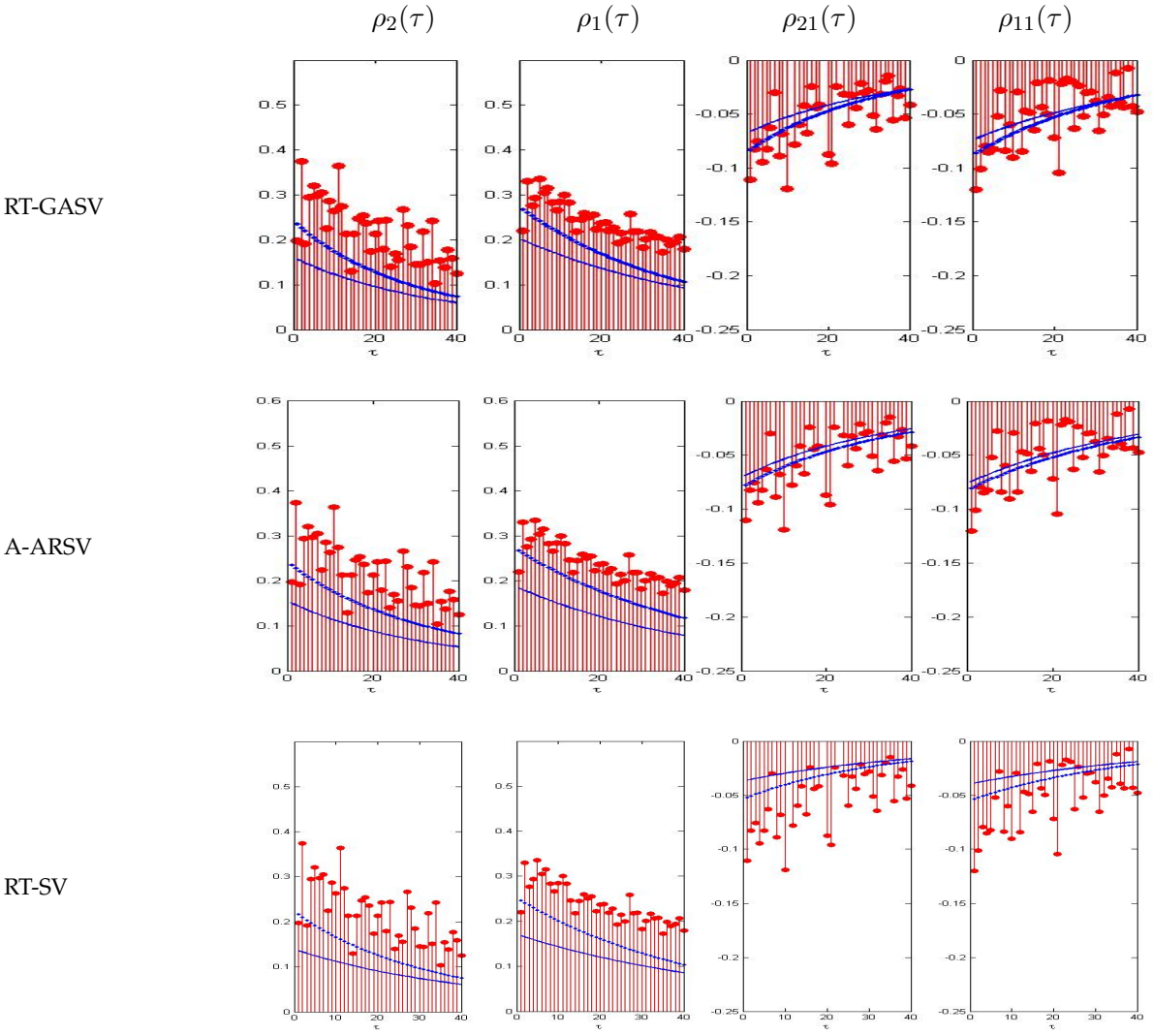

Figure 2.9: Sample autocorrelations of squares (first column), autocorrelations of absolute returns (second column), cross-correlations of returns and future squared returns (third column) and cross-correlations between absolute returns and lagged returns (fourth column) together with the corresponding 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.

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 ϵ_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.[11] Moreover, one-step-ahead VaRs are obtained starting on January 2, 2014 until July 25, 2014 as

$$VaR_{t+1|t}(m) = q\hat{\sigma}_{t+1|t}, \tag{2.24}$$

with q being the 5% quantile of the distribution with parameter ν 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 (2.24) for each of the models, Table 2.6 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 the level 0.05. Therefore, our RT-GASV model with GED error provides the best prediction of volatilities for this S&P500 return series.

		Failure Rate
$\epsilon_t \sim GED(\nu)$	A-ARSV	0.056
	RT-SV	0.085
	RT-GASV	0.049
$\epsilon_t \sim N(0,1)$	A-ARSV	0.092
	RT-SV	0.099
	RT-GASV	0.070

Table 2.6: Failure rates.

2.7 Conclusions

In this chapter, 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

[11] Checking the estimates obtained, we observe that all the estimates are very stable over the year considered in the rolling window estimation.

modeling heteroscedastic series with leverage effect can be included within the GASV family. We propose a new model named T-GASV which belongs to the GASV family and nests these particular specifications. 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 nested by the T-GASV model. Closed-form expressions of the statistical properties of T-GASV model are obtained. 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.

As a marginal outcome, we are able to obtain the statistical properties of those nested models, some of which were previously unknown. We find that, first, the parameter γ_2 in E-SV model allows to capture more volatility clustering than the A-ARSV model. Furthermore, by adding the threshold parameter α, the T-GASV model adds flexibility to capture the leverage effect. Finally, the degrees of freedom of the GED errors enforce the model's flexibility to capture volatility clustering. The ability of the T-SV model to explain the empirical properties of financial returns is also analyzed. Through extensive simulation studies, we show that allowing the autoregressive parameter and/or the variance of the log-volatility disturbance to be different depending on the sign of past returns do not generate leverage effect. However, changing the constant in the volatility equation allows the model to capture asymmetric conditional heteroscedasticity with volatility clustering. We derive the analytical properties of the T-SV model in which only the constant changes, named as RT-SV. It is found that the RT-SV model generates returns with smaller autocorrelations and absolute cross-correlations than the T-GASV model with the same values of parameters.

Another contribution of this chapter 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 RT-GASV model using the BUGS software. We show first that the parameters and volatilities of the RT-GASV model can be estimated appropriately. Second, fitting the proposed RT-GASV model allows to correctly identify the true data generating process. Finally, the RT-GASV model as well as its nested models, A-ARSV and RT-SV, are fitted to estimate the volatilities of S&P500 daily returns. For this particular data set, all the models with GED errors provide similar in-sample performance and better than their counterparts with Gaussian return errors. When estimating the VaRs our RT-GASV model with GED errors outperforms the benchmarks considered.

Chapter 3

Score Driven Asymmetric SV models

3.1 Introduction

It is well acknowledged that the standardized financial returns are heavy-tailed distributed. In order to capture this feature, the SV models have been extended by assuming fat-tailed return errors, for instance, the E-SV model with Student-t distribution of Asai and McAleer (2011). However, in some of the traditional asymmetric SV models, the volatility is specified as being driven by the past return error. Therefore, when the return errors are fat-tailed, the traditional asymmetric SV models could attribute a large realisation of the return errors to an increase in volatility. In this chapter, we propose a new class of asymmetric SV models, which specifies the volatility as a function of the score of the lagged return distribution as in the Generalized Autoregressive Score (GAS) model of Creal et al. (2013). The score-driven models can automatically correct for the influential observations which are judged as outliers by the Gaussian yardstick. We propose three score-driven SV models, namely, GAS^{2}V-N, GAS^{2}V-T, and GAS^{2}V-G corresponding to the return errors following either the Normal, Student-t or the GED distribution. The closed-form expressions of their statistic properties are derived.

We show that the MCMC procedure described in Section 2.6 can estimate the parameters of some restricted score driven SV models adequately. Finally, the models are fitted both daily and

weekly financial returns and evaluated in terms of their in-sample and out-of-sample performance.

3.2 Score driven asymmetric SV models

In this section, we propose the GAS^{2}V model and derive its statistical properties when the errors are distributed as Normal, Student-t and GED. In particular, we obtain the closed-form expressions of the marginal variance, the kurtosis, acf of power-transformed absolute returns and the ccf between returns and future power-transformed absolute returns.

3.2.1 The GAS^{2}V models

Let y_t be modeled as in the equation (2.1). The GAS^{2}V specifies the volatility as

$$h_t - \mu = \phi(h_{t-1} - \mu) + f(u_{t-1}) + \eta_{t-1}, \tag{3.1}$$

where η_t is a Gaussian white noise with variance σ_η^2 and ϵ_t is a strict white noise with variance one which is distributed independently of η_t for all leads and lags. μ is a scale parameter related with the marginal variance of returns while the parameter ϕ is related with the persistence of the volatility shocks. Finally, $f(\cdot)$ is a function of the scaled conditional score of the lagged return, u_{t-1}, which is defined as follows

$$u_t = C\frac{\partial ln P(y_t|h_t)}{\partial h_t}, \tag{3.2}$$

where C is any real number introduced to simplify the expression of the score and $P(y_t|h_t)$ is the density of returns conditional on volatilities. Denoting by $\psi(\epsilon_t)$ the probability density function (pdf) of ϵ_t, the density function of y_t conditional on h_t is given by $P(y_t|h_t) = \exp(-h_t/2)\psi(y_t \exp(-h_t/2))$. It follows immediately that

$$u_t = -\frac{C}{2} + \frac{C}{2}\frac{\epsilon_t \psi'(\epsilon_t)}{\psi(\epsilon_t)}, \tag{3.3}$$

where $\psi'(\epsilon_t)$ denotes the derivative of $\psi(\epsilon_t)$ with respect to ϵ_t. Thus, u_t depends on ϵ_t and, consequently, after writing $f(u_{t-1}) = f\left(-\frac{C}{2} + \frac{C}{2}\frac{\epsilon_t\psi'(\epsilon_t)}{\psi(\epsilon_t)}\right)$ in equation (2.2), the GAS^{2}V model in equations (2.1) and (3.1) can be obtained as a particular case of the GASV family defined in Chapter 2 and the results on the properties of this family can be directly used. In particular, according to Theorem 2.1, when $|\phi| < 1$ and the distribution of ϵ_t is such that $E(\exp(f(\epsilon_t))) < \infty$, the GAS^{2}V model is stationary. Moreover, for any non-negative integer c, if the distribution of ϵ_t is such that $E(\exp(0.5cf(\epsilon_t))) < \infty$ and $E(|\epsilon_t|^c) < \infty$, both y_t and $|y_t|$ have finite moments of order c. The autocorrelation function of $|y_t|^c$ is also finite. Finally, the finiteness of the cross-correlation function between y_t and $|y_{t+\tau}|^c$, for $\tau = 1, 2, \cdots$, is guaranteed when further $E(|\epsilon_t|^{2c}) < \infty$. The general expressions of these moments are given by Theorem 2.1, Theorem 2.2 and Theorem 2.3.

Later in this chapter, we obtain closed-form expressions of these moments for particular assumptions on the function $f(\cdot)$ and on the error distribution. In particular, in order to represent the leverage effect often observed when dealing with time series of financial returns, we consider the following specification of $f(\cdot)$

$$f(u_{t-1}) = \alpha I(\epsilon_{t-1} < 0) + ku_{t-1} + k^* sign(-\epsilon_{t-1})(u_{t-1} + 1), \tag{3.4}$$

where $I(\cdot)$ is an indicator function that takes value one when the argument is true and zero otherwise. The parameter k represents an ARCH effect while the parameters α and k^* represent the leverage effect with α dealing with changes in the scale parameter depending on the sign of past returns and k^* with changes in the dynamics involving the score. Note that the last term in (3.4) is based on the proposal of Harvey (2013) in the context of asymmetric score GARCH models. As pointed out by Harvey (2013), although the statistical validity of the model does not require it, proper restriction may be imposed on k and k^* in order to ensure that an increase in the absolute value of a standardized observation does not lead to a decrease in volatility.

Finally, the SNIS of GAS^{2}V model is given by

$$SNIS_t = \exp((1-\phi)\mu)\sigma_y^{2\phi}\exp(f(u_{t-1}) + \eta_{t-1}), \tag{3.5}$$

where σ_y^2 is the marginal variance of y_t and $f(u_{t-1})$ is given in (3.4). It is important to note that the score, u_t, is different depending on the particular assumption on the error distribution. Several distributions of return errors have been proposed in the related literature being the Gaussian distribution the most popular; see, for example, Jacquier et al. (1994) and Harvey and Shephard (1996). When the errors are Gaussian, the score is given by

$$u_{t-1} = \epsilon_{t-1}^2 - 1. \tag{3.6}$$

The corresponding SNIS is plotted in the top panel of Figure 3.1 when the GAS^{2}V model has parameters $\{\alpha, \phi, k^*, k, \sigma_\eta^2\} = \{0.07, 0.98, 0.08, 0.1, 0.05\}$. The scale parameter, μ is chosen so that $\exp((1-\phi)\mu)\sigma_y^{2\phi} = 1$. It shows that the volatility response is larger when the lagged return is negative than when it is positive. Therefore, this model is able to capture the leverage effect. Moreover, the difference in the response of the volatility to positive and negative ϵ_{t-1} depends on the log-volatility noise, η_{t-1}. Stronger leverage effect is observed when η_{t-1} is positive and large. The News Impact Curve (NIC), defined by Engle and Ng (1993), is obtained when $\eta_{t-1} = 0$, which is also plotted in Figure 3.1. The inclusion of η_{t-1} in the model allows it to be more flexible in representing the leverage effect.

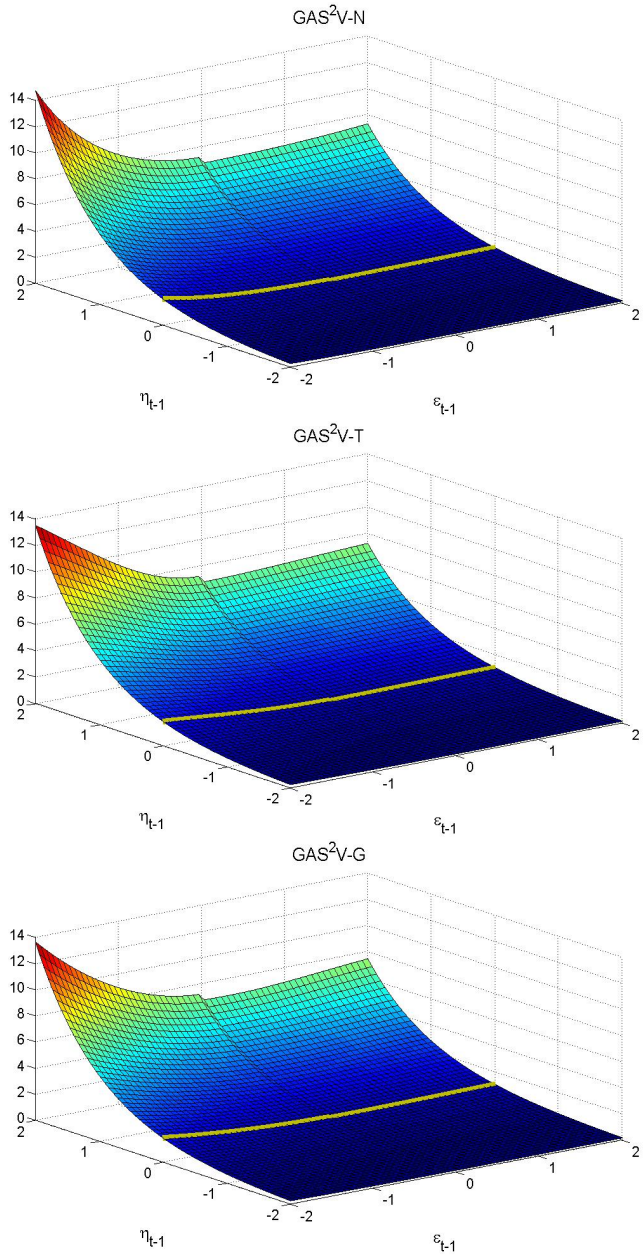

Figure 3.1: SNIS of GAS^{2}V-N (top panel) with parameters $(\alpha, \phi, k^*, k, \sigma_\eta^2) = (0.07, 0.98, 0.08, 0.1, 0.05)$ and $\exp((1-\phi)\mu)\sigma_y^{2\phi} = 1$, GAS^{2}V-T (middle panel) with $\nu_0 = 6$ and GAS^{2}V-G (bottom panel) with $\nu = 1.5$

However, the Gaussian distribution does not fully capture the fat tails of financial time series often observed in practice and may suffer from a lack of robustness in the presence of extreme outlying observations. Consequently, several authors consider heavy-tailed distributions such as

the Student-t or the GED distributions;[1] see, for example, Chen et al. (2008), Choy et al. (2008) and Wang et al. (2011, 2013). Consider first the GAS^{2}V model when ϵ_t has a Student-t distribution with ν_0 degrees of freedom. In this case, the score is given by

$$u_t = (\nu_0 + 1)\frac{\epsilon_t^2}{\nu_0 - 2 + \epsilon_t^2} - 1. \tag{3.7}$$

The SNIS of the GAS^{2}V model with Student-t errors is plotted in the middle panel of Figure 3.1 for the same parameters as above and $\nu_0 = 6$. The asymmetric response of volatility to ϵ_{t-1} is similar to that of the GAS^{2}V model with Gaussian errors.

Finally, when ϵ_t is assumed to follow a GED(ν) distribution, then the score function is given by

$$u_t = \frac{\nu}{2}\left|\frac{\epsilon_t}{\varphi}\right|^{\nu} - 1, \tag{3.8}$$

with $\varphi = \sqrt{2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/\nu)}$. The SNIS of the GAS^{2}V model with GED errors when $\nu = 1.5$ is plotted in the bottom panel of Figure 3.1. The volatility responds asymmetrically to the positive and negative returns errors. However, no big difference can be observed among the SNISs of all the three GAS^{2}V models.

3.2.2 Different GAS^{2}V models

In this subsection, we analyze the properties of three GAS^{2}V models corresponding to three different return error distributions.

[1] There are also proposals to include simultaneously leptokurtosis and skewness in the distribution of ϵ_t, such as the skewed-Normal and skew-Student-t in Nakajima and Omori (2012) and the asymmetric GED in Cappuccio et al. (2004). It is not straightforward to capture the moments of returns when the distribution of ϵ_t is asymmetric. Consequently, we leave this extension for future research and focus on symmetric distributions.

GAS^{2}V-N

If ϵ_t follows a Gaussian distribution, then, the scaled score, u_t is given by expression (3.6) and the specification of the log-volatility with $f(\cdot)$ defined as in (3.4) reduces to

$$h_t - \mu = \phi(h_{t-1} - \mu) + \alpha I(\epsilon_{t-1} < 0) + k(\epsilon_{t-1}^2 - 1) + k^* sign(-\epsilon_{t-1})\epsilon_{t-1}^2 + \eta_{t-1}. \tag{3.9}$$

The resulting model is denoted as GAS^{2}V-N. It is important to note that although the specification of the volatility in (3.9) is closely related to that in the T-GASV model in Chapter 2, the way in which the leverage is introduced is different in both cases. In (3.9), the log-volatility depends on squared returns and the leverage effect is introduced in the same fashion as in the TGARCH model of Zakoian (1994). However, the log-volatility in the T-GASV model depends on past absolute returns and the leverage is introduced as in the EGARCH model. Rodríguez and Ruiz (2012) show that the TGARCH and EGARCH models are very similar. Therefore, we expect that, if ϵ_t is Gaussian, the GAS^{2}V-N and T-GASV models have very similar properties. The analytical expressions of $E(\epsilon_t^c \exp(bf(\epsilon_t)))$ and $E(|\epsilon_t|^c \exp(bf(\epsilon_t)))$ are given in Appendix B.1.1. Using these expressions we can verify that when $|\phi| < 1$ and $k + |k^*| < 1/2$, the model is stationary, y_t and $|y_t|$ have finite moments of order c and the acf of $|y_t|^c$ and ccf between y_t and $|y_{t+h}|^c$ are finite when $ck + |ck^*| < 1$.

We first explore the kurtosis of the GAS^{2}V-N model. It is the kurtosis of the ARSV(1) model proposed by Harvey et al. (1994), $k_\epsilon \exp\left(\frac{\sigma_\eta^2}{1-\phi^2}\right)$, multiplying the factor $r = \frac{\prod_{i=1}^{\infty} E\left(2\phi^{i-1} f(u_{t-i})\right)}{\prod_{i=1}^{\infty} E^2(\phi^{i-1} f(u_{t-i}))}$. As an illustration, Figure 3.2 plots R as a function of the leverage parameters α and k^* when $k = 0$ and 0.1 for three different persistence parameters, namely, $\phi = 0.5$, 0.9 and 0.98. For these particular parameter values, we can observe that the ratio is always larger than 1. Therefore, the GAS^{2}V-N model generates returns with larger kurtosis than the corresponding basic ARSV(1). Furthermore, the kurtosis increases with α, k^* and k. The increment is more prominent when ϕ is larger.

In order to illustrate how the autocorrelations and the cross-correlations depend on the parameters,

we have considered a particular GAS^{2}V-N model with parameters $\phi = 0.98$ and $\sigma_\eta^2 = 0.05$. The leverage parameters α and k^* take values between 0 and 0.2 and 0 and 0.1, respectively. Figure 3.3 plots the first order autocorrelations of squares, $\rho_2(1)$ (top left panel), the first order autocorrelations of absolute returns, $\rho_1(1)$ (top right panel), and the first order cross-correlations between returns and future squared returns, $\rho_{21}(1)$ (bottom left panel), and future absolute returns, $\rho_{11}(1)$ (bottom right panel) when $k = 0$. These moments are also plotted in Figure 3.4 when $k = 0.1$. We can observe that they have very similar patterns as those of the GASV model; see Figure 2.3. First, the first order autocorrelations are positive and the surface is rather flat and it is not affected by the leverage effect parameters k^* and α. However, the first order autocorrelation of absolute returns is larger than that of the squared returns and increases with the two parameters. Finally, the cross-correlations are negative and decrease with the two leverage effect parameters, α and k^* linearly. By comparing Figure 3.3 and Figure 3.4, we can observe that larger value of k gives larger first order autocorrelations but negligible difference in cross-correlations.

To illustrate the shape of these moments for different lags, Figure 3.5 plots the first twenty orders of these moments for a GAS^{2}V-N model with parameters $\mu = 0$, $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\alpha = 0.07$, $k^* = 0.1$ when $k = 0$, while Figure 3.6 illustrates these moments when $k = 0.1$. The values of the parameters are chosen to be very similar to those obtained when fitting these models to financial data; see Section 3.4. The figures show that both the acf and absolute ccf decay exponentially towards zero. The absolute values of the moments related with absolute returns are larger than those of the squared returns. Therefore, we can conclude that the model is able to capture the Taylor Effect, phenomenon characterised by the autocorrelations of absolute returns to be larger than those of squares. Moreover, the larger value of k allows the model to capture larger autocorrelations of squared and absolute returns, therefore, volatility clustering.

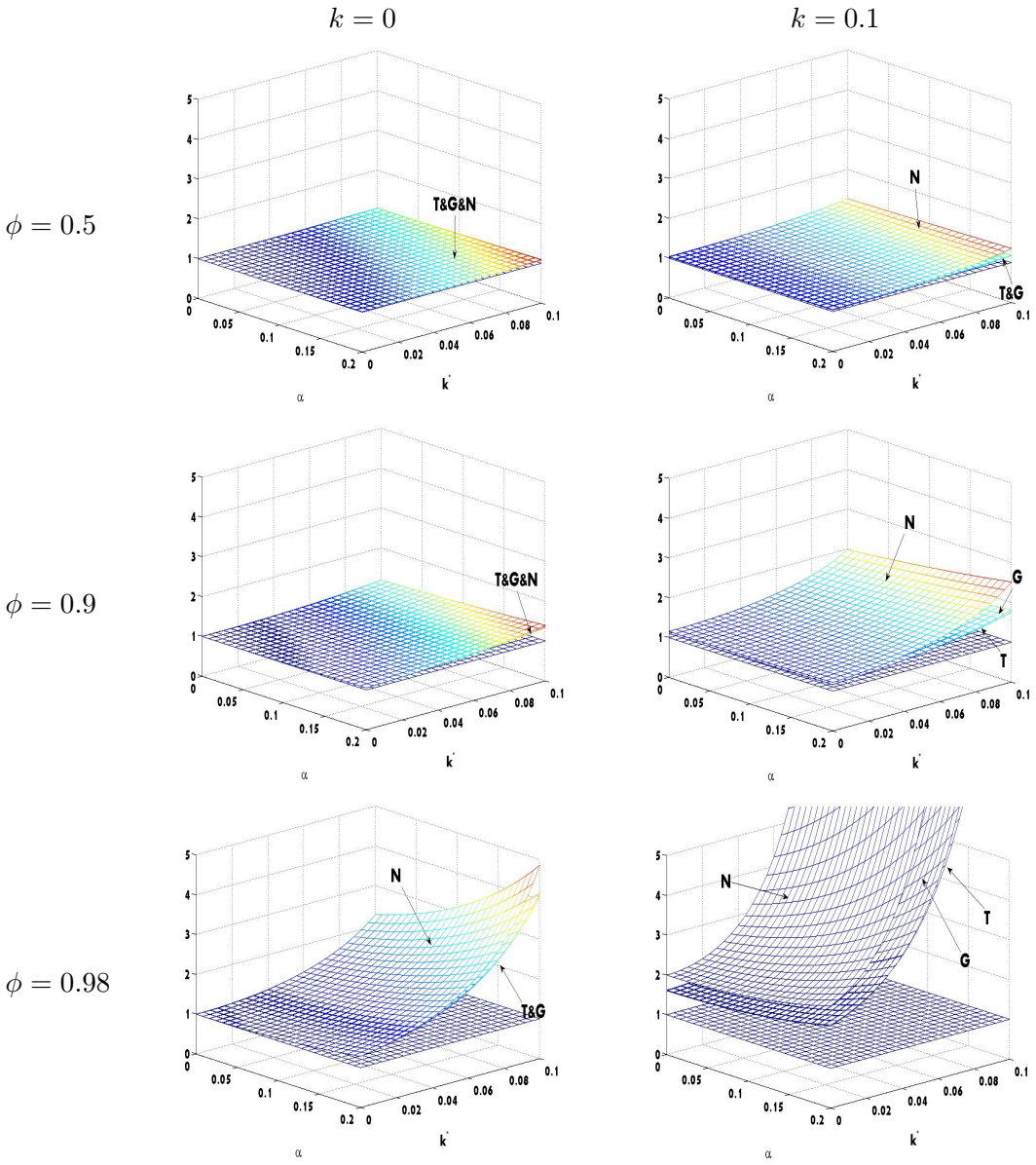

Figure 3.2: Ratio between the kurtoses of the GAS^{2}V model and the symmetric ARSV(1) model with Gaussian (N), GED (G) and Student-t (T) errors when $k = 0$ (left column) and 0.1 (right column) for three different values of the persistence parameter, $\phi = 0.5$ (first row), $\phi = 0.9$ (middle row) and $\phi = 0.98$ (bottom row).

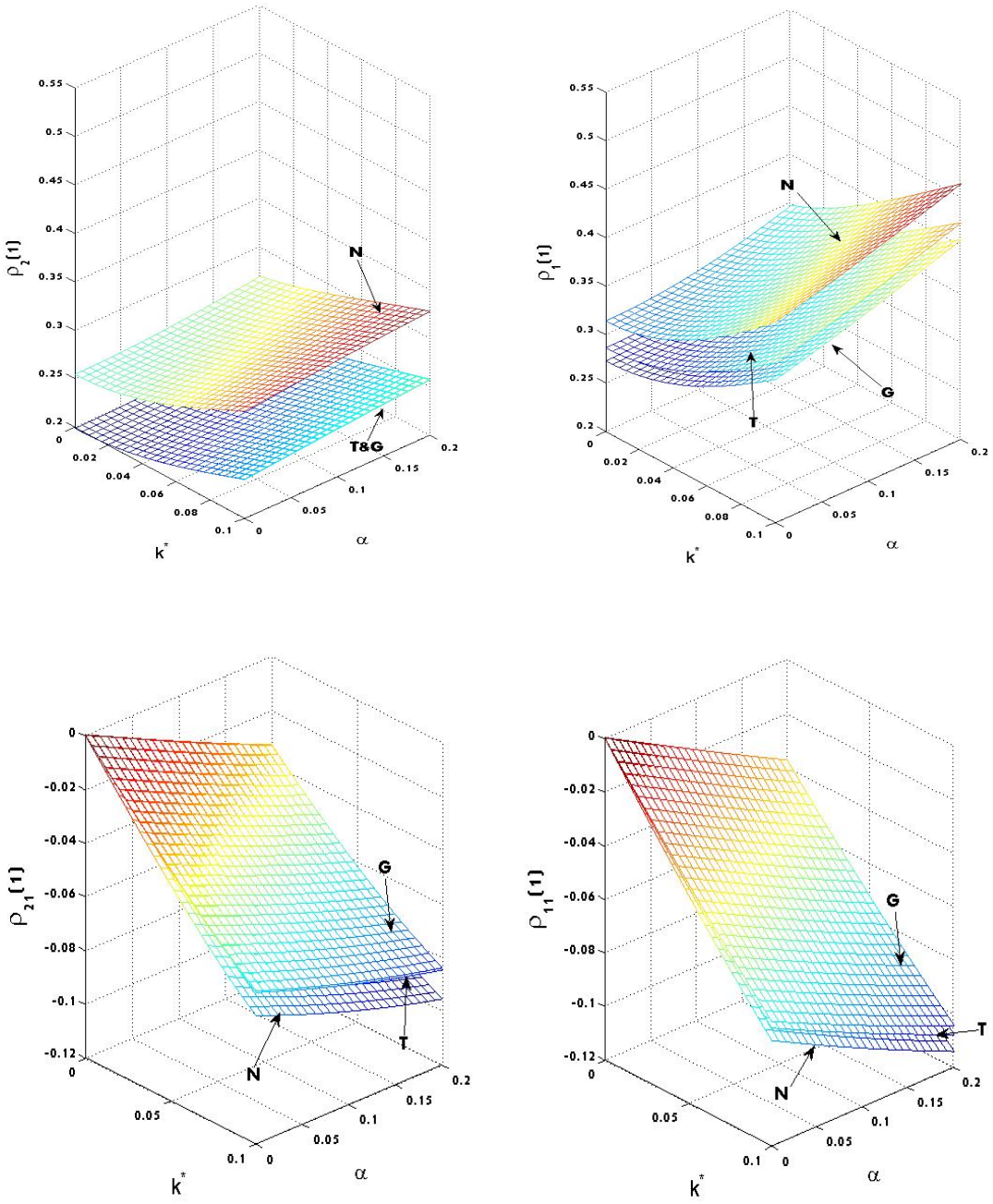

Figure 3.3: First order autocorrelations of squares (top left), first order autocorrelations of absolute returns (top right), first order cross-correlations between returns and future squared returns (bottom left) and first order cross-correlations between returns and future absolute returns (bottom right) of different GAS^{2}V models when $\mu = 0$, $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\nu = 1.5$, $\nu_0 = 11.8745$ and $k = 0$. The surface N represents the moments of the GAS^{2}V-N model, T represents the moments of the GAS^{2}V-T model and G represents the moments of the GAS^{2}V-G model.

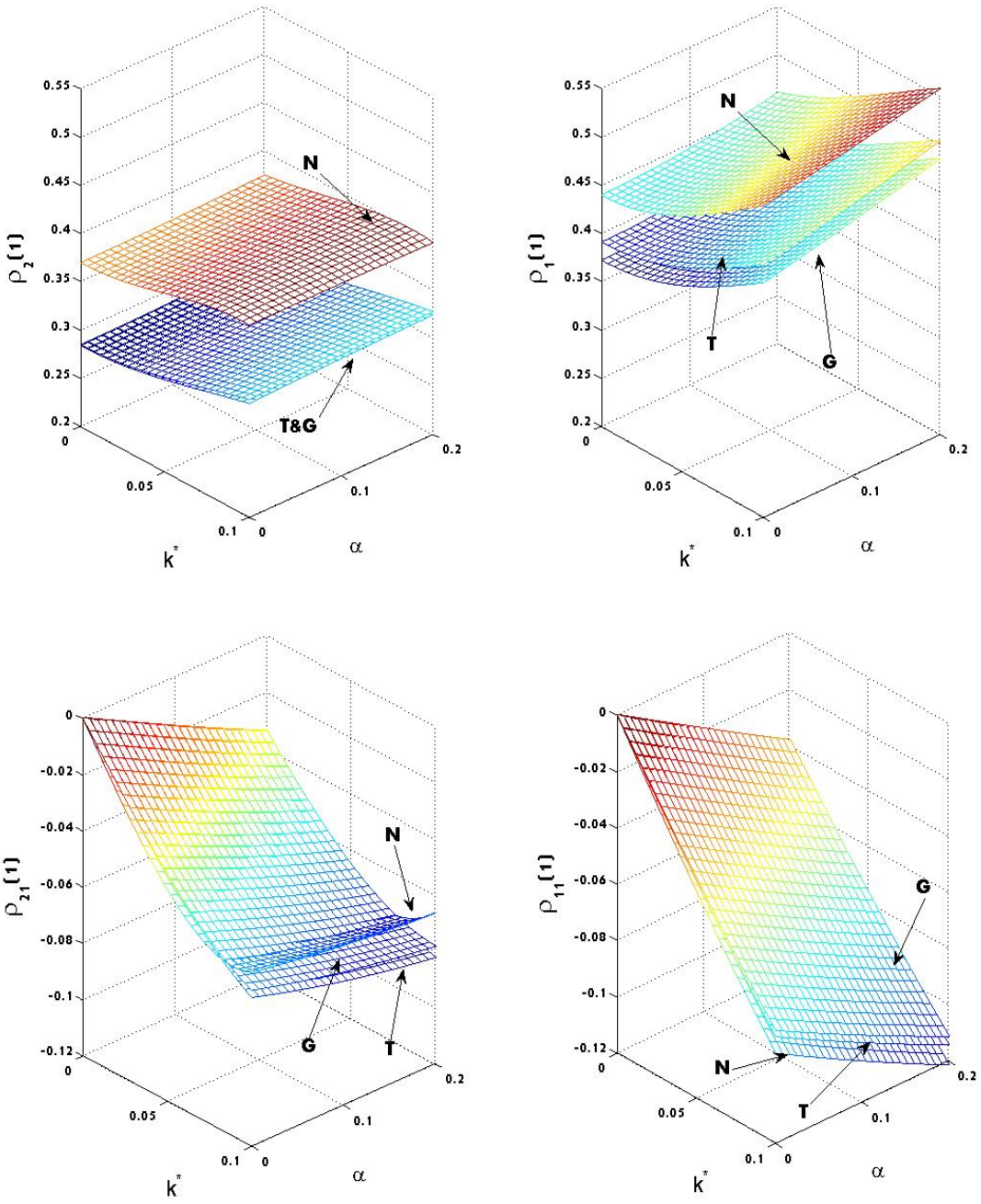

Figure 3.4: First order autocorrelations of squares (top left), first order autocorrelations of absolute returns (top right), first order cross-correlations between returns and future squared returns (bottom left) and first order cross-correlations between returns and future absolute returns (bottom right) of different GAS^{2}V models when $\mu = 0$, $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\nu = 1.5$, $\nu_0 = 11.8745$ and $k = 0.1$. The surface N represents the moments of the GAS^{2}V-N model, T represents the moments of the GAS^{2}V-T model and G represents the moments of the GAS^{2}V-G model.

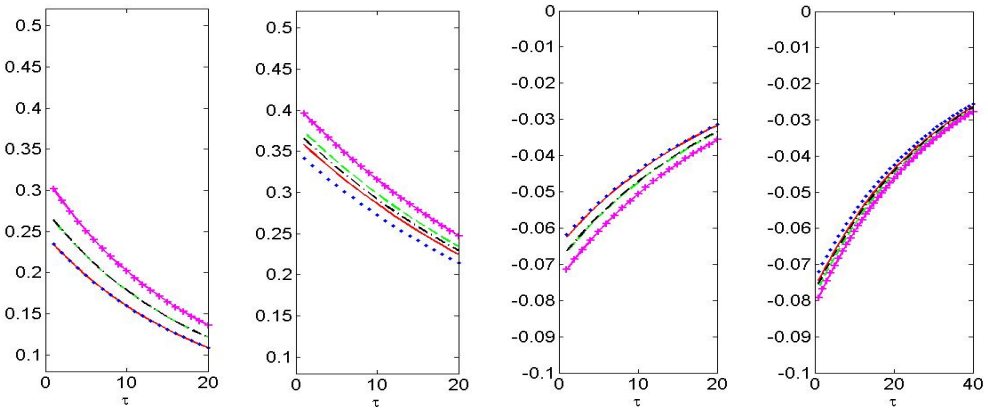

Figure 3.5: Autocorrelations of squares (first column), autocorrelations of absolute returns (second column), cross-correlations between returns and future squared returns (third column) and cross-correlations between returns and future absolute returns (fourth column) for different specifications of GAS^{2}V models when $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\alpha = 0.07$, $k^* = 0.08$ and $k = 0$. The solid line corresponds to the moments of the GAS^{2}V-T model with $\nu_0 = 11.8745$ while $\nu_0 = 19.8387$ for dashed lines. The dotted and dashdot lines corresponds to the moments of the GAS^{2}V-G model when $\nu = 1.5$ and 1.7, respectively. Finally, the '+-' line represents the moments of the GAS^{2}V-N model.

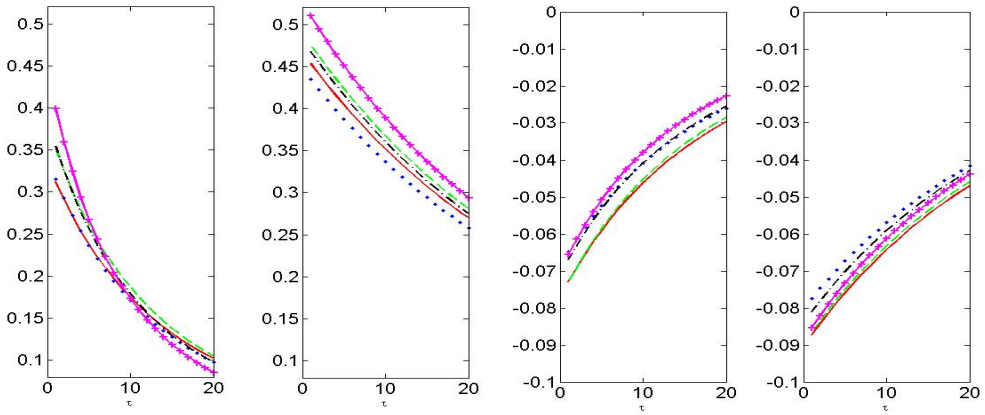

Figure 3.6: Autocorrelations of squares (first column), autocorrelations of absolute returns (second column), cross-correlations between returns and future squared returns (third column) and cross-correlations between returns and future absolute returns (fourth column) for different specifications of GAS^{2}V models when $\phi = 0.98$, $\sigma_\eta^2 = 0.05$, $\alpha = 0.07$, $k^* = 0.08$ and $k = 0.1$. The solid line corresponds to the moments of the GAS^{2}V-T model with $\nu_0 = 11.8745$ while $\nu_0 = 19.8387$ for dashed lines. The dotted and dashdot lines corresponds to the moments of the GAS^{2}V-G model when $\nu = 1.5$ and 1.7, respectively. Finally, the '+-' line represents the moments of the GAS^{2}V-N model.

GAS^{2}V-T

Alternatively, if ϵ_t is distributed as a standardized Student-t distribution with degrees of freedom $\nu_0 > 2$, pdf $\psi_0(\epsilon_t) = \frac{\Gamma((\nu_0+1)/2)}{\sqrt{\pi\nu_0}\Gamma(\nu_0/2)\varphi_0}\left(1+\frac{y_t^2}{\nu_0\varphi_0^2}\right)^{-\frac{\nu_0+1}{2}}$ with $\varphi_0 = \sqrt{\frac{\nu_0-2}{\nu_0}}$, then u_t is given by (3.7). We denote the model specified by equations (2.1), (3.1), (3.4) and (3.7) as GAS^{2}V-T. When $|\phi| < 1$, the model is stationary. Moreover, for some non-negative integer c, if $\nu_0 > c$, then the acf of $|y_t|^c$ is finite. If further, $\nu_0 > 2c$, the ccf between y_t and $|y_{t+\tau}|^c$ for a positive integer τ is also finite. The expectations needed to obtain the analytical expressions of the moments are derived in Appendix B.1.2.

Analogously, we illustrate the kurtosis of GAS^{2}V-T by plotting the factor R in Figure 3.2, for the same parameters chosen for the GAS^{2}V-N model and $\nu_0 = 11.8745$. Note that ν_0 guarantees ϵ_t to have the same kurtosis when it follows a GED distribution with degrees of freedom $\nu = 1.5$. We can observe that the ratio of the GAS^{2}V-T model is smaller than that of the GAS^{2}V-N when $\phi = 0.98$, while they are indistinguishable when ϕ is small.

As previously, we illustrate the first order of the acfs and ccfs of GAS^{2}V-T models in Figure 3.3 and Figure 3.4 when $\nu_0 = 11.8745$. The other parameters are the same as those chosen for the GAS^{2}V-N model. We observe that the GAS^{2}V-N model generates larger first order autocorrelations for both absolute and squared returns than the corresponding GAS^{2}V-T models. Moreover, the absolute values of the cross-correlations are also larger for the GAS^{2}V-N model than for the GAS^{2}V-T when $k = 0$. However, the absolute cross-correlation between returns and future squared returns are smaller in the case of the GAS^{2}V-N when $k = 0.1$ and k^* approximates 0.1.

We illustrate the first twenty orders of acfs and ccfs in Figure 3.5 and Figure 3.6 for the same parameter values used in the illustrations of the GAS^{2}V-N model while considering two different values of the degrees of freedom, namely 11.8745 and 19.8387, which guarantee the same kurtoses of ϵ_t when $\epsilon_t \sim GED$ with $\nu = 1.5$ and $\nu = 1.7$, respectively. We observe that the autocorrelations and cross-correlations of the absolute values are smaller than those of GAS^{2}V-N models for the considered parameter values. Moreover, larger degrees of freedom imply larger autocorrelations

and larger cross-correlation of absolute values. Therefore, we may conclude that fatter tails of ϵ_t imply smaller autocorrelations of both absolute and squared returns, which coincides with the conclusion of Carnero et al. (2004).

GAS^{2}V-GED

Finally, we assume that ϵ_t follows a $GED(\nu)$ distribution with probability density function (pdf) $\psi(\epsilon_t) = \frac{1}{2^{1+\frac{1}{\nu}}\varphi\Gamma(1+1/\nu)} \exp\left(-\frac{1}{2}\left|\frac{\epsilon_t}{\varphi}\right|^{\nu}\right)$ with $\varphi = \sqrt{2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/\nu)}$. Then u_t is given by (3.8) where $g_t \equiv \left|\frac{\epsilon_t}{\varphi}\right|^{\nu}$ follows a Gamma $(2, 1/\nu)$ distribution; see Harvey (2013). The model defined by equations (2.1), (3.1), (3.4) and (3.8) is denoted as GAS^{2}V-G. It is strictly stationary if $|\phi| < 1$ and if further $k + |k^*| < \frac{2}{\nu c}$, y_t and $|y_t|$ have finite and time-invariant moments of non-negative integer order c. Under these conditions, the acfs and ccfs are also finite. The analytical expressions of the two expectations are given in Appendix B.1.3.

In Figure 3.2, we also plot the ratio of the kurtoses between GAS^{2}V-G and ARSV(1) for the same parameter values specified for the GAS^{2}V-N model while $\nu = 1.5$. Though this GAS^{2}V-G always generates returns with higher kurtosis than the ARSV(1) model, its kurtosis is smaller than that of the corresponding GAS^{2}V-N with similar parameter values. As the Gaussian distribution is a special case of the GED distribution with $\nu = 2$, we might conclude that a fatter tailed GED generates less kurtosis. Moreover, the ratio of GAS^{2}V-G is indistinguishable from that of the GAS^{2}V-T model when the return errors are assumed to have the same kurtosis in both models. Apparently, the kurtosis of the return generated by the GAS^{2}V model depends on the kurtosis of the errors and barely on the type of distribution.

We also analyse the first order acfs and ccfs of the returns generated by the GAS^{2}V-G model when $\nu = 1.5$ in Figure 3.3 and Figure 3.4. We find that when the kurtoses of return errors are the same as in GAS^{2}V-T, these moments related with squared returns are indistinguishable for both models. The first order autocorrelation of absolute returns and first order cross-correlation between returns and future absolute returns of GAS^{2}V-T models are larger than those of the GAS^{2}V-G model.

Figure 3.5 and Figure 3.6 illustrate the first twenty orders of these moments for two different GAS^{2}V-G models with two different values of the GED parameter, $\nu = 1.5$ and 1.7. As expected, the acfs of $|y_t|$ and y_t^2 have both an exponential decay. Furthermore, fatter tails of ϵ_t imply smaller autocorrelations, but it has very mild influence on the cross-correlations. It verifies again that the acf of squared returns and ccf between returns and future squared return are indistinguishable to those of GAS^{2}V-T model with ϵ_t having the identical kurtosis.

3.3 Finite Sample performance of the MCMC estimator for the GAS^{2}V models

We adopt the same MCMC procedure described in Section 2.6 to estimate the parameters of the GAS^{2}V model with restriction $k = 0$. We assume the prior distribution for k^* as $k^* \sim N(0.05, 10)$ and for ν as $\nu \sim \chi^2_8$ when $\epsilon_t \sim t_\nu$ following the suggestion of Meyer and Yu (2000).

To check the reliability of this MCMC estimator, we simulate data from the three GAS^{2}V models, GAS^{2}V-N, GAS^{2}V-T and GAS^{2}V-G, with parameter values $\{\mu, \phi, \sigma_\eta^2, \alpha, k^*, \nu, \nu_0\} = \{0, 0.98, 0.05, 0.07, 0.08, 1.5, 11.8745\}$ while imposing the restriction $k = 0$. Recall that, in the previous section, we show that the GAS^{2}V-G and GAS^{2}V-T model generate returns with very similar properties when the parameters of both distributions are chosen to have the same kurtoses. Hence, there could be potential identification problem. Therefore, we consider the restricted GAS^{2}V models in the rest of this chapter. For each model, $T = 1000$ observations are simulated. The posterior mean and standard deviation of each parameter in the model is obtained by fitting the model to these simulated data using the MCMC estimator. The total number of iterations is 30,000 with the first 10,000 iterations used as burn-in. We replicate the experiment for $R = 200$ times.

Table 3.1 reports the Monte Carlo average of these posterior means and standard deviations together with the standard deviation of these posterior means based on these R replicates for each model. We find that the MCMC estimator is quite reliable for all parameters in all cases.

	μ	ϕ	α	k^*	σ_η^2	ν
GAS^{2}V-N						
True	**0**	**0.98**	**0.07**	**0.08**	**0.05**	-
Mean	0.131	0.976	0.067	0.083	0.054	-
	(1.259)	(0.010)	(0.056)	(0.021)	(0.018)	-
s.d.	1.548	0.007	0.060	0.020	0.014	-
GAS^{2}V-T						
True	**0**	**0.98**	**0.07**	**0.08**	**0.05**	**11.8745**
Mean	0.108	0.974	0.076	0.084	0.059	10.602
	(1.274)	(0.010)	(0.056)	(0.025)	(0.027)	(2.007)
s.d.	1.362	0.008	0.058	0.022	0.016	2.845
GAS^{2}V-G						
True	**0**	**0.98**	**0.07**	**0.08**	**0.05**	**1.5**
Mean	0.257	0.973	0.071	0.081	0.055	1.522
	(1.438)	(0.011)	(0.073)	(0.029)	(0.025)	(0.147)
s.d.	1.529	0.008	0.067	0.024	0.016	0.104

Table 3.1: Monte Carlo results of the MCMC estimator of the parameters of the GAS^{2}V model. The value reported are the Monte Carlo average and standard deviation (in parenthesis) of the posterior means together with the Monte Carlo average of the posterior standard deviation.

3.4 Empirical application

3.4.1 Estimation results from daily data

In this subsection, we fit the restricted GAS^{2}V models to the series of the S&P500 returns described in Chapter 1.

Table 3.2 reports the posterior mean, the 95% credible interval for each parameter and the marginal log-likelihood. From the table, several conclusions can be drawn. First, all the parameter estimates are different from zero. The credible intervals of the degrees of freedom in both GAS^{2}V-T and GAS^{2}V-G model exclude the case of the Normal distribution, which implies that the return

error follows a fat-tailed distribution. Regarding the goodness of fit of the models, we observe, analysing the log-likelihood values, that GAS^{2}V-N model outperforms the other two models and that the GAS^{2}V-G model fits the data better than the GAS^{2}V-T model.

	GAS^{2}V-N	GAS^{2}V-T	GAS^{2}V-G
μ	-2.401	-1.435	-1.892
	(-3.530, -0.215)	(-1.782, -0.669)	(-2.518, -0.169)
ϕ	0.978	0.980	0.982
	(0.966,0.989)	(0.968, 0.989)	(0.973, 0.992)
α	0.104	0.080	0.067
	(0.003, 0.145)	(0.048, 0.108)	(0.050, 0.093)
k^*	0.056	0.087	0.073
	(0.041, 0.071)	(0.069, 0.104)	(0.060, 0.080)
σ_η^2	0.020	0.011	0.008
	(0.009, 0.030)	(0.007, 0.021)	(0.001, 0.002)
ν	-	3.929	1.395
	-	(2.733, 3.176)	(1.267, 1.422)
Log-Likelihood	-5590.031	-5743.086	-5696.456

Table 3.2: Estimation results from daily S&P500. The values reported are the mean and 95% credible interval (parenthesis) of the posterior distributions.

3.4.2 Estimation results from weekly data

In this subsection, we fit the restricted GAS^{2}V models to the mean-adjusted weekly return series of S&P500 and NIKKEI225 observed from January 13, 1992 to December 27, 2010. The number of observations are $T_1 = 990$ and $T_2 = 986$, respectively. Although the sample size is relatively small, according to our Monte Carlo experiments, we can obtain reliable estimation results. For completeness, we also fit RT-GASV models with GED (RT-GASV-G) and Gaussian (RT-GASV-N) errors. Some of the relevant statistical moments are reported in the Table 3.3. We observe that the sample autocorrelations of the squared returns are significantly positive and the cross-correlations between returns and future squared returns are significantly negative, confirming the volatility clustering and leverage effect.

Estimation results are reported in Table 3.4. According to the log-likelihood and credible intervals, we can observe that γ_1 and k^* are not statistically significant for the S&P 500 which means that the models with normal errors are similar. This does not happen to the case of

NIKKEI225. Moreover, if we observe the estimates of the degrees of freedom, we can find that the distribution is not fat-tailed for weekly data. Finally, according to the log-likelihood, the RT-GASV-G model provides the best fit for both S&P500 and NIKKEI225 series of returns.

	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	$\rho_2(1)$	$\rho_1(1)$	$\rho_{21}(1)$	$\rho_{11}(1)$
S&P500	0.125	11.245	-20.195	2.404	-0.813^*	10.354^*	0.297^*	0.332^*	-0.254^*	-0.229^*
NIKKEI225	0.136	11.529	-27.805	3.113	-0.741^*	9.945^*	0.120^*	0.171^*	-0.125^*	-0.139^*

* Significant at 1% level.

Table 3.3: Sample moments of mean adjusted weekly S&P500 and NIKKEI225 returns observed from Jan 13, 1992 to Dec 27, 2010.

Data	Model	Log MargLik	μ	ϕ	σ_η^2	α	γ_1/k^*	ν
S&P500	GAS^{2}V-N	-2041.397	-2.534	0.964	0.030	0.281	0.012	
			(-3.914, -1.171)	(0.947, 0.982)	(0.013, 0.047)	(0.204, 0.380)	(-0.015, 0.037)	
	GAS^{2}V-T	-2056.212	-2.032	0.971	0.018	0.205	0.022	15.490
			(-3.651, -1.097)	(0.9542, 0.9839)	(0.004856, 0.02706)	(0.1375, 0.3335)	(-0.005294, 0.04658)	(9.894, 20.05)
	GAS^{2}V-G	-2044.018	-2.419	0.966	0.029	0.245	0.014	2.010
			(-5.262, -0.492)	(0.937, 0.982)	(0.015, 0.054)	(0.106, 0.339)	(-0.017, 0.051)	(1.802, 2.215)
	T-GASV-N	-2041.107	-1.548	0.962	0.032	0.230	-0.047	
			(-3.576, 0.09023)	(0.9436, 0.974)	(0.01539, 0.05349)	(0.09363, 0.3669)	(-0.1132, 0.007723)	
	T-GASV-G	-2033.651	-1.394	0.957	0.040	0.221	-0.058	2.257
			(-4.066, 0.516)	(0.9283, 0.9788)	(0.01514, 0.07397)	(0.07043, 0.353)	(-0.1335, 0.02006)	(1.969, 2.526)
NIKKEI225	GAS^{2}V-N	-2401.129	1.288	0.882	0.060	0.164	0.057	
			(0.599, 1.873)	(0.832, 0.932)	(0.029, 0.105)	(0.024, 0.306)	(0.006, 0.107)	
	GAS^{2}V-T	-2417.644	1.505	0.917	0.033	0.090	0.074	13.050
			(0.821, 2.125)	(0.848, 0.957)	(0.015, 0.065)	(-0.024, 0.222)	(0.029, 0.112)	(8.252, 21.04)
	GAS^{2}V-G	-2401.129	1.348	0.877	0.064	0.152	0.064	2.002
			(0.793, 2.017)	(0.797, 0.929)	(0.031, 0.119)	(0.016, 0.276)	(0.025, 0.114)	(1.824,2.252)
	T-GASV-N	-2399.766	1.800	0.893	0.059	0.041	-0.150	
			(0.700, 2.630)	(0.833, 0.932)	(0.033, 0.100)	(-0.175, 0.258)	(-0.276, -0.041)	
	T-GASV-G	-2391.140	1.897	0.873	0.074	0.030	-0.158	2.164
			(1.068, 2.666)	(0.801, 0.924)	(0.039, 0.126)	(-0.170, 0.229)	(-0.273, -0.055)	(1.964, 2.518)

Table 3.4: Estimation results from weekly S&P500 and NIKKEI225. The values reported are the mean and 95% credible interval (parenthesis) of the posterior distributions.

3.4.3 Forecasting results from weekly data

Good model in-sample performance does not necessary imply good model out-of-sample performance. In this section, we compare the out-of-sample performance of the proposed models using the two weekly return series described above. The three GAS^{2}V and the RT-GASV models are fitted to the return data and used to obtain one-period-ahead out-of-the-sample forecasts of weekly

volatility. We split the weekly sample into an in-sample estimation period and an out-of-sample forecast evaluation period. For estimation we use the rolling window scheme, where the size of the sample, which is used to estimate the competing models, is fixed at T_i with $i = 1$ and 2. The first forecast is made for the first week of January, 2011. When a new observation is added to the sample, we discard the first observation and re-estimate all the models. The re-estimated models are then used to forecast volatility. This process is repeated until we reach the end of the sample, December 30, 2013. In total, we obtain 157 forecasts from each model.

Two alternative criteria are considered in this chapter to compare the out-of-sample performances of these models, namely Mean Absolute Error (MAE) of the volatility forecasts and the Log Predictive Score (LPS), which is computed using the MCMC output. In Table 3.5, we report the MAE of the volatility forecasts. First, we calculated the weekly realized volatility (RV) obtained from the sum of daily squared returns. Let RV_t denote the weekly RV and $p(t,k)$ denote the k-th daily log-price in week t. Then RV_t is defined as $\sqrt{\sum_{k=1}^{N_t}(p(t,k)-p(t,k-1))^2}$, where N_t is the number of trading days in week t and $p(t,0) = p(t-1,N_{t-1})$. We match each volatility forecast with the corresponding realized volatility. Table 3.5 summarizes the MAE of the volatility forecasts. We can see that all the models perform nearly equally in forecasting the volatility of the S&P500 and NIKKEI225 returns.

On the other hand, LPS is a scoring rule introduced by Good (1952) that examines the model's performance when its implied predictive distribution is compared with observations not used in the inference sample. In this sense, it evaluates the out-of-sample behaviour of different models by mean of their divergence between the actual sampling density and the predictive density. The formula for the LPS is given as follows

$$LPS = \frac{1}{K}\sum_{k=1}^{K} log(f(y_{T+k}|y_k,\cdots,y_{T+k-1})), \tag{3.10}$$

where K is the total number of forecasts we've obtained. The one-step-ahead LPS are reported in the lower panel of Table 3.5. According to the LPS, the best model in forecasting the volatility is

the GAS^{2}V-T although the difference among the alternative models are almost negligible.

MAE*1000					
	GAS^{2}V-N	GAS^{2}V-T	GAS^{2}V-G	T-GASV-N	T-GASV-G
S&P500	6.220	6.378	6.248	6.206	6.220
NIKKEI 225	9.171	9.536	9.157	9.235	8.992
LPS					
S&P500	-2.047	-2.047	-2.064	-2.049	-2.062
NIKKEI 225	-2.572	-2.524	-2.764	-2.579	-2.672

Table 3.5: Forecasting results from weekly data. MAE refers to the mean absolute forecasting error and LPS refers to the log-predictive likelihood.

3.5 Conclusion

In this chapter, we propose to extend the asymmetric SV models by specifying the volatility as being driven by the conditional score of lagged return. This type of models, denoted as GAS^{2}V, can automatically correct the influential observations, which are outliers judged by the Gaussian yardstick and usually attribute to an increase in the volatility in the traditional SV models.

Three GAS^{2}V models are proposed, namely, GAS^{2}V-N, GAS^{2}V-T and GAS^{2}V-G corresponding to the return errors following a Normal, Student-t and GED distribution, respectively. The closed-form expressions of their statistical properties are derived and analyzed. We find that the GAS^{2}V model with Student-t error generates returns with similar moments as those generated by the GAS^{2}V model with GED error for fixed kurtosis of return errors.

Finally, the new proposals are fitted to both daily and weekly financial data and we observe that the GAS^{2}V-T model provides the best fit in-sample for the daily S&P500 return series. Regarding the out-of-sample performance of the models in forecasting the volatility of the weekly financial returns of the S&P500 and NIKKEI225, all models provide similar mean absolute forecast errors when the volatility forecasts are compared with a consistent measure of volatility, the realized volatility. Using the Log Predictive Score criterion, the best model in forecasting the volatility of the two series of financial returns is the GAS^{2}V-T, although the difference among models are almost negligible.

Chapter 4

Conclusions and Future Research

4.1 Conclusions

In this dissertation, we propose a family of asymmetric Stochastic Volatility (SV) models, named GASV. This family is very general and includes some of the most famous asymmetric SV models available in the literature as, for instance, the A-ARSV model of Taylor (1994) and Harvey and Shephard (1996), the exponential SV (E-SV) model proposed by Demos (2002) and Asai and McAleer (2011) and a restricted version of the Threshold SV (RT-SV) model of Breidt (1996) and So et al. (2002). The statistical properties of the GASV models are derived, namely, marginal variance, kurtosis, autocorrelations of power transformed absolute returns and cross-correlations between returns and future power transformed absolute absolute returns. These statistical properties are important for evaluating the ability of the models to explain the empirical properties of interest when dealing with real financial time series.

We show that some of the parameters of the proposed T-GASV model cannot be identified when the parameter of the GED distribution is allowed to change as, in this case, the moments of returns can be indistinguishable for different combinations of the parameters and distributions. As a byproduct, we obtain the statistical properties of those nested models, some of which were previously unknown. We find that the E-SV model is able to capture more volatility clustering

compared to the A-ARSV while the T-GASV model is more flexible to represent the leverage effect than the A-ARSV and E-SV models. Finally, we show that allowing the autoregressive parameter and/or the variance of the log-volatility disturbance in the T-SV model to be different depending on the sign of past returns do not generate leverage effect. However, changing the constant in the volatility equation allows the T-SV model to capture asymmetric conditional heteroscedasticity.

Besides the T-GASV model, we also propose another kind of GASV models, the score-driven GASV, denoted as GAS^{2}V models, aiming at robustifying the traditional SV models, which might suffer from a potential drawback that a large realisation of the return error, that can be due to the heavy-tailed nature of the distributions, will be attributed to an increase in volatility. The analytical expressions of the statistical moments of the GAS^{2}V models are derived when the return errors follow either Normal, Student-t or GED distribution. It is important to point out that analytical expressions of these moments of the GAS^{2}V model with Student-t return errors can be derived, in opposition to the traditional specifications of the volatility where they are hardly possible to be derived. We also show that the GAS^{2}V model with Student-t errors generates returns with very similar moments as those generated by the GAS^{2}V model with GED errors for fixed kurtosis of return errors.

Another contribution of this thesis is the proposal of the Stochastic News Impact Surface (SNIS) to describe the asymmetric response of volatility to positive and negative past returns in the context of SV model. It is a surface relating the volatility with the level and volatility disturbances. From the SNIS, we can observe the asymmetric response of the volatility and this asymmetry is different on the values of the volatility error.

We consider a MCMC estimator that is implemented by the user-friendly available software BUGS to estimate the parameters and volatility of the restricted T-GASV (RT-GASV) and GAS^{2}V models. Through extensive Monte Carlo studies, we show the adequacy of the finite sample properties of this estimator. Moreover, by fitting the general RT-GASV model, we are able to identify the true data generating process. Therefore, in empirical applications, researches will be better off by fitting the RT-GASV model and letting the data choose the preferred specification of

the volatility instead of choosing a particular ad hoc specification. The RT-GASV model and its nested models are fitted to one series of daily S&P500 returns and used to forecast the one-step-ahead VaR. For this particular data set, the models with GED errors give better fit than those with Gaussian models. Moreover, our RT-GASV model with GED errors provides the best estimates of the VaR. The three GAS^{2}V models are also fitted to this series. It turns out that the GAS^{2}V models with GED and Student-t models fit the data as well as the RT-GASV model with GED errors and better than the GAS^{2}V model with Gaussian errors. We also fit the GAS^{2}V models to two weekly data, S&P500 and NIKKEI225, and forecast the one-step-ahead volatility. The out-of-sample results, according to the LPS, slightly favor the GAS^{2}V models in comparison to the RT-GASV model.

4.2 Future research

In this section, we discuss several possible extensions of the ideas proposed in this thesis. First, 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. In this thesis, we compare alternative asymmetric SV models in terms of their statistical properties. Hence, it is interesting to compare the properties of these alternative SV models with those of the best candidates within the GARCH family including the robust score-driven GARCH models, such as the Beta-t-EGARCH and Gamma-GED-EGARCH models of Harvey (2013).

Second, we propose the score driven GASV models aiming at robustifying the traditional asymmetric SV models. Hence, it is important to analyze the performance of these models in the presence of outliers and compare to the traditional ones.

Third, although our focus is on univariate models, 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 future research

line.

Fourth, Bandi and Renò (2012) and Yu (2012) argue that the leverage effect found in many real time series of financial returns can be time-varying. Extending the model and results derived in this paper to include time-varying leverage effect is also in our research agenda.

Finally, the MCMC procedure is usually time-consuming. Some alternative estimation methods can be considered to estimate the GASV models. First, the efficient importance sampling (EIS) method of Liesenfeld and Richard (2003) and Richard and Zhang (2007) might be an alternative. Particularly, we would like to extend the numerically accelerated importance sampling (NAIS) proposed by Koopman et al. (2014) to estimate the GASV models. The NAIS extends the global approximating method of Richard and Zhang (2007) by solving for the parameters of the importance sampling distribution using Gauss-Hermit quadrature rather than simulation. They show that their NAIS method produces reliable results in a numerically and computationally efficient way. Another possible alternative is the Approximate Bayesian Computation (ABC) method which is a class of simulation-based algorithms and methods developed to perform inference by circumventing of explicit evaluation of the likelihood. Moreover, the Particle Learning (PL) approach is also under our consideration. It was firstly introduced by Carvalho et al. (2010) and implemented by Virbickaite et al. (2014) to estimate a Bayesian non-parametric SV models with Markov switching jumps.

References

Abanto-Valle, C., Bandyopadhyay, D., Lachos, V., and Enriquez, I. (2010). Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Computational Statistics & Data Analysis*, 54(12):2883–2898.

Andersen, T., Bollerslev, T., Diebold, F., and Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.

Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Asai, M. and McAleer, M. (2004). Dynamic leverage and threshold effects in stochastic volatility models. *Unpublished paper, Faculty of Economics, Tokyo Metropolitan University*.

Asai, M. and McAleer, M. (2005). Dynamic asymmetric leverage in stochastic volatility models. *Econometric Reviews*, 24:317–332.

Asai, M. and McAleer, M. (2006). Asymmetric multivariate stochastic volatility. *Econometric Reviews*, 25(2-3):453–473.

Asai, M. and McAleer, M. (2009). Multivariate stochastic volatility, leverage and news impact surfaces. *The Econometrics Journal*, 12(2):292–309.

Asai, M. and McAleer, M. (2011). Alternative asymmetric stochastic volatility models. *Econometric Reviews*, 30:548–564.

Asai, M., McAleer, M., and Medeiros, M. C. (2012). Asymmetry and long memory in volatility modeling. *Journal of Financial Econometrics*, 10(3):495–512.

Bandi, F. and Renò, R. (2012). Time-varying leverage effects. *Journal of Econometrics*, 169:94–113.

Bartolucci, F. and De Luca, G. (2003). Likelihood-based inference for asymmetric stochastic volatility models. *Computational Statistics & Data Analysis*, 42(3):445–449.

Black, F. (1976). Studies of stock price volatility changes, proceedings of the 1976 meetings of the business and economic statistics section. 177-191. In *American Statistical association*.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

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

Bollerslev, T., Litvinova, J., and Tauchen, G. (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics*, 4(3):353–384.

Bollerslev, T. and Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109(1):33–65.

Breidt, F. J. (1996). A threshold autoregressive stochastic volatility model. In *VI Latin American Congress of Probability and Mathematical Statistics (CLAPEM), Valparaiso, Chile*. Citeseer.

Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys*, 18(5):613–649.

Caporin, M. and McAleer, M. (2011). Thresholds, news impact surfaces and dynamic asymmetric multivariate GARCH. *Statistica Neerlandica*, 65(2):125–163.

Cappuccio, N., Lubian, D., and Raggi, D. (2004). MCMC Bayesian estimation of a skew-GED stochastic volatility model. *Studies in Nonlinear Dynamics & Econometrics*, 8(2).

Carnero, M., Peña, D., and Ruiz, E. (2004). Persistence and kurtosis in GARCH and stochastic volatility models. *Journal of Financial Econometrics*, 2(2):319–342.

Carvalho, C. M., Johannes, M. S., Lopes, H. F., Polson, N. G., et al. (2010). Particle learning and smoothing. *Statistical Science*, 25(1):88–106.

Cavaliere, G. (2006). Stochastic volatility: Selected readings. *The Economic Journal*, 116(512):F326–F327.

Chan, D., Kohn, R., and Kirby, C. (2006). Multivariate stochastic volatility models with correlated errors. *Econometric Reviews*, 25(2-3):245–274.

Chen, C., Liu, F., and So, M. (2008). Heavy-tailed-distributed threshold stochastic volatility models in financial time series. *Australian & New Zealand Journal of Statistics*, 50(1):29–51.

Chesney, M. and Scott, L. (1989). Pricing European currency options: A comparison of the modified Black-Scholes model and a random variance model. *Journal of Financial and Quantitative Analysis*, 24(03):267–284.

Chib, S., Nardari, F., and Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371.

Choy, B., Wai, Y., and Chan, C. (2008). Bayesian Student-t stochastic volatility models via scale mixtures. *Advances in Econometrics*, 23:595–618.

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

Danielsson, J. (1998). Multivariate stochastic volatility models: Estimation and a comparison with VGARCH models. *Journal of Empirical Finance*, 5(2):155–173.

Demos, A. (2002). Moments and dynamic structure of a time-varying parameter stochastic volatility in mean model. *Econometrics Journal*, 5(2):345–357.

Ding, Z., Granger, C., and Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

Durbin, J. and Koopman, S. J. (1997). Monte carlo maximum likelihood estimation for non-gaussian state space models. *Biometrika*, 84(3):669–684.

Elliott, R., Liew, C., and Siu, T. (2011). On filtering and estimation of a threshold stochastic volatility model. *Applied Mathematics and Computation*, 218(1):61–75.

Engle, R. and Ng, V. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1779–1801.

Engle, R. F. (1995). *ARCH: selected readings*. Oxford University Press.

Fridman, M. and Harris, L. (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business & Economic Statistics*, 16(3):284–291.

Garcia, R., Lewis, M.-A., Pastorello, S., and Renault, É. (2011). Estimation of objective and risk-neutral distributions based on moments of integrated volatility. *Journal of Econometrics*, 160(1):22–32.

Ghysels, E., Harvey, A., and Renault, E. (1996). Stochastic Volatility. In Maddala, G. S., Rao, C. R., and Vinod, H. D., editors, *Statistical Methods in Finance*. Amsterdam: North-Holland.

Giraitis, L., Leipus, R., and Surgailis, D. (2007). Recent advances in ARCH modelling. In *Long memory in economics, 3–38*. Springer.

Glosten, L., Jaganathan, R., and Runkle, D. (1993). On the relation between the expected value and the volatility of the nominal excess returns on stocks. *Journal of Finance*, 48:1779–1801.

Good, I. J. (1952). Rational decisions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 14(1):107–114.

Harvey, A. (1990). *The Econometric Analysis of Time Series*. Cambridge: MIT Press, Second edition.

Harvey, A. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press.

Harvey, A., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.

Harvey, A. and Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–34.

Hibbert, A. M., Daigler, R. T., and Dupoyet, B. (2008). A behavioral explanation for the negative asymmetric return–volatility relation. *Journal of Banking & Finance*, 32(10):2254–2266.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300.

Jacquier, E., Polson, N., and Rossi, P. (1994). Bayesian analysis of stochastic volatility models: Reply. *Journal of Business & Economic Statistics*, 12(4):413–417.

Jacquier, E., Polson, N., and Rossi, P. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212.

Jeffreys, H. (1961). *The theory of probability*. Oxford University Press.

Jungbacker, B. and Koopman, S. J. (2006). Monte Carlo likelihood estimation for three multivariate stochastic volatility models. *Econometric Reviews*, 25(2-3):385–408.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65(3):361–393.

Koopman, S. J., Lucas, A., and Scharth, M. (2014). Numerically accelerated importance sampling for nonlinear non-gaussian state space models. *Journal of Business & Economic Statistics, Forthcoming*.

Lien, D. (2005). A note on asymmetric stochastic volatility and futures hedging. *Journal of Futures Markets*, 25(6):607–612.

Liesenfeld, R. and Jung, R. C. (2000). Stochastic volatility models: Conditional normality versus heavy-tailed distributions. *Journal of Applied Econometrics*, 15(2):137–160.

Liesenfeld, R. and Richard, J. (2003). Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *Journal of Empirical Finance*, 10(4):505–531.

Maheu, J. M. and McCurdy, T. H. (2004). News arrival, jump dynamics, and volatility components for individual stock returns. *The Journal of Finance*, 59(2):755–793.

Meyer, R. and Yu, J. (2000). BUGS for a Bayesian analysis of stochastic volatility models. *The Econometrics Journal*, 3(2):198–215.

Montero, J., Fernández-Avilés, G., and García, M. (2010). Estimation of asymmetric stochastic volatility models: Application to daily average prices of energy products. *International Statistical Review*, 78(3):330–347.

Muñoz, M. P., Marquez, M. D., and Acosta, L. M. (2007). Forecasting volatility by means of threshold models. *Journal of Forecasting*, 26(5):343–363.

Nakajima, J. and Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics & Data Analysis*, 53(6):2335–2353.

Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Students t-distribution. *Computational Statistics & Data Analysis*, 56(11):3690–3704.

Nandi, S. (1998). How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the S&P 500 index options market. *Journal of Banking & Finance*, 22(5):589–610.

Nelson, D. (1991). Conditional heteroscedasticity in asset pricing: A new approach. *Econometrica*, 59:347–370.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449.

Omori, Y. and Watanabe, T. (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics & Data Analysis*, 52(6):2892–2910.

Pérez, A., Ruiz, E., and Veiga, H. (2009). A note on the properties of power-transformed returns in long-memory stochastic volatility models with leverage effect. *Computational Statistics & Data Analysis*, 53(10):3593–3600.

Richard, J.-F. and Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141(2):1385–1411.

Rodríguez, M. and Ruiz, E. (2012). GARCH models with leverage effect: differences and similarities. *Journal of Financial Econometrics*, 10(4):637–668.

Ruiz, E. and Pérez, A. (2012). Maximally autocorrelated power transformations: A closer look at the properties of stochastic volatility models. *Studies in Nonlinear Dynamics & Econometrics*, 16(3).

Ruiz, E. and Veiga, H. (2008). Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH. *Computational Statistics & Data Analysis*, 52(6):2846–2862.

Ryzhik, I., Jeffrey, A., and Zwillinger, D. (2007). *Table of integrals, series and products*. London: Academic Press, Seventh edition.

Sandmann, G. and Koopman, S. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics*, 87(2):271–301.

Savva, C. (2009). International stock markets interactions and conditional correlations. *Journal of International Financial Markets, Institutions and Money*, 19(4):645–661.

Sentana, E. (1995). Quadratic ARCH models. *The Review of Economic Studies*, 62(4):639–61.

Shephard, N. and Pitt, M. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, 84(3):653–667.

Smith, D. (2009). Asymmetry in stochastic volatility models: Threshold or correlation? *Studies in Nonlinear Dynamics & Econometrics*, 13(3).

So, M., Li, W., and Lam, K. (2002). A threshold stochastic volatility model. *Journal of Forecasting*, 21(7):473–500.

Stout, W. F. (1974). *Almost Sure Convergence*. London: Academic Press.

Takahashi, M., Omori, Y., and Watanabe, T. (2013). News impact curve for stochastic volatility models. *Economics Letters*, 120(1):130–134.

Taylor, S. (1986). *Modelling Financial Time Series*. Chichester: Wiley.

Taylor, S. (1994). Modelling stochastic volatility: A review and comparative study. *Mathematical Finance*, 4:183–204.

Taylor, S. (2007). *Asset price dynamics, volatility, and prediction*. Princeton, NJ: Princeton University Press.

Teräsvirta, T. (2009). An introduction to univariate garch models. In *Handbook of Financial Time Series*, pages 17–42. Springer.

Tsiotas, G. (2012). On generalised asymmetric stochastic volatility models. *Computational Statistics & Data Analysis*, 56(1):151–172.

Virbickaite, A., Lopes, H. F., Ausín, C., and Galeano, P. (2014). Particle learning for bayesian non-parametric markov switching stochastic volatility model. *WP*.

Walker, S. and Gutiérrez-Peña, E. (1999). Robustifying Bayesian procedures. *Bayesian Statistics*, 6:685–710.

Wang, J. J., Chan, J. S., and Choy, S. B. (2011). Stochastic volatility models with leverage and heavy-tailed distributions: A Bayesian approach using scale mixtures. *Computational Statistics & Data Analysis*, 55(1):852–862.

Wang, J. J., Chan, J. S., and Choy, S. B. (2013). Modelling stochastic volatility using generalized t distribution. *Journal of Statistical Computation and Simulation*, 83:340–354.

Watanabe, T. and Omori, Y. (2004). A multi-move sampler for estimating non-Gaussian time series models: Comments on Shephard & Pitt (1997). *Biometrika*, 91(1):246–248.

Wiggins, J. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19(2):351–372.

Yu, J. (2012). A semiparametric stochastic volatility model. *Journal of Econometrics*, 167(2):473–482.

Yu, J. and Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, 25(2-3):361–384.

Yu, J., Yang, Z., and Zhang, X. (2006). A class of nonlinear stochastic volatility models and its implications for pricing currency options. *Computational Statistics & Data Analysis*, 51(4):2218–2231.

Zakoian, J. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955.

Appendix A

Appendix to Chapter 2

A.1 Proof of Theorems

A.1.1 Proof of Theorem 2.1

Consider y_t, which, according to equation (2.1), is given by $y_t = \epsilon_t \exp(h_t/2)$. From equation (2.2), h_t can be written as

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

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 $|x - x_0| = \sqrt{\sum_{i=1}^{\infty}(x_i - x_i^0)^2} < \delta$, we have $|\Psi(x) - \Psi(x_0)| = |\sum_{i=1}^{\infty} \phi^{i-1}(x_i - x_i^0)|$. Using the Cauchy-Schwarz inequality, it follows that $|\Psi(x) - \Psi(x_0)| \leq \sqrt{\sum_{i=1}^{\infty} \phi^{2i-2}} \sqrt{\sum_{i=1}^{\infty}(x_i - x_i^0)^2} < \frac{\delta}{\sqrt{1-\phi^2}} = \varsigma$. Therefore, $\Psi(x)$ is continuous, and consequently, measurable.

Second, given that ϵ_t and η_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 σ_t is a continuous function of h_t, σ_t is also strictly stationary. The level noise ϵ_t is independent of σ_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 σ_t^2 are strictly stationary and, consequently, any existing moments are time invariant. Next we show that σ_t has finite moments of arbitrary positive order c when ϵ_t follows a distribution such that $E(\exp(0.5cf(\epsilon_t))) < \infty$.

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

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

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

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

As η_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 η_t and ϵ_t are both IID sequences, it is easy to show that (A.3) becomes

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

We need to show that $P(0.5c\phi^{i-1}) \equiv \prod_{i=1}^{\infty} E\left[\exp\left(0.5c\phi^{i-1} f(\epsilon_{t-i})\right)\right]$ is finite when $E\left(\exp\left(0.5cf(\epsilon_{t-i})\right)\right]) < \infty$. 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 $P(b_i) \equiv \prod_{i=1}^{\infty} E\left[\exp\left(b_i f(\epsilon_{t-i})\right)\right]$ 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

$$a_i - 1 = O(b_i) \; as \; b_i \to 0.$$

Consequently, for some $\varsigma > 0$, there exist a finite M independent of i such that

$$\sup_{|b_i|<\varsigma, b_i \neq 0} |O(b_i)| < M|b_i|.$$

$\sum_{i=1}^{\infty} |b_i| < \infty$ 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, $E(\sigma_t^c)$ 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 $E(\exp(0.5cf(\epsilon_t))) < \infty$, $E(\sigma_t^c)$ is finite.

Finally, consider y_t, which, according to equation (2.1), is given by $y_t = \sigma_t \epsilon_t$. Therefore, given that σ_t and ϵ_t are contemporaneously independent, the following expressions are obtained

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

$$E(y_t^c) = E(\sigma_t^c) E(\epsilon_t^c). \tag{A.6}$$

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

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

where $P(b_i) \equiv \prod_{i=1}^{\infty} E(\exp(b_i f(\epsilon_{t-i})))$. Therefore, if further ϵ_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

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

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

A.1.2 Proof of Theorem 2.2

Consider y_t as given in equations (2.1) and (2.2). We first compute the τ-th order auto-covariance of $|y_t|^c$ which is given by

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

Note that from equation (2.2), $\sigma_t^c = \exp\{0.5ch_t\}$ can be written as follows

$$\sigma_t^c = \exp\{0.5c\mu(1-\phi^\tau)\}\exp\left\{0.5c\sum_{i=1}^{\tau}\phi^{i-1}(f(\epsilon_{t-i}) + \eta_{t-i})\right\}\sigma_{t-\tau}^{c\phi^\tau}. \tag{A.10}$$

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

$$\begin{aligned} &cov(|y_t|^c, |y_{t-\tau}|^c) = \\ &E\left(|\epsilon_t|^c|\epsilon_{t-\tau}|^c\exp(0.5c\mu(1-\phi^\tau))\exp\left(\sum_{i=1}^{\tau}0.5c\phi^{i-1}(f(\epsilon_{t-i}) + \eta_{t-i})\right)\sigma_{t-\tau}^{c(\phi^\tau+1)}\right) \\ &-\left\{\exp(0.5c\mu)E\left(|\epsilon_t|^c\right)\exp\left(\frac{c^2\sigma_\eta^2}{8\left(1-\phi^2\right)}\right)P(0.5c\phi^{i-1}))\right\}^2. \end{aligned} \tag{A.11}$$

Given that ϵ_t and η_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 σ_t in (A.4), equation (A.11) can be written as follows

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

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

A.1.3 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.1.2. That is

$$cov\left(|y_t|^c, y_{t-\tau}\right) = \exp(0.5(c+1)\mu)E\left(|\epsilon_t|^c\right)\exp\left(\frac{1+c^2+2c\phi^\tau}{8\left(1-\phi^2\right)}\sigma_\eta^2\right)E\left(\epsilon_t\exp\left(0.5c\phi^{\tau-1}f\left(\epsilon_t\right)\right)\right)$$
$$\cdot\prod_{i=1}^{\infty}E\left(\exp\left(0.5\left(1+c\phi^\tau\right)\phi^{i-1}f\left(\epsilon_{t-i}\right)\right)\right)\prod_{i=1}^{\tau-1}E\left(\exp\left(0.5c\phi^{i-1}f\left(\epsilon_{t-i}\right)\right)\right). \tag{A.12}$$

Finally, $\rho_{c1}(\tau) = \frac{cov(|y_t|^c,y_{t-\tau})}{\sqrt{E(|y_t|^{2c})-E^2(|y_t|^c)}\sqrt{E(y_t^2)}}$ together with (A.7) and (A.12) yields the required equation (2.8).

A.2 Expectations

A.2.1 Expectations needed to compute $E(|y_t|^c)$, $corr(|y_t|^c, |y_{t+\tau}|^c)$ and $corr(y_t, |y_{t+\tau}|^c)$ when $\epsilon \sim GED(\nu)$

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

of ϵ is symmetric with support $(-\infty, \infty)$, if p is a nonnegative finite integer, then

$$\begin{aligned} E(|\epsilon|^p) &= C_0 \int_{-\infty}^{+\infty} |\epsilon|^p \exp\left(-\frac{|\epsilon|^\nu}{2\lambda^\nu}\right) d\epsilon \\ &= 2C_0 \int_0^{+\infty} \epsilon^p \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) d\epsilon. \end{aligned}$$

Substituting $s = \frac{\epsilon^\nu}{2\lambda^\nu}$ and solving the integral yields

$$E(|\epsilon|^p) = 2^{\frac{p}{\nu}} \lambda^p \Gamma\left((p+1)/\nu\right) / \Gamma\left(1/\nu\right). \tag{A.13}$$

On the other hand,

$$\begin{aligned} &E(|\epsilon|^p \exp(bf(\epsilon))) \\ &= \int_{-\infty}^{+\infty} |\epsilon|^p \exp(b\alpha I(\epsilon < 0) + b\gamma_1 \epsilon + b\gamma_2 |\epsilon|) C_0 \exp\left(-\frac{|\epsilon|^\nu}{2\lambda^\nu}\right) d\epsilon \\ &= C_0 \left[\int_{-\infty}^{0} (-\epsilon)^p \exp(b\alpha) \exp(b(\gamma_1 - \gamma_2)\epsilon) \exp\left(-\frac{(-\epsilon)^\nu}{2\lambda^\nu}\right) d\epsilon \right. \\ &\qquad \left. + \int_0^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) d\epsilon \right]. \end{aligned}$$

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

$$\begin{aligned} E(|\epsilon|^p \exp(bf(\epsilon))) &= C_0 \left[\int_0^{+\infty} s^p \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)s) \exp\left(-\frac{s^\nu}{2\lambda^\nu}\right) ds \right. \\ &\qquad \left. + \int_0^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) d\epsilon \right] \\ &= C_0 \int_0^{+\infty} \epsilon^p \exp\left(-\frac{\epsilon^\nu}{2\lambda^\nu}\right) \left[\exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\epsilon) + \exp(b(\gamma_1 + \gamma_2)\epsilon)\right] d\epsilon. \end{aligned} \tag{A.14}$$

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

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

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

$$E(|\epsilon|^p \exp(bf(\epsilon))) = \frac{\lambda^p 2^{\frac{p}{\nu}-1}}{\Gamma\left(\frac{1}{\nu}\right)} \int_0^{+\infty} \sum_{k=0}^{+\infty} \left[\exp(b\alpha)\left(b\lambda 2^{\frac{1}{\nu}}(\gamma_2-\gamma_1)\right)^k + \left(b\lambda 2^{\frac{1}{\nu}}(\gamma_1+\gamma_2)\right)^k\right] \frac{y^{-1+\frac{1+p+k}{\nu}}\exp(-y)}{k!} dy. \quad \text{(A.15)}$$

Define $\Delta = max\left\{|b\lambda 2^{1/\nu}(\gamma_1+\gamma_2)|, max(\exp(b\alpha),1)|b\lambda 2^{1/\nu}(\gamma_2-\gamma_1)|\right\}$. Then, we can use the results in Nelson (1991) to show that if $\nu > 1$ then the summation and integration in (A.15) can be interchanged. Further, applying Formula 3.381 #4 of Ryzhik et al. (2007) yields the following required expression[1]

$$E(|\epsilon|^p \exp(bf(\epsilon))) = 2^{p/\nu}\lambda^p \sum_{k=0}^{\infty} (2^{1/\nu}\lambda b)^k \left[(\gamma_1+\gamma_2)^k + \exp(b\alpha)(\gamma_2-\gamma_1)^k\right] \frac{\Gamma((p+k+1)/\nu)}{2\Gamma(1/\nu)k!} < \infty. \quad \text{(A.16)}$$

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

$$E(\epsilon^p \exp(bf(\epsilon))) = 2^{p/\nu}\lambda^p \sum_{k=0}^{\infty} (2^{1/\nu}\lambda b)^k \left[(\gamma_1+\gamma_2)^k + (-1)^p\exp(b\alpha)(\gamma_2-\gamma_1)^k\right] \frac{\Gamma((p+k+1)/\nu)}{2\Gamma(1/\nu)k!} < \infty. \quad \text{(A.17)}$$

Note that the expectations (A.16) and (A.17) 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 (A.14) 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 $E(|\epsilon|^p \exp(bf(\epsilon)))$ are $b(\gamma_2-\gamma_1) < \frac{1}{2\lambda}$ and $b(\gamma_2+\gamma_1) < \frac{1}{2\lambda}$. That is

[1] See Nelson (1991) for the proof of finiteness of the formula.

$b(\gamma_2 - \gamma_1) < \sqrt{2}$ and $b(\gamma_2 + \gamma_1) < \sqrt{2}$. The conditions for the infinity of $E(\epsilon^p \exp(bf(\epsilon)))$ are the same as those for $E(|\epsilon|^p \exp(bf(\epsilon)))$ $0 < \nu \leq 1$.

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

$$\begin{aligned} & E(|\epsilon|^p \exp(bf(\epsilon))) \\ & = C_1 \left[\int_0^{+\infty} \epsilon^p \exp(b\alpha) \exp(b(\gamma_2 - \gamma_1)\epsilon) \left(1 + \frac{\epsilon^2}{d-2}\right)^{-\frac{d+1}{2}} d\epsilon \right. \\ & \qquad \left. + \int_0^{+\infty} \epsilon^p \exp(b(\gamma_1 + \gamma_2)\epsilon) \left(1 + \frac{\epsilon^2}{d-2}\right)^{-\frac{d+1}{2}} \right] d\epsilon, \end{aligned} \tag{A.18}$$

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 $b(\gamma_2 + \gamma_1) \leq 0$.

A.2.2 Expectations needed to compute $E(|y_t|^c)$, $corr(|y_t|^c, |y_{t+\tau}|^c)$ and $corr(y_t, |y_{t+\tau}|^c)$ when $\epsilon \sim N(0,1)$

Assume that all the parameters are defined as in equations (2.1) and (2.2). When $\epsilon \sim N(0,1)$, using the expression (A.14) 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

$$\begin{aligned} E(|\epsilon|^p \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \Bigg\{ & \exp(b\alpha)\Gamma(p+1) \exp\left(\frac{b^2(\gamma_1 - \gamma_2)^2}{4}\right) D_{-p-1}(b(\gamma_1 - \gamma_2)) \\ & + \Gamma(p+1) \exp\left(\frac{b^2(\gamma_1 + \gamma_2)^2}{4}\right) D_{-p-1}(-b(\gamma_1 + \gamma_2)) \Bigg\} \end{aligned} \tag{A.19}$$

and

$$\begin{aligned} E(\epsilon^p \exp(bf(\epsilon))) = \frac{1}{\sqrt{2\pi}} \Bigg\{ & (-1)^p \exp(b\alpha)\Gamma(p+1) \exp\left(\frac{b^2(\gamma_1 - \gamma_2)^2}{4}\right) D_{-p-1}(b(\gamma_1 - \gamma_2)) \\ & + \Gamma(p+1) \exp\left(\frac{b^2(\gamma_1 + \gamma_2)^2}{4}\right) D_{-p-1}(-b(\gamma_1 + \gamma_2)) \Bigg\}, \end{aligned} \tag{A.20}$$

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

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

and

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

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

Appendix B

Appendix to Chapter 3

B.1 Closed-form expressions of $E(\epsilon_t^c \exp(bf(\epsilon_t)))$ and $E(|\epsilon_t|^c \exp(bf(\epsilon_t)))$

B.1.1 $\epsilon_t \sim Normal$

Proposition B.1. *Let c be a non-negative integer and $b \in \mathbb{R}$ and ϵ_t and $f(\epsilon_t)$ are defined as in GAS^{2}V-N model. If $bk + |bk^*| < \frac{1}{2}$, then*

$$E\left(|\epsilon_t|^c \exp(bf(\epsilon_t))\right) = \frac{\exp(-bk)}{2\sqrt{2\pi}}\Gamma\left(\frac{c+1}{2}\right)\left[\exp(b\alpha)\left(\frac{1}{2}-b(k+k^*)\right)^{-\frac{c+1}{2}} + \left(\frac{1}{2}-b(k-k^*)\right)^{-\frac{c+1}{2}}\right] \tag{B.1}$$

and

$$E\left(\epsilon_t^c \exp(bf(\epsilon_t))\right) = \frac{\exp(-bk)}{2\sqrt{2\pi}}\Gamma\left(\frac{c+1}{2}\right)\left[(-1)^c\exp(b\alpha)\left(\frac{1}{2}-b(k+k^*)\right)^{-\frac{c+1}{2}} + \left(\frac{1}{2}-b(k-k^*)\right)^{-\frac{c+1}{2}}\right]. \tag{B.2}$$

Proof.

$$
\begin{aligned}
&E(|\epsilon_t|^c \exp(bf(\epsilon_t))) \\
&= \int_{-\infty}^{0} (-\epsilon_t)^c \exp(b\alpha + bk\epsilon_t^2 - bk + bk^*\epsilon_t^2) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\epsilon_t^2}{2}\right) d\epsilon_t \\
&+ \int_{0}^{\infty} (\epsilon_t)^c \exp(bk\epsilon_t^2 - bk - bk^*\epsilon_t^2) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\epsilon_t^2}{2}\right) d\epsilon_t
\end{aligned} \tag{B.3}
$$

Integrating by substitution with $s_t = -\epsilon_t$ in the finite integral, we obtain

$$
\begin{aligned}
&E(|\epsilon_t|^c \exp(bf(\epsilon_t))) \\
&= \frac{\exp(b(\alpha - k))}{\sqrt{2\pi}} \int_0^{\infty} (s_t)^c \exp((b(k + k^*) - \frac{1}{2})s_t^2) ds_t + \frac{\exp(-bk)}{\sqrt{2\pi}} \int_0^{\infty} (\epsilon_t)^c \exp((b(k - k^*) - \frac{1}{2})\epsilon_t^2) d\epsilon_t.
\end{aligned} \tag{B.4}
$$

According to the formula 3.326-2 of Ryzhik et al. (2007), when $c \geq 0$ and $bk + |bk^*| < \frac{1}{2}$, the former equation reduces to

$$
\begin{aligned}
E\left(|\epsilon_t|^c \exp(bf(\epsilon_t))\right) &= \frac{\exp(b(\alpha - k))}{\sqrt{2\pi}} \frac{\Gamma\left(\frac{c+1}{2}\right)}{2\left(\frac{1}{2} - b(k + k^*)\right)^{\frac{c+1}{2}}} + \frac{\exp(-bk)}{\sqrt{2\pi}} \frac{\Gamma\left(\frac{c+1}{2}\right)}{2\left(\frac{1}{2} - b(k - k^*)\right)^{\frac{c+1}{2}}} \\
&= \frac{\exp(-bk)}{2\sqrt{2\pi}} \Gamma\left(\frac{c+1}{2}\right) \left[\exp(b\alpha)\left(\frac{1}{2} - b(k + k^*)\right)^{-\frac{c+1}{2}} + \left(\frac{1}{2} - b(k - k^*)\right)^{-\frac{c+1}{2}}\right].
\end{aligned} \tag{B.5}
$$

Following the same steps, we can obtain the analytical expression of $E\left(\epsilon_t^c \exp(bf(\epsilon_t))\right)$ as follows:

$$
\begin{aligned}
&E(\epsilon_t^c \exp(bf(\epsilon_t))) \\
&= \int_{-\infty}^{0} \epsilon_t^c \exp(b\alpha + bk\epsilon_t^2 - bk + bk^*\epsilon_t^2) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\epsilon_t^2}{2}\right) d\epsilon_t \\
&+ \int_{0}^{\infty} (\epsilon_t)^c \exp(bk\epsilon_t^2 - bk - bk^*\epsilon_t^2) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\epsilon_t^2}{2}\right) d\epsilon_t \\
&= \frac{\exp(b(\alpha-k))}{\sqrt{2\pi}} \int_{0}^{\infty} (-s_t)^c \exp((b(k+k^*) - \frac{1}{2})s_t^2) ds_t \\
&+ \frac{\exp(-bk)}{\sqrt{2\pi}} \int_{0}^{\infty} (\epsilon_t)^c \exp((b(k-k^*) - \frac{1}{2})\epsilon_t^2) d\epsilon_t \\
&= \frac{\exp(b(\alpha-k))}{\sqrt{2\pi}} \frac{(-1)^c \Gamma\left(\frac{c+1}{2}\right)}{2\left(\frac{1}{2} - b(k+k^*)\right)^{\frac{c+1}{2}}} + \frac{\exp(-bk)}{\sqrt{2\pi}} \frac{\Gamma\left(\frac{c+1}{2}\right)}{2\left(\frac{1}{2} - b(k-k^*)\right)^{\frac{c+1}{2}}} \\
&= \frac{\exp(-bk)}{2\sqrt{2\pi}} \Gamma\left(\frac{c+1}{2}\right) \left[(-1)^c \exp(b\alpha) \left(\frac{1}{2} - b(k+k^*)\right)^{-\frac{c+1}{2}} + \left(\frac{1}{2} - b(k-k^*)\right)^{-\frac{c+1}{2}} \right]. \qquad \text{(B.6)}
\end{aligned}
$$

□

B.1.2 $\epsilon_t \sim t_\nu$

Proposition B.2. *Let c be a nonnegative integer and $b \in R$ and ϵ_t and $f(\epsilon_t)$ defined as in GAS^{2}V-T model, then when $\nu > c$*

$$
\begin{aligned}
E(|\epsilon_t|^c \exp(bf(\epsilon_t))) = \frac{(\nu-2)^{c/2} \exp(-bk)}{2} \frac{B(\frac{1+c}{2}, \frac{\nu-c}{2})}{B(\frac{1}{2}, \frac{\nu}{2})} \\
\cdot \left\{ \exp(b\alpha) \left[1 + \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \frac{c+1+2j}{\nu+1+2j} \right) \frac{(b(\nu+1)(k+k^*))^i}{i!} \right] \right. \\
\left. + \left[1 + \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \frac{c+1+2j}{\nu+1+2j} \right) \frac{(b(\nu+1)(k-k^*))^i}{i!} \right] \right\} \qquad \text{(B.7)}
\end{aligned}
$$

and

$$E(\epsilon_t^c \exp(bf(\epsilon_t))) = \frac{(\nu-2)^{c/2}\exp(-bk)}{2}\frac{B(\frac{1+c}{2},\frac{\nu-c}{2})}{B(\frac{1}{2},\frac{\nu}{2})} \cdot \left\{ (-1)^c \exp(b\alpha) \left[1 + \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \frac{c+1+2j}{\nu+1+2j} \right) \frac{(b(\nu+1)(k+k^*))^i}{i!} \right] + \left[1 + \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \frac{c+1+2j}{\nu+1+2j} \right) \frac{(b(\nu+1)(k-k^*))^i}{i!} \right] \right\}. \tag{B.8}$$

Proof. The probability density function of ϵ_t is $\psi_0(\epsilon_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\varphi_0\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}(1+\frac{\epsilon_t^2}{\nu\varphi_0^2})^{-\frac{\nu+1}{2}}$ where $\varphi_0 = \sqrt{\frac{\nu-2}{\nu}}$, then $u_t = (\nu+1)b_t - 1$ and $b_t = \frac{\epsilon_t^2/(\nu\varphi_0^2)}{1+\epsilon_t^2/(\nu\varphi_0^2)} \sim Beta(\frac{1}{2},\frac{\nu}{2})$, see Harvey (2013).

$$\begin{aligned} E(|\epsilon_t|^c \exp(bf(\epsilon_t))) &= \int_{-\infty}^{0} (-\epsilon_t)^c \exp\left(b(\alpha-k)\right) \exp\left(b(\nu+1)(k+k^*)b_t\right) \psi_0(\epsilon_t) d\epsilon_t \\ &+ \int_{0}^{+\infty} \epsilon_t^c \exp(-bk) \exp\left(b(\nu+1)(k-k^*)b_t\right) \psi_0(\epsilon_t) d\epsilon_t \\ &= \exp(b(\alpha-k)) \int_{-\infty}^{0} \epsilon_t^c \exp\left(b(\nu+1)(k+k^*)b_t\right) \psi_0(\epsilon_t) d\epsilon_t \\ &+ \exp(-bk) \int_{0}^{\infty} \epsilon_t^c \exp\left(b(\nu+1)(k-k^*)b_t\right) \psi_0(\epsilon_t) d\epsilon_t \\ &= \frac{\exp(b(\alpha-k))}{2} E\left(|\epsilon_t|^c \exp\left(b(\nu+1)(k+k^*)b_t\right)\right) \\ &+ \frac{\exp(-bk))}{2} E\left(|\epsilon_t|^c \exp\left(b(\nu+1)(k-k^*)b_t\right)\right), \end{aligned} \tag{B.9}$$

We proceed to work out the expectation $E(|\epsilon_t|^c \exp(mb_t))$ with respect to ϵ_t. Note that $E(|\epsilon_t|^c \exp(mb_t)) =$

$\varphi_0^c E(\nu^{c/2} b_t^{c/2}/(1-b_t)^{c/2} \exp(mb_t))$ with respect to $b_t \sim Beta(\frac{1}{2}, \frac{\nu}{2})$. It follows that

$$\begin{aligned}
&E(|\epsilon_t|^c \exp(m) b_t) \\
&= \varphi_0^c \int_0^1 \nu^{c/2} b_t^{c/2}/(1-b_t)^{c/2} \exp(mb_t) \frac{b_t^{\frac{1}{2}-1}(1-b_t)^{\frac{\nu}{2}-1}}{B(\frac{1}{2}, \frac{\nu}{2})} db_t \\
&= \varphi_0^c \nu^{c/2} \frac{B(\frac{1+c}{2}, \frac{\nu-c}{2})}{B(\frac{1}{2}, \frac{\nu}{2})} \int_0^1 \exp(mb_t) \frac{b_t^{\frac{1+c}{2}-1}(1-b_t)^{\frac{\nu-c}{2}-1}}{B(\frac{1+c}{2}, \frac{\nu-c}{2})} db_t \\
&= \varphi_0^c \nu^{c/2} \frac{B(\frac{1+c}{2}, \frac{\nu-c}{2})}{B(\frac{1}{2}, \frac{\nu}{2})} E(\exp(m\tilde{b}_t))
\end{aligned} \tag{B.10}$$

with the expectation taken with respect to a $Beta(\frac{1+c}{2}, \frac{\nu-c}{2})$ when $\nu > c$, which is the moment generating function of $\tilde{b}_t \sim Beta(\frac{1+c}{2}, \frac{\nu-c}{2})$. It yields that

$$E(|\epsilon_t|^c \exp(mb_t) = \varphi_0^c \nu^{c/2} \frac{Beta(\frac{1+c}{2}, \frac{\nu-c}{2})}{Beta(\frac{1}{2}, \frac{\nu}{2})} \left\{ 1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^{k-1} \frac{c+1+2r}{\nu+1+2r} \right) \frac{m^k}{k!} \right\}. \tag{B.11}$$

Combing equation (B.9) and (B.11) gives the expression. On the other hand,

$$\begin{aligned}
E(\epsilon_t^c \exp(bf(\epsilon_t))) &= \int_{-\infty}^{0} \epsilon_t^c \exp(b(\alpha-k)) \exp(b(\nu+1)(k+k^*)b_t) \psi_0(\epsilon_t) d\epsilon_t \\
&+ \int_0^{+\infty} \epsilon_t^c \exp(-bk) \exp(b(\nu+1)(k-k^*)b_t) \psi_0(\epsilon_t) d\epsilon_t \\
&= (-1)^c \exp(b(\alpha-k)) \int_{-\infty}^{0} \epsilon_t^c \exp(b(\nu+1)(k+k^*)b_t) \psi_0(\epsilon_t) d\epsilon_t \\
&+ \exp(-bk) \int_0^{\infty} \epsilon_t^c \exp(b(\nu+1)(k-k^*)b_t) \psi_0(\epsilon_t) d\epsilon_t \\
&= (-1)^c \frac{\exp(b(\alpha-k))}{2} E(|\epsilon_t|^c \exp(b(\nu+1)(k+k^*)b_t)) \\
&+ \frac{\exp(-bk))}{2} E(|\epsilon_t|^c \exp(b(\nu+1)(k-k^*)b_t))
\end{aligned} \tag{B.12}$$

The proof is completed. □

B.1.3 $\epsilon_t \sim GED(\nu)$

Proposition B.3. *Let c be a nonnegative integer and $b \in \mathbb{R}$. ϵ_t and $f(\epsilon_t)$ defined as in GAS^{2}V-G model. Then, when $bk + |bk^*| < 1/\nu$,*

$$E(|\epsilon_t|^c \exp(bf(\epsilon_t))) = \frac{\exp(b(\alpha-k))}{2} \frac{(\Gamma(1/\nu))^{c/2-1}\,\Gamma(\frac{c+1}{\nu})}{\left(\Gamma(\frac{3}{\nu})\right)^{c/2}} (1-\nu b(k+k^*))^{-\frac{c+1}{\nu}}, \\ + \frac{\exp(-bk)}{2} \frac{(\Gamma(1/\nu))^{c/2-1}\,\Gamma(\frac{c+1}{\nu})}{\left(\Gamma(\frac{3}{\nu})\right)^{c/2}} (1-\nu b(k-k^*))^{-\frac{c+1}{\nu}}, \tag{B.13}$$

and

$$E(\epsilon_t^c \exp(bf(\epsilon_t))) = \frac{(-1)^c \exp(b(\alpha-k))}{2} \frac{(\Gamma(1/\nu))^{c/2-1}\,\Gamma(\frac{c+1}{\nu})}{\left(\Gamma(\frac{3}{\nu})\right)^{c/2}} (1-\nu b(k+k^*))^{-\frac{c+1}{\nu}} \\ + \frac{\exp(-bk)}{2} \frac{(\Gamma(1/\nu))^{c/2-1}\,\Gamma(\frac{c+1}{\nu})}{\left(\Gamma(\frac{3}{\nu})\right)^{c/2}} (1-\nu b(k-k^*))^{-\frac{c+1}{\nu}}. \tag{B.14}$$

Proof.

$$\begin{aligned}
&E(|\epsilon_t|^c \exp(bf(\epsilon_t))) \\
&= \int_{-\infty}^{0} (-\epsilon_t)^c \exp(b\alpha + bku_t + bk^*(u_t+1))\psi(\epsilon_t)d\epsilon_t + \int_0^{+\infty} \epsilon_t^c \exp(bku_t - bk^*(u_t+1))\psi(\epsilon_t)d\epsilon_t \\
&= \exp(b(\alpha-k)) \int_0^{+\infty} \epsilon_t^c \exp(b(k+k^*)\frac{\nu}{2}g_t)\psi(\epsilon_t)d\epsilon_t + \exp(-bk) \int_0^{+\infty} \epsilon_t^c \exp(b(k-k^*)\frac{\nu}{2}g_t)\psi(\epsilon_t)d\epsilon_t \\
&= \frac{\exp(b(\alpha-k))}{2} E(|\epsilon_t|^c \exp(b(k+k^*)\frac{\nu}{2}g_t)) + \frac{\exp(-bk)}{2} E(|\epsilon_t|^c \exp(b(k-k^*)\frac{\nu}{2}g_t)) \\
&= \frac{\exp(b(\alpha-k))}{2} E(\varphi^c g_t^{\frac{c}{\nu}} \exp(b(k+k^*)\frac{\nu}{2}g_t)) + \frac{\exp(-bk)}{2} E(\varphi^c g_t^{\frac{c}{\nu}} \exp(b(k-k^*)\frac{\nu}{2}g_t)) \\
&= \frac{\varphi^c \exp(b(\alpha-k))}{2} E(g_t^{\frac{c}{\nu}} \exp(b(k+k^*)\frac{\nu}{2}g_t)) + \frac{\varphi^c \exp(-bk)}{2} E(g_t^{\frac{c}{\nu}} \exp(b(k-k^*)\frac{\nu}{2}g_t))
\end{aligned}$$

According to the Appendix B.2 of Harvey (2013), when $E(\exp(b(k+k^*)\frac{\nu}{2}g_t)) < \infty$ and $E(\exp(b(k-k^*)\frac{\nu}{2}g_t)) < \infty$, the previous equation can be written

$$\frac{\varphi^c \exp(b(\alpha-k))}{2}\frac{2^{\frac{c}{\nu}}\Gamma(\frac{c+1}{\nu})}{\Gamma(\frac{1}{\nu})}E\left(\exp\left(\frac{\nu b(k+k^*)}{2}\tilde{g}_t\right)\right) + \frac{\varphi^c \exp(-bk)}{2}\frac{2^{\frac{c}{\nu}}\Gamma(\frac{c+1}{\nu})}{\Gamma(\frac{1}{\nu})}E\left(\exp\left(\frac{\nu b(k-k^*)}{2}\tilde{g}_t\right)\right),$$

where $\tilde{g}_t \sim Gamma(2, \frac{c+1}{\nu})$. When $bk + |bk^*| < \frac{1}{\nu}$, both $E(\exp(b(k+k^*)\frac{\nu}{2}g_t))$ and $E(\exp(b(k-k^*)\frac{\nu}{2}g_t))$ are finite and given by the generating moments function of the Gamma distribution, then

$$\begin{aligned} E(|\epsilon_t|^c \exp(bf(\epsilon_t))) = &\frac{\exp(b(\alpha-k))}{2}\frac{(\Gamma(1/\nu))^{c/2-1}\Gamma(\frac{c+1}{\nu})}{\left(\Gamma(\frac{3}{\nu})\right)^{c/2}}\left(1-\nu b(k+k^*)\right)^{-\frac{c+1}{\nu}} \\ &+ \frac{\exp(-bk)}{2}\frac{(\Gamma(1/\nu))^{c/2-1}\Gamma(\frac{c+1}{\nu})}{\left(\Gamma(\frac{3}{\nu})\right)^{c/2}}\left(1-\nu b(k-k^*)\right)^{-\frac{c+1}{\nu}} \end{aligned}$$

The expression for $E(|\epsilon_t|^c \exp(bf(\epsilon_t)))$ can be obtained following the similar steps. □